WHAT IS THE DISTRIBUTIONAL BURDEN OF TAXING CONSUMPTION?

JOHN SABELHAUS*

To what extent would the distributional burden of taxes change if income were replaced by consumption as the primary base for collecting federal revenues? It is reasonable to assume that a consumption tax would be more regressive, because intuitive and empirical analyses suggest that the ratio of consumption to income falls as income rises. However, this paper shows that the data commonly used to evaluate consumption taxes probably overstate the extent to which the tax burden would change.

The focus of this paper is on the statistical basis for estimating tax burdens, and, hence, conceptual issues about how to measure the burden of a consumption tax take a back seat. This is not meant to imply that the theoretical issues are unimportant, as estimates of lifetime and annual burdens of consumption taxes are known to differ markedly.[1] The estimates in this paper are based on realized cash tax liability divided by realized cash income. This type of cash-basis burden measure is a standard approach and has the advantages of being easily estimated and interpretable.[2]

*Congressional Budget Office, Washington, D.C. 20515.

There are several variants of consumption tax systems in use around the world and being discussed in the United States.[3] This paper focuses on a proposal to allow the deduction of new net saving from gross income to arrive at taxable income. The resulting tax base (income less saving) is current consumption, and, hence, this method of collecting revenues is sometimes referred to as a "consumed-income" tax. Partial deduction of new net saving already exists in the United States, as employer contributions to pension funds, taxpayer contributions to some IRAs and 401k plans, and all interest earned on these retirement accounts are excluded from the tax base.

The consumed-income tax has a desirable property with respect to distributional burden, because a multiple bracket structure is maintained. In fact, given estimates of how saving varies with income, a set of brackets and rates can be found that maintains the distributional burden currently in place. The exact set of brackets and rates needed to maintain the current distributional burden depend on how saving varies with income. Thus, point estimates of saving rates by income class would play a crucial role in

any equity evaluation of a policy shift toward consumption taxes.

Two theoretically equivalent approaches can be used to estimate the joint distribution of income and saving using household-level data. First, expenditures and taxes can be subtracted from household income to create a residual measure of saving, similar to the approach in the National Income and Product Accounts (NIPA). A second approach is to measure net worth for a household at two points in time and then compute cash saving as the change in assets less the change in liabilities, adjusted for capital gains. Although these two measures are theoretically equivalent, actual estimates show very divergent results.

An example of how residual and net worth saving estimates diverge is found in Bosworth *et al.* (1991). They used the Survey of Consumer Finances (SCF) for 1983 and 1986 and found net worth saving rates ranging from −2.4 percent of income for the bottom income quintile up to 12.5 percent for the top three quintiles. In contrast, the same authors used data from the Consumer Expenditure Surveys (CEX) for 1982–1985 to create estimates of residual saving, and average saving varied from −92.1 percent of income in the bottom quintile to 25.8 percent in the top quintile. The SCF and CEX surveys show similar aggregate saving rates but yield very different conclusions about the distributional burden of a consumption tax.

This paper builds on the analysis described above but controls for possible differences in the surveys. Saving is measured using both the residual and net worth methods with one data set, the 1988–1989 CEX. The differences in the joint distribution of saving and income still exist when changes in net worth and residual measures are created for the same set of households. Stated differently, there are large statistical discrepancies in the survey data between residual and net worth saving measures.

No final solution to the discrepancy problem is offered in this paper. Evidence is presented that the net worth saving rates, which indicate that a consumption tax is less regressive than previously thought, are more reasonable than the residual measures commonly used. The evidence implies that the current statistical basis for evaluating the distributional burden of consumption tax proposals is at least very questionable. More research is needed on this topic, or policy proposals may be rejected for the wrong reasons.

RESIDUAL AND NET WORTH SAVING ESTIMATES

Cash-basis saving can be measured in two theoretically equivalent ways. The residual measure of saving is defined by

1

$$S^r = Y + R - T - I - C$$

where Y is cash income, R is private transfers and other money received, T is the amount of personal taxes paid, I is social insurance contributions, and C is consumption. An alternative to the residual saving construct is a net worth saving measure, given by

2

$$S^{nw} = \Delta A - \Delta L + \Delta H$$

where ΔA is the change in nonhousing assets, ΔL is the change in nonhousing liabilities, and ΔH is the change in housing net worth. Housing wealth is distinguished from nonhousing wealth because of the differential treatment in the existing tax code, *i.e.*, the fact that mortgage interest paid is excluded from the income tax base. If the components of each saving measure are conceptually consistent, then S^r and S^{nw} should be equal for any household. To the extent that they are not, define the statistical discrepancy as

3

$D = S^r - S^{nw}$.

If the statistical discrepancy is zero, the household's books balance. Another way to say this is that sources of funds ($Y + R + \Delta L$) are equal to the uses of funds ($T + I + C + \Delta A + \Delta H$).

The next section is devoted to estimates of residual and net worth saving across income groups. If the components of each saving measure are conceptually equivalent, then the statistical discrepancy should be zero for every household. Two possible problems with comparing residual and net worth based measures arise from the failure of surveys to be comprehensive and the treatment of capital gains. Some examples will help clarify exactly what is not measured in the survey data and when we can expect the alternative measures to be equivalent.

The survey data used in this paper do not include information about employer pension contributions or balances in pension funds. Therefore, it is impossible to include employer contributions or interest earned by pensions as part of household saving, as is done in the NIPA. In my notation, Y is missing employer contributions and imputed interest on pensions. This is consistent with S^r and S^{nw} being equal if we do not count the change in pension fund balance as part of ΔA, which we do not. The employer contributions and interest earned by pensions are certainly part of saving, but we cannot measure them.

A second example of what may be counted incorrectly in the survey is the purchase of consumer durables. If a household borrows money to buy a car, then residual saving (which has the purchase price of the car subtracted) and net worth saving (which has the value of the loan originated subtracted) will both be lower by the purchase price of the car. In economic terms, the household's residual saving should only fall by the "consumption flow" produced by the car during the survey year, and ΔA should include the undepreciated value of the car at year end. It can be argued that saving is mismeasured, but S^r and S^{nw} should be mismeasured by the same amount, that is, the statistical discrepancy should be zero.

These first two examples of what is not counted in the surveys and why the treatment should not affect the statistical discrepancy may give you the wrong idea about the goal of this paper. These transactions are treated incorrectly from the point of view of measuring saving, and the lack of information needed to do the accounting of the transactions is a problem. However, employer pension fund contributions and interest earned on pension assets are already excluded from the tax base. Also, no consumption-based tax proposal considers allowing subtraction of saving in the form of consumer durables saving from the tax base. Therefore, the improper treatment biases the saving rate estimates, but not conclusions about how the tax base would change.

Employer or government provided transfers for medical care or education are not covered by the survey either. This omission does not bias the estimates of saving levels, though it will affect saving rates. Saving levels are unaffected because the transfers are not counted as part of income or consumption.[4] If the level of saving is unaffected, then S^r and S^{nw} should still be equal. The saving rate is biased upward because the level of income, which should include the transfers, is lower.

Realized capital gains measured in the asset data may cause bias in the saving estimates but should not introduce a wedge between S^r and S^{nw}. As an example, assume a person sells $10,000 worth of stock during the survey, for which they had paid $5,000 in some year prior to the survey. There is a $5,000 capital gain

which is not counted in the survey income measure, so any saving estimate will be biased down by $5,000. The impact of the stock sale on S^r and S^{nw} depends on what the person does with the proceeds of the sale. If the person deposits the $10,000 proceeds in a saving account, there are two offsetting changes in ΔA, so S^r and S^{nw} are not affected. If all other saving were zero, both S^r and S^{nw} would still register a value of zero, though each should be $5,000. If the person buys a car with the proceeds, C rises by $10,000, so S^r falls by $10,000. The value of S^{nw} will also fall by $10,000, because the drop in one ΔA category is not offset by an increase in another. In this case, if all other saving is zero, true saving is −$5,000 but registered saving will be −$10,000. Again, this example is not meant to defend the survey measure of saving on economic grounds. The example just points out that both measures of saving should be biased in the same way.

Only one type of economic event I can think of could cause S^r and S^{nw} to be different. If a household reports a positive ΔA that is due to unrealized capital gains, there will be no offsetting transaction in the residual saving calculation. The only asset category for which this may cause a problem is stocks and bonds, where households are asked how much the market value of stocks and bonds they are holding has changed in the last year. The way this question is asked could potentially cause problems with the statistical discrepancy, but its effect is attenuated by the empirical observation that most households hold little, if any, stocks and bonds.

HOW DOES SAVING VARY ACROSS INCOME GROUPS?

Average residual and net worth saving in the CEX survey are reasonably close for the entire sample but differ markedly across income quintiles. Table 1 presents consumer-unit averages for several components of

TABLE 1
MEASURING SAVING USING THE RESIDUAL AND NET WORTH METHODS, CONSUMER EXPENDITURE SURVEY (1989)

	Consumer Unit Average
Cash income (Y)	$32,978
Wages and salaries	24,775
Self-employment income	2,302
Rent, interest, and dividends	1,473
Government transfers	2,787
Private transfers received	1,638
+ Other money receipts (R)	25
Support paid to other CUs (outflow)	−591
Insurance refunds	149
Inheritances and bequests	378
Food stamp benefits	87
− Personal taxes (T)	4,368
− Social insurance taxes (I)	2,095
− Consumption (C)	25,081
Food	4,143
Clothing	1,095
Rent and utilities	3,389
Out-of-pocket medical expenses	1,458
Motor vehicles and parts	2,633
Furniture and household equipment	777
Housing interest	1,799
Housing property taxes	586
Housing intermediate goods	615
Life insurance premiums paid	361
Gifts and contributions to organizations	462
Personal interest expense	571
Other goods	3,264
Other services	3,921
= Residual saving (S^r)	1,457
Increase in nonhousing assets (ΔA)	1,258
Private pension contributions	216
Deductible IRA/Keogh contributions	186
Other retirement contributions	137
Change in checking accounts	−17
Change in saving accounts	194
Change in stocks and bonds	247
Investments to own farm/business	294
− Increase in nonhousing liabilities (ΔL)	462
Vehicle loans originated	1,565

TABLE 1
(CONTINUED)

	Consumer Unit Average
Less: vehicle loan principal paid	1,163
Change in other installment credit	113
Less: change in other amount owed to CU	52
+ Change in housing net worth (ΔH)	−3
Properties purchased	1,246
Less: properties sold	506
Additions and alterations	436
Payment of mortgage principal	937
Less: mortgage loans originated	2,116
= Net worth saving (S^{nw})	792
Statistical discrepancy	665
Residual saving as a percent of cash income	4.4 percent
Net worth saving as a percent of cash income	2.4 percent
Discrepancy as a percent of cash income	2.0 percent
Sample size	5976

Source: Author's calculations using 1988–1989 CEXs.

the income, tax, consumption, and wealth change measures used to construct the saving estimates.[5] The bottom line ratios of residual and net worth saving to cash income are 4.4 and 2.4 percent for the entire sample. The two estimates, in turn, are similar to the NIPA rate computed using the residual method (adjusted for the omissions described in the last section), which was 2.6 percent in 1989.

The similarity in residual and net worth saving rates disappears when the sample is divided by income class. Table 2 shows saving rate estimates by income quintiles, and the differences between residual and net worth saving are striking.[6] Residual saving is −72.5 percent of cash income in the bottom quintile, but net worth saving is estimated at −6.7 percent of income. At the top of the income distribution, residual saving is estimated at 16.8 percent of cash income, while net worth saving is only 5.7 percent. The residual and net worth estimates imply very different changes in distributional burden if the tax base is shifted toward consumption. Which, if either, of these sets of estimates is correct?

The CEX data used to estimate saving by income quintile in Table 2 are often criticized on the grounds that the income information is poor. This criticism is consistent with the findings in Table 2. Any negative measurement error in the income data will be reflected in very low (*e.g.*, negative) residual saving rates. It is possible to evaluate this critique by comparing the CEX income data to another benchmark. Table 3 presents a comparison of the CEX cash income distribution to the CPS/SOI income distribution developed by the Congressional Budget Office.[7]

The distribution of population across quintiles in Table 3 is expected to be even, but the results show that the top two quintiles are underrepresented in the CEX. The details show that the top 1 percent of the income distribution is effectively missing or topcoded in the CEX, the next four percentiles of the population have only 82.5 percent (3.3 divided by the expected 4.0) represented, and the next five percentiles have only 92.0 percent (4.6 divided by the expected 5.0) represented. The bottom of the income distribution is consistent with CPS tabulations. I find 12.6 percent of the population living below the poverty level in the CEX data, whereas Census reports 12.8 percent.[8] The differences between the CEX and CPS/SOI income distribution are apparently concentrated at the top end and due to undersampling of the wealthy.

The CEX income data are consistent with the CPS/SOI up to the top percentiles of the distribution. This observation is due to the fact that the CEX and the CPS ask basically the same questions of basically the same people. The fact that CEX and CPS data are similar does not refute the point made above that Table 2 is consistent with income reporting errors in the data. The

TABLE 2
SAVING AND STATISTICAL DISCREPANCIES BY INCOME QUINTILES, CONSUMER EXPENDITURE SURVEY (1989)

	Income Quintile					
	First	Second	Third	Fourth	Fifth[a]	Total
Cash income	8,025	17,903	28,612	40,795	69,992	32,978
+ Other money receipts	327	−54	72	20	−224	25
− Personal taxes	92	806	2,618	5,009	13,382	4,368
− Social insurance taxes	331	995	1,880	2,825	4,492	2,095
− Consumption	13,745	18,258	23,654	29,585	40,434	25,081
= Residual saving	−5,815	−2,210	531	3,394	11,459	1,457
Increase in nonhousing assets	119	−42	1,012	1,922	3,354	1,258
− Increase in nonhousing liabilities	276	473	538	457	558	462
+ Change in housing net worth	−380	−75	−699	−86	1,194	−3
= Net worth saving	−538	−591	−226	1,378	3,990	792
Statistical discrepancy	−5,277	−1,619	757	2,016	7,468	665
Residual saving as a percent of cash income	−72.5	−12.3	1.9	8.3	16.8	4.4
Net worth saving as a percent of cash income	−6.7	−3.3	−0.8	3.4	5.7	2.4
Discrepancy as a percent of cash income	−65.8	−9.0	2.6	4.9	10.7	2.0
Sample size	1,077	1,238	1,145	1,214	1,302	5,976
Percent of population	19.7	21.8	20.8	19.4	18.3	100.0

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model.
[a]Sample is missing top 1 percent of the income distribution.

TABLE 3
DISTRIBUTION OF INCOME IN THE CONSUMER EXPENDITURE SURVEY (1989)

Quintiles	Percent of Population	Percent of Consumer Units and Unrelated Individuals	Sample Size	Percent with Topcoded Income
Lowest	20.6	21.6	1261	0.6
Second	21.8	21.2	1270	0.2
Third	21.0	20.3	1198	0.1
Fourth	18.8	18.1	1240	0.7
Fifth	17.7	18.8	1292	6.8
All	100.0	100.0	6261	1.6
Distribution tails				
Poverty level	12.6	13.0	750	0.9
91–95 percent	4.6	5.1	346	5.1
95–99 percent	3.3	3.8	270	19.3
Top 1 percent	0.1	0.2	10	86.3

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model. Sampling basis is consumer units and unrelated individuals. Quintile rankings are based on AFI.

similarity just adds an interesting dimension to this research, because it suggests that, if income reporting is a problem in the CEX, then the CPS might have the same problems.

The fact that the top 1 percent of the distribution is missing complicates the search for which piece(s) of the discrepancy puzzle are badly reported. We know that the sum of any income, consumption, or wealth variable in the CEX should not match the aggregate value of the variable from any other source, because the CEX

sample is not representative. The approach of comparing CEX totals by income, consumption, and wealth change component will have to wait until further information about the top 1 percent can be used to aggregate totals to derive truncated sums the CEX sample should match. Though aggregate sums cannot be compared at this point, a good deal more can be learned by looking at the CEX more closely and a Canadian survey which does not have the discrepancy problem.

SAVING PROFILES IN THE CANADIAN FAMILY EXPENDITURE SURVEY

Some evidence about why residual and net worth saving diverge across income groups can be found by looking at tabulations from another survey, the Canadian Family Expenditure Survey (FEX). The FEX is similar to the CEX in its basic mission to collect detailed expenditure information for developing a consumer price index. It differs in the way the survey is conducted and the emphasis that Statistics Canada places on "balancing" residual and net worth saving measures for a households in the sample.

The Canadian FEX survey is conducted using only one interview, and respondents are asked about income, expenditure, and wealth change during the previous 12 months. In contrast, the CEX interviewers visit respondents for each of 4 consecutive quarters and ask about the previous 3 months.[9] The FEX interviewers do a rough balancing calculation while at the respondent's home, and central office personnel call on a respondent if their calculations show large discrepancies between the sources and uses of funds. In contrast, each quarter of CEX data is processed independently, and the Bureau of Labor Statistics explicitly states that microlevel consistency is not a goal of the survey.

This alternative approach to collecting, processing, and reconciling the data leads to much less discrepancy between residual and net worth saving estimates across income groups, as the results in Table 4 indicate. The saving rate estimates in the FEX are close to the CEX residual measure at the top and close to the CEX net worth measure at the bottom. Unfortunately, this suggests multiple sources of error in the CEX saving estimates, *i.e.*, too little income at the bottom and too little positive asset change at the top. This comparison is a little ambiguous, however, because the CEX is missing the top of the income distribution.

The lack of discrepancy in the FEX is partially manufactured, however, as Statistics Canada procedures call for any households with large discrepancies (greater than 10 percent of sources or use of funds) to be dropped from tabulations and the public-use data. In 1986, about 5 percent of the FEX sample was excluded on this basis. It would be useful to look at these observations, but the point that the FEX is much cleaner can be made by noting that only about 40 percent of the CEX sample meets the FEX discrepancy criteria.

CAN THE DISCREPANCIES BE RESOLVED?

The Canadian FEX data show much less discrepancy between residual and net worth saving across income groups than the CEX data. Is there any systematic bias causing discrepancies in the CEX? The Canadian FEX data excludes households with extreme discrepancies, and one could take that approach with the CEX. Before that solution is invoked, however, it is interesting to consider subsamples of the population whose discrepancies are expected to be smaller.

One subset of the CEX sample is considered in Table 5. Any household that changed income quintile during the survey year is omitted.[10] The statistical discrepancy for the omitted group is still expected to be zero, but the subsetting exercise is

TABLE 4
SAVING AND STATISTICAL DISCREPANCIES BY INCOME QUINTILES, CANADIAN FAMILY EXPENDITURE SURVEY (1986)

	Income Quintile					
	First	Second	Third	Fourth	Fifth	Total
Cash income	11,886	23,812	33,543	44,384	67,290	35,661
+ Other money receipts	820	714	860	934	1,058	878
− Personal taxes	472	2,786	5,489	8,409	15,756	6,484
− Social insurance taxes	148	577	912	1,193	1,476	838
− Consumption	13,560	22,108	27,790	33,434	42,266	27,431
= Residual saving	−1,474	−945	211	2,281	8,849	1,785
Change in nonhousing net worth	−1,287	−1,385	−148	737	6,002	801
+ Change in housing net worth	12	691	645	1,716	2,592	1,109
= Net worth saving	−1,275	−694	497	2,453	8,594	1,910
Statistical discrepancy	−200	−250	−286	−174	253	−126
Residual saving as a percent of cash income	−12.4	−4.0	0.6	5.1	13.1	5.0
Net worth saving as a percent of cash income	−10.7	−2.9	1.5	5.5	12.8	5.4
Discrepancy as a percent of cash income	−1.7	−1.1	−0.9	−0.4	0.4	−0.4
Sample size	2,467	1,973	1,949	1,851	2,116	10,356
Percent of population	20.0	20.0	20.0	20.0	20.0	100.0
Percent of purchasing units	23.7	18.6	18.3	18.4	20.9	100.0

Source: Author's calculations using the 1986 FEX. Income distribution is based on AFI.

TABLE 5
SAVING AND STATISTICAL DISCREPANCIES BY INCOME QUINTILES, CONSUMER EXPENDITURE SURVEY (1989): SAMPLE EXCLUDES CONSUMER UNITS THAT CHANGED INCOME QUINTILE

	Income Quintile					
	First	Second	Third	Fourth	Fifth[a]	Total
Cash income	7,831	17,569	29,263	41,168	72,090	34,738
+ Other money receipts	651	−62	143	190	−796	−4
− Personal taxes	72	783	2,877	5,352	13,864	4,942
− Social insurance taxes	281	934	1,974	2,948	4,724	2,237
− Consumption	11,761	16,973	23,496	29,504	42,760	25,373
= Residual saving	−3,633	−1,185	1,058	3,553	9,943	2,180
Increase in nonhousing assets	−13	−90	1,153	1,210	4,085	1,377
− Increase in nonhousing liabilities	268	104	379	396	448	321
+ Change in housing net worth	−65	−392	211	88	1,087	225
= Net worth saving	−348	−588	986	902	4,725	1,280
Statistical discrepancy	−3,285	−597	72	2,650	5,217	899
Residual saving as a percent of cash income	−46.4	−6.7	3.6	8.6	13.8	6.3
Net worth saving as a percent of cash income	−4.4	−3.3	3.4	2.2	6.6	3.7
Discrepancy as a percent of cash income	−42.0	−3.4	0.2	6.4	7.2	2.6
Sample size	751	695	641	708	1001	3,796
Percent of population	22.5	19.3	19.0	17.8	21.5	100.0

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model.
[a]Sample is missing the top 1 percent of income distribution.

based on the idea that the survey may have trouble reconciling sources and uses for households with large income swings. The extreme differences between residual and net worth saving are still apparent in Table 5, however. Residual saving is somewhat less negative in the bottom quintile, but the discrepancy is by no means eliminated.

I have not included tables on several other subsetting exercises that were done and are worth mentioning. Households where members other than the head and spouse were working, where self-employment income was present, where the head was over a critical age, where properties were bought or sold, or where one or more imputed variables existed were all excluded as a test of whether systematic discrepancies were specific to certain groups. The reduction of statistical discrepancy in each case was less than that reported in Table 5.

Table 6 reports the results of applying the Canadian approach to the CEX data. Households are excluded from the sample if their statistical discrepancies are large. Because the CEX and FEX surveys are fundamentally different, I excluded households if the absolute value of their statistical discrepancy exceeded 25 percent of the greater of sources or uses of funds, rather than 10 percent, as in the FEX. Even this less restrictive criteria excludes 45 percent of the sample, a much larger fraction than the 5 percent dropped from the FEX.

The omission of high-discrepancy observations in the CEX dramatically flattens estimated residual saving profiles. Saving rates are estimated to range from −4.6 percent in the lowest quintile to 11.4 percent in the highest. These estimates do not differ much from the FEX estimates nor do they differ much from the Bosworth *et al.* (1991) estimates using the SCF.

Some concern about the saving estimates in Table 6 arises, because the remaining sample is clearly no longer representative. The fraction of population estimated to be in the bottom quintile falls dramatically to 12.5 percent. Accepting the saving rates in Table 6 as point estimates implicitly involves rejecting the CPS income distribution, because the CEX incomes match the CPS incomes and we are throwing out a disproportionate number of observations in the lowest quintile. Without stronger evidence about why we should exclude low-income households, I can only argue that the CEX residual saving rates estimated using a sample dominated by discrepancies should be treated with skepticism.

The last few tables are meant to illustrate an aspect of measuring saving in these data sets which researchers should keep in mind should they go searching for point estimates of saving rates. Table 7 shows the distribution of CEX and FEX observations by residual saving, net worth saving, and discrepancy, all relative to cash income. The most important point in Table 7 is that there is a great deal of inherent variability in these types of saving measures. The FEX, which forces sources and uses of funds to be equal, shows that 12.3 percent of the population saves less than −25 percent of their cash income, while 11.3 percent save more than +25 percent of their cash income. This occurs even though there is virtually no discrepancy in those ranges.

The CEX shows more dispersion than the FEX, as 23.0 percent of households save less than −25 percent of their cash income, while 23.5 percent save more than +25 percent. The dispersion is larger than in the FEX by a factor of about two in each tail. This is meant to illustrate that excluding people because their saving rates are "unreasonable" is not ultimately the solution to the discrepancy problem. The pure financial measure of saving, which, for example, treats durable expenditures as consumption, leads to highly dispersed saving rates.

TABLE 6
SAVING AND STATISTICAL DISCREPANCIES BY INCOME QUINTILES, CONSUMER EXPENDITURE SURVEY (1989): SAMPLE EXCLUDES CONSUMER UNITS WITH ABSOLUTE STATISTICAL DISCREPANCY EXCEEDING 25 PERCENT OF THE GREATER OF SOURCES OR USES OF FUNDS

	Income Quintile					
	First	Second	Third	Fourth	Fifth[a]	Total
Cash income	8,776	18,368	28,051	41,164	68,325	35,220
+ Other money receipts	642	61	−14	91	−840	−70
− Personal taxes	57	940	2,709	5,366	13,325	4,896
− Social insurance taxes	325	1,056	1,919	2,958	4,630	2,354
− Consumption	9,653	16,857	22,501	29,998	41,256	25,415
= Residual saving	−619	−423	906	2,933	8,272	2,483
Increase in nonhousing assets	−146	185	519	1,214	4,331	1,347
− Increase in nonhousing liabilities	234	503	520	354	258	386
+ Change in housing net worth	97	258	541	1,198	2,009	890
= Net worth saving	−282	−58	540	2,057	6,082	1,851
Statistical discrepancy	−336	−364	365	875	2,189	631
Residual saving as a percent of cash income	−7.1	−2.3	3.2	7.1	12.1	7.1
Net worth saving as a percent of cash income	−3.2	−0.3	1.9	5.0	8.9	5.3
Discrepancy as a percent of cash income	−3.8	−2.0	1.3	2.1	3.2	1.8
Sample size	420	693	736	812	812	3,473
Percent of population	12.7	22.0	22.5	23.0	19.8	100.0

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model.
[a]Sample is missing top 1 percent of the income distribution.

TABLE 7
DISTRIBUTION OF CONSUMER UNITS BY RESIDUAL SAVING, NET WORTH SAVING, AND DISCREPANCY, CONSUMER EXPENDITURE SURVEY (1989) AND CANADIAN FAMILY EXPENDITURE SURVEY (1986)

	Residual Saving		Net Worth Saving		Discrepancy	
Relative to Cash Income	CEX	FEX	CEX	FEX	CEX	FEX
Less than −100 percent	7.8	1.4	3.1	1.3	8.7	0.1
−100 to −50 percent	6.8	3.5	3.7	2.9	7.0	0
−50 to −25 percent	9.7	7.4	5.0	6.5	10.0	0.2
−25 to −10 percent	10.6	13.4	8.0	10.9	12.0	14.1
−10 to 0 percent	10.9	18.3	13.6	17.1	10.3	37.4
0 to 10 percent	13.4	22.6	38.0	29.3	12.9	41.4
10 to 25 percent	20.2	22.1	17.1	22.5	19.1	6.5
25 to 50 percent	18.0	9.7	6.6	7.9	13.4	0.1
50 to 100 percent	2.2	1.1	2.8	1.1	4.0	0
More than 100 percent	0.4	0.5	2.1	0.5	2.6	0

Source: Author's calculations using 1988–1989 CEXs and the 1986 FEX.

Tables 8A through 8C and 9A through 9C further refine the observations about Table 7. The residual, net worth, and discrepancy distributions for the FEX and CEX are presented by quintiles. The Table 8 group, for example, shows how much the bottom quintile is dominated by discrepancy. Table 8A shows that 54.4 percent of the bottom quintile is in a residual saving class of less than −25 percent of income, but Table 8C shows that this is because 56.5 percent of the bottom quintile is in a discrepancy class of less than −25 percent of income. Table 9A shows, by contrast, that only 19.6 percent of the FEX bottom quintile saves less than −25 percent of their income on a residual basis.

TABLE 8A
DISTRIBUTION OF CONSUMER UNITS BY RESIDUAL SAVING CLASS AND INCOME QUINTILE, CONSUMER EXPENDITURE SURVEY (1989)

Saving Relative to Cash Income	Income Quintile					
	First	Second	Third	Fourth	Fifth[a]	Total
Less than −100 percent	31.5	5.5	1.6	0.6	0.3	7.8
−100 to −50 percent	13.7	10.7	5.0	3.3	1.1	6.8
−50 to −25 percent	14.0	13.4	10.5	7.3	3.0	9.7
−25 to −10 percent	11.1	13.2	10.8	9.8	7.7	10.6
−10 to 0 percent	7.0	13.3	12.7	11.0	10.4	10.9
0 to 10 percent	5.7	13.9	17.0	15.0	15.3	13.4
10 to 25 percent	9.5	15.1	21.7	28.5	26.7	20.2
25 to 50 percent	6.3	13.6	18.4	21.4	30.4	18.0
50 to 100 percent	1.1	1.3	1.8	2.8	4.2	2.2
More than 100 percent	0.2	0.1	0.6	0.5	0.8	0.4

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model.
[a]Sample is missing top 1 percent of the income distribution.

TABLE 8B
DISTRIBUTION OF CONSUMER UNITS BY NET WORTH SAVING CLASS AND INCOME QUINTILE, CONSUMER EXPENDITURE SURVEY (1989)

Saving Relative to Cash Income	Income Quintile					
	First	Second	Third	Fourth	Fifth[a]	Total
Less than −100 percent	6.8	2.6	2.2	1.5	2.2	3.1
−100 to −50 percent	3.1	6.7	3.8	1.9	2.5	3.7
−50 to −25 percent	3.9	5.1	6.7	5.1	4.4	5.0
−25 to −10 percent	5.1	7.4	8.5	10.1	9.2	8.0
−10 to 0 percent	11.3	14.6	13.9	15.9	12.2	13.6
0 to 10 percent	47.7	38.7	35.4	33.2	34.8	38.0
10 to 25 percent	9.3	15.5	17.8	22.5	20.8	17.1
25 to 50 percent	5.0	5.3	7.7	6.4	8.6	6.6
50 to 100 percent	3.8	2.6	2.5	2.2	3.0	2.8
More than 100 percent	3.9	1.6	1.7	1.2	2.2	2.1

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model.
[a]Sample is missing top 1 percent of the income distribution.

TABLE 8C
DISTRIBUTION OF CONSUMER UNITS BY STATISTICAL DISCREPANCY CLASS AND INCOME QUINTILE, CONSUMER EXPENDITURE SURVEY (1989)

Discrepancy Relative to Cash Income	Income Quintile					
	First	Second	Third	Fourth	Fifth[a]	Total
Less than −100 percent	29.8	6.5	3.3	2.0	2.5	8.7
−100 to −50 percent	12.9	9.9	5.5	3.9	2.5	7.0
−50 to −25 percent	15.1	12.3	9.4	7.1	5.8	10.0
−25 to −10 percent	9.3	13.6	13.5	12.6	11.0	12.0
−10 to 0 percent	4.8	12.8	11.5	12.3	10.1	10.3
0 to 10 percent	6.9	12.1	13.9	16.3	15.4	12.9
10 to 25 percent	9.2	14.4	22.6	24.9	24.6	19.1
25 to 50 percent	6.0	11.9	14.2	15.7	19.2	13.4
50 to 100 percent	3.4	4.1	3.4	3.4	5.6	4.0
More than 100 percent	2.5	2.4	1.9	1.9	3.5	2.6

Source: Author's calculations using 1988–1989 CEXs. Income distribution is based on AFI classes from the CBO tax simulation model.
[a]Sample is missing top 1 percent of the income distribution.

TABLE 9A
DISTRIBUTION OF CONSUMER UNITS BY RESIDUAL SAVING CLASS AND INCOME QUINTILE, CANADIAN FAMILY EXPENDITURE SURVEY (1986)

Saving Relative to Cash Income	Income Quintile First	Second	Third	Fourth	Fifth	Total
Less than −100 percent	4.6	0.8	0.5	0	0.2	1.4
−100 to −50 percent	6.1	5.8	3.3	1.4	0.7	3.5
−50 to −25 percent	8.9	10.0	8.5	6.0	3.6	7.4
−25 to −10 percent	17.9	16.3	12.4	12.7	7.2	13.4
−10 to 0 percent	23.9	19.2	19.0	15.8	12.6	18.3
0 to 10 percent	21.0	23.0	23.9	25.1	20.6	22.6
10 to 25 percent	12.1	16.6	22.4	26.3	34.7	22.1
25 to 50 percent	4.2	7.4	8.6	11.5	17.5	9.7
50 to 100 percent	0.7	0.7	0.8	0.8	2.4	1.1
More than 100 percent	0.8	0.2	0.5	0.4	0.4	0.5

Source: Author's calculations using 1986 FEXs. Income distribution is based on AFI.

TABLE 9B
DISTRIBUTION OF CONSUMER UNITS BY NET WORTH SAVING CLASS AND INCOME QUINTILE, CANADIAN FAMILY EXPENDITURE SURVEY (1986)

Saving Relative to Cash Income	Income Quintile First	Second	Third	Fourth	Fifth	Total
Less than −100 percent	4.3	0.7	0.4	0	0.4	1.3
−100 to −50 percent	5.4	4.2	2.7	1.4	0.5	2.9
−50 to −25 percent	7.5	9.3	7.0	5.6	3.2	6.5
−25 to −10 percent	11.8	13.0	11.5	11.4	6.8	10.9
−10 to 0 percent	23.5	18.0	16.8	14.0	12.0	17.1
0 to 10 percent	33.7	31.4	29.1	28.0	23.7	29.3
10 to 25 percent	9.3	15.7	23.8	29.2	36.4	22.5
25 to 50 percent	3.2	6.6	7.3	9.2	14.0	7.9
50 to 100 percent	0.5	0.7	0.8	0.9	2.5	1.1
More than 100 percent	0.8	0.2	0.6	0.3	0.5	0.5

Source: Author's calculations using 1986 FEXs. Income distribution is based on AFI.

TABLE 9C
DISTRIBUTION OF CONSUMER UNITS BY STATISTICAL DISCREPANCY CLASS AND INCOME QUINTILE, CANADIAN FAMILY EXPENDITURE SURVEY (1989)

Discrepancy Relative to Cash Income	Income Quintile First	Second	Third	Fourth	Fifth	Total
Less than −100 percent	0.5	0	0	0	0	0.1
−100 to −50 percent	0.1	0	0	0	0	0
−50 to −25 percent	0.8	0.1	0.1	0	0	0.2
−25 to −10 percent	18.9	16.2	13.9	11.8	9.1	14.1
−10 to 0 percent	35.5	36.5	39.4	39.6	36.8	37.4
0 to 10 percent	36.6	40.4	40.3	43.0	47.6	41.4
10 to 25 percent	6.9	6.8	6.3	5.7	6.6	6.5
25 to 50 percent	0.6	0	0	0	0	0.1
50 to 100 percent	0.1	0	0	0	0	0
More than 100 percent	0	0	0	0	0	0

Source: Author's calculations using 1986 FEXs. Income distribution is based on AFI.

Conclusions

This project set out to estimate the joint distribution of consumption and income for the purpose of assessing the distributional burden of taxes and found that the estimates are dominated by unresolved statistical discrepancy. However, the circumstantial evidence suggests that analysis of consumption taxation based on residual

measures of saving should be treated with skepticism. This critique applies to any studies of excise, VAT, or sales tax burdens which rely on an estimated distribution of consumption and income. It is reasonable to infer that existing studies using the residual method to compute saving are biased toward determining that consumption taxes are more regressive than what is probably the case.

ENDNOTES

I am grateful to Frank Sammartino for help with the CBO tax model data and to seminar participants at the Congressional Budget Office for useful comments and ideas. Opinions here do not necessarily reflect those of the Congressional Budget Office.

[1] See Davies *et al.* (1984), Fullerton and Rogers (1991), and Poterba (1989).

[2] For examples of tax burden estimates, see Pechman (1985), Kasten *et al.* (1993), and Congressional Budget Office (1987, 1988, 1992a, 1992c).

[3] For a discussion of the various types of consumption taxes, see, for example, Gravelle (1992), Aaron and Galper (1985), and Bradford (1986).

[4] Medical expenditures, for example, are measured as consumption only if they are "out of pocket." For a discussion of how these omitted transactions affect the NIPA saving measure, see the Appendix to Bosworth *et al.* (1991).

[5] Consumer units are the basic sampling unit in the CEX. A consumer unit is similar to the basic Census sampling units of households but adjusted for whether individuals share resources. In 1989, there were 93.4 million households and 95.8 million consumer units. See the Appendix for a discussion of how the CEX data are used in this paper.

[6] The allocation of households across quintiles is based on Adjusted Family Income (AFI), which is an equivalence-scale adjusted measure of income. The quintile breaks are from the Congressional Budget Office tax simulation data base, which combines Current Population Survey (CPS) and Statistics of Income (SOI) information into one consistent data base. The CPS/SOI data set is the basis for CBO studies of tax burden (1987, 1988) and changes in distribution income over time (1992b).

[7] The sampling basis for the CEX has to be adjusted somewhat to make this comparison. When the Census and Congressional Budget Office estimate income distribution, individuals living in households, who are not related to the household head, are treated as separate units. I split out unrelated individuals in order to create Table 3, which is why the population distribution does not match that in Table 2.

[8] This finding that poverty rates in the CEX are similar to poverty rates in the CPS implies that some critical analysis of U.S. poverty estimates may be possible. In particular, the percent of people living under the income poverty line and also having consumption below the poverty line (which is not justified by net worth change, such as borrowing) is only 5.6%. This estimate should be treated cautiously until a more detailed comparison of the CPS poor and CEX poor can be undertaken.

[9] The survey only covers certain topics during certain interviews. For example, income is asked about at the first and last interview and financial wealth change is asked about at only the last interview. The Appendix has some details about how this distribution of questions affects the sample selection process.

[10] Respondents are asked at the first and fourth interviews about income during the prior 12 months. The change in income quintile is based on the change between the first and fourth interviews.

REFERENCES

Aaron, Henry J. and Harvey Galper. *Assessing Tax Reform.* Washington, D.C.: The Brookings Institute, 1985.

Bradford, David F. *Untangling the Income Tax.* Cambridge: Harvard University Press, 1986.

Bosworth, Barry, Gary Burtless, and John Sabelhaus. "The Decline in Saving: Evidence from Household Surveys." *Brookings Papers on Economic Activity (1)* (1991).

Congressional Budget Office. *The Changing Distribution of Federal Taxes: 1975–1990.* Washington, D.C.: U.S. Government Printing Office, October, 1987.

Congressional Budget Office. *The Changing Distribution of Federal Taxes: A Closer Look at 1980.* Washington, D.C.: U.S. Government Printing Office, July, 1988.

Congressional Budget Office. *Effects of Adopting a Value Added Tax.* Washington, D.C.: U.S. Government Printing Office, February, 1992a.

Congressional Budget Office. "Measuring the Distribution of Income Gains." Staff Memorandum. Washington, D.C.: U.S. Government Printing Office, March, 1992b.

Congressional Budget Office. "Distributional Effects of Substituting a Flat-Rate Income Tax and a Value Added Tax for Current Federal Income, Payroll, and Excise Taxes." Staff Memorandum. Washington, D.C.: U.S. Government Printing Office, April, 1992c.

Congressional Budget Office. "New Tax Proposals: Flat, VAT, and Variations." Congressional Research Service Report for Congress No. 92-386 S. Washington, D.C.: U.S. Government Printing Office, April, 1992.

Davies, James, France St. Hilaire, and John Whalley. "Some Calculations of Lifetime Tax Incidence." *American Economic Review 74,* No. 4 (September, 1984).

Fullerton, Don and Diane Lim Rogers. "Lifetime Versus Annual Perspectives on Tax Incidence." *National Tax Journal XLIV* (No. 3, 1991).

Kasten, Richard, Frank Sammartino, and Eric Toder. "Trends in Federal Tax Progressivity, 1980–93." In *Tax Pro-*

gressivity and Income Inequality, edited by Joel Slemrod. New York, Cambridge University Press, 1993, in press.

Pechman, Joseph A. *Who Paid the Taxes, 1966–85.* Washington, D.C.: Brookings Institution, 1987.

Poterba, James M. "Lifetime Incidence and the Distributional Burden of Excise Taxes." *American Economic Review 79,* No. 2 (May, 1989).

APPENDIX

The CEX data used in this paper are based on matches of quarterly information from the 1988–1990 public-release interview data sets. Any consumer unit who was interviewed about expenditures or income in at least part of 1989 is included in the sample. For example, a consumer unit whose interviews began in July, 1988, answered questions about expenditures during their four interviews for the periods April–June, 1988, July–September, 1988, October–December, 1988, and January–March, 1989. This consumer unit is part of the 1989 sample but with a weight about one-fourth the weight of a consumer unit whose interviews covered the calendar year 1989. Thus, there are 5,976 consumer units in my sample, and they have from 1 to 12 months of (weighted) information pertaining to 1989.

The four quarterly CEX interviews for any given consumer unit vary in content, and a complete sources and uses of funds accounting requires that the consumer unit remain in the survey for all four quarters. The Bureau of Labor Statistics views the survey as quarterly cross sections, and any respondent who moves or refuses to participate as the survey progresses is simply dropped from the sample, without any attention given to attrition bias. This makes sense from the viewpoint of quarterly representation but biases any matched data set which requires four quarters of participation.

A second potential bias in the data arises, because some consumer units do not give a full accounting of their income, either because they do not know or because they refuse. The CEX designates people who answer income questions as "complete-income reporters," and this is the subset used in this paper. The potential bias arises because complete-income reporters may not be representative of the whole sample.

The basic CEX weights are adjusted to reflect these factors. The sample of complete-income reporters in the sample for a full year is compared to the complete CEX sample, and weights are adjusted by age of household head and housing tenure (owner vs. renter) to reverse high rates of attrition among young renters. Average consumption in the complete-income full-year sample with unadjusted weights differs substantially from the complete CEX sample, because the homeowning older people who dominate the restricted sample spend more and spend differently. The weight adjustment results in total consumption estimates which are closer to those in the complete CEX sample, particularly for items such as rent payments.

All items except personal income taxes paid by the consumer unit are tabulated from the raw data. Income taxes are estimated because the CEX does not collect the tax data with much precision. Inspection of the data indicates that estimated taxes are based on extrapolating the amount of tax withheld from a recent paycheck stub. This causes particular problems in households with nonwage income. The tax estimates used in this paper have been evaluated relative to other sources and are much closer to the distribution of actual taxes collected.