

THE MISMEASURE OF MAN' S WELL-BEING: Refining Realized Income Measures with Wealth, Portfolio, and Mortality Information

Barry Johnson, Statistics of Income Division, Internal Revenue Service
Jenny Wahl, Economics Department, Carleton College

Economists and policymakers often rely on realized income to gauge individual well-being. Attractive for its ease of calculation, this measure is nonetheless seriously flawed, in part because people have some ability to choose how much income to realize at a given time. Income from capital is particularly susceptible to manipulation. In this paper, we build upon the path-breaking work of Projector and Weiss (1969) to ascertain the degree of mismatch between realized income and wealth and to suggest ways to construct better indices of wellbeing.

Empirical evidence from a recently compiled Internal Revenue Service data set shows just how imperfect realized income can be as a measure of economic well-being. Linked federal estate and income tax returns reveal that wealthy individuals – particularly those in their prime working years -- realize very low returns on capital. What is more, less-wealthy retirees tend to realize larger returns than more-wealthy retirees.

Our data also allow us to impute wealth on the basis of realized income, portfolio allocation, and other important factors. We offer here some initial results that suggest how one might use income data to predict an individual's wealth.

EXISTING RESEARCH

The research most closely associated with ours is a set of studies conducted by C. Eugene Steuerle (1983, 1985). Steuerle used a database containing wealth and income

data from federal tax returns to examine the relationship between realized income and the underlying wealth that generates at least a portion of that income.

Steuerle's work serves as a partial blueprint for ours, although our data are much more extensive. Not only are Steuerle's samples smaller and more restricted than ours, they fail to contain weights that reflect the probability of a match between estate and income tax records. Nor did Steuerle's data have weights to conform the decedent population more closely to the living population.

Other research has informed ours as well. The years since Steuerle's work have witnessed the advent of data sets such as the Federal Reserve Board of Governor's triennial Survey of Consumer Finance (SCF), the University of Michigan's Panel Study of Income Dynamics, and the Census Bureau's Survey of Income and Program Participation.

One result of this inquiry is the creation of measures of well-being that blend components of income and annuitized values of certain assets (see for example Ringen 1988, Radner 1990, Wolff et al. 2004). Much of this work — particularly in the federal government — centers upon measures of poverty (see for example Bauman 1999, Mishra et al. 2002). The appropriate treatment of wealth in poverty indices remains a significant source of debate, however (see Short and Ruggles 2004).

Recent work (especially Kennickell 1999, 2001) explores the possibility of modeling the relationship between wealth and income for the very wealthy as well. The work we describe here suggests that estate tax data may prove particularly useful in this effort.

DATA

U.S. federal tax records offer a bountiful source of intergenerational data on income and wealth. For our ongoing research, the Statistics of Income Division of the Internal Revenue Service has selected a set of federal estate tax returns and matched it to income tax returns filed by the decedent in the year before death, to gift tax returns filed during the decedent's lifetime, and to income tax returns filed by beneficiaries. These data are referred to collectively as the Estate Collation Study. The core data for this paper come from a stratified sample of federal estate tax returns (form 706) filed in 1992 and 1993 for people who died in 1992 and left estates of at least \$600,000. Matched to the estate tax returns are income tax returns (form 1040) filed by the decedent in 1991. The final Estate Collation sample consists of returns for 3,767 decedents.

Before analyzing the data, we needed to account for certain factors. Estate tax returns did not have equal probabilities of being matched to income tax returns due to primarily to late filing and errors in the secondary Social Security numbers on returns of joint filers. We therefore generated weights for the sample that reflect the unequal probabilities of a successful match. The first step was to create an adjustment factor to balance to the original population totals, essentially treating unmatched records as non-respondents. We then used auxiliary data, post-stratification, and raking to adjust the sampling weights and compare them to estimates from other sources of administrative data. Johnson and McMahon (2002) describe this process in greater detail.

A second feature of the data that requires adjustment has to do with differences in traits between the living population and the decedents. The 1992 estate tax population consisted of some 60,000 individuals with gross estates of at least \$600,000, the estate-

tax filing threshold in effect at that time. These decedents represented less than 1 percent of the U.S. population in 1992, according to the U.S. Census Bureau, and accounted for 2.8 percent of all 1992 deaths. While female decedents comprised 51.2 percent of the total U.S. resident population in 1992, female decedents made up only 43.5 percent of the 1992 estate tax population. The majority of male decedents -- 65.8 percent -- were married, as compared to 56.8 percent in the general population. Most female estate-tax decedents -- 61.5 percent -- were widowed, much higher than the 11.2 percent observed for their living counterparts in 1992. More than 87 percent of male decedents were 60 years old or older, while 14.4 percent of the living male population was in that age group in 1992. Likewise, 94.5 percent of female estate tax decedents were age 60 or older, while just 18.9 percent of living women were in that age group in 1992. Eller et al. (1992) contains a more complete description of the 1992-estate-tax population.

These statistics highlight one of the potential deficiencies of using data from estate tax returns to study the living population. As Smith (1985) points out, estate-tax data provide an excellent means of making statements about the deceased, but do not of themselves allow inferences about the living population. To compensate for the age bias and produce estimates more representative of the living population, we re-weight the file using reciprocals of mortality rates (by age and sex), adjusted by a differential that reflects the lower mortality rates experienced by the wealthy. Richer people tend to live longer because they enjoy access to better healthcare, safer occupations, and superior nutrition. Johnson and Woodburn (1994) provide a full discussion of weight adjustments.

Another potential limitation of the Estate Collation file concerns married decedents. While the estate tax return should contain complete information on the

decedent's portfolio, many, perhaps most, married decedents filed income tax returns jointly with a surviving spouse. Yet we do not directly capture the assets of the surviving spouse for the purpose of calculating returns to capital. We make a partial adjustment for this by including the full value of any property owned jointly by the decedent and surviving spouse in our asset base, including all community property and property owned as tenants-in-common. But we still miss the value of assets owned solely by the surviving spouse. While we have experimented with imputing values for these assets, we make no adjustment for them here. In some of our analysis, we do try to account for possible differences between married and other decedents – for example, by including dummy variables in various regressions.

One further data concern: the reporting of certain assets on federal estate tax returns is idiosyncratic. For example, the full face value of life insurance is included in the decedent's total gross estate for tax purposes. In addition, the tax code allows certain adjustments in asset value, such as the special valuation of real estate used for farming or certain business purposes. Where possible, we modify the data to compensate for these reporting anomalies. In the case of life insurance, for instance, we impute an equity value using data from the 1992 SCF.

Table 1 shows income and assets by source for our matched data, using weights appropriate for the estate-tax population. It also reports estimates generated using weights and asset values adjusted to represent the living population. As might be expected, the share of net income subject to tax attributable to salaries and wages is substantially higher for the living population than for those in the decedent group. Likewise, income from businesses is much higher for the living group. The portfolio

estimates for these two groups reveal differences as well. The share of the portfolio pertaining to business assets is higher in estimates for the living population than that for decedents, as is the share made up of real estate. The proportion of investments in tax-exempt bonds is significantly higher in the estimates for the decedent group than for those in the living population. A comparison of the adjusted data set with estimates from the 1992 SCF indicates that the adjusted estate values are consistent with patterns seen in the 1992 living population.

But other data issues remain. Certainly, many decedents must have been aware that they were close to death, so their portfolios could differ from those of the general population. Decedents (and their executors) naturally had incentives to report the lowest possible legal values for tax purposes on both income and estate tax returns. We believe that the high audit rate for estate-tax returns ensures that evasion is relatively rare, although informal transfers of small items such as jewelry surely take place. In addition, the truncation of the distribution due to the estate-tax filing threshold means that we must be cautious in generalizing from these data to populations other than the relatively wealthy. Finally, limitations due to the timing and retention of IRS masterfile data (the source of income-tax data for this study) mean that the only income data available are for income earned during calendar year 1991. Some income-producing assets could have been sold in 1992 prior to a decedent's death, but we cannot track that transaction. What is more, Kennickell (2001), among others, has suggested that pooling several years of data smoothes out year-to-year fluctuations in income caused by events such as change in employment status, receipt of inheritances, or realization of capital gains. Our one year of income data could therefore contain substantial transitory components. Despite these

flaws, we think our data are more promising than other micro-level data sources for realized property income because they do not suffer from the amount of underreporting and item non-response present in most survey data.

EMPIRICAL FINDINGS

Figure 1 shows the distributions of income and wealth across net-worth deciles for individuals in the Estate Collation sample. The graph shows that, as wealth increases, income also increases. However, the rate of growth for income is significantly less than that of wealth, as evidenced by the flatter slope of the income line. This means that the realized rate of return on assets actually tends to decrease as individuals accumulate more wealth.

Calculating Realized Rates of Return on Capital

Realized rates of return differ from actual economic rates of return by the amount of unrealized income or other income from capital not reported on a tax return. For our sample, the mean return on capital for all ages is 4.6 percent, with the mean return for those aged 50 to 59 at 2.9 percent. These rates are lower than those associated with a reasonably risk-free, low-paying savings account. For instance, six-month CDs generated an average 5.91 percent return during 1991. What is more, our measure for realized return to capital is likely biased upward because some included income items overstate the actual return.

Economic theory suggests that higher-risk, less-liquid assets generate higher economic rates of return. We think it implausible that these wealthy and presumably

investment-savvy decedents would have been satisfied with the relatively low rates they realized on complex portfolios of stock, bonds, real estate, partnerships, and the like. Instead, we believe that the pattern of realized returns offers evidence of careful tax planning, which became more important the more the individual had at stake. Indeed, the fact that people appear in high-wealth categories suggests that these individuals were successful in generating both a high economic return and a low realized return (and thus low taxes).

Table 2 shows estimated average realized rates of return on capital across different classes of wealth for decedents of various ages as well as for the living population. For those aged 70 or older, the table reveals that those with estates of \$10 million or more realized lower returns than those with estates less than \$1 million. Also notable is that individuals in their prime work-years tend to realize lower returns on capital than retirees. This finding reinforces our planning argument: if people earn taxable labor income, they may wish to realize relatively less capital income than those who are not working.

Rate-of Return Regression Analysis

Regression analysis might allow us to say more about the influence of one's portfolio upon realized rates of return on capital and on stock. Consistent with Steuerle, we find that realized rates of return varied inversely with the value of the particular asset in question, holding other relevant variables constant.

Yet modeling rates of return from estate and income tax data is fraught with problems. Income generated from various assets that could appear on an estate tax return

can be categorized in many different ways, for instance. Consequently, we do not draw conclusions from this rate-of-return analysis.

Estimating Wealth from Components of Gross Income

Rather than refine the rate-of-return analysis, we construct a model that predicts wealth from components of realized income and adjustments to income. A very simple model poses total assets as a function of various types of income reported on the 1040, along with the value of interest deductions and an index for the importance of deductions and adjustments to income. Table 3 reveals these results. For the overall sample, total assets are an increasing function of age in the relevant range. Nearly all income components have a positive relationship to total assets, with the largest coefficient associated with dividend income. The regression weighted to the living population suggests that an extra dollar of dividend income implies an increment to total assets of \$83. In turn, this result gives us a point estimate of only 1.2 percent for the rate of return on assets that yield dividends. The coefficient on taxable interest implies a higher estimated yield of 10.7 percent. The negative coefficient on tax-exempt interest seems odd; as we shall see, however, all but the highest-income people generate a positive relationship between tax-exempt interest and asset value. A similar result holds for capital gains and other income.

Deductions from income as well as income components plausibly might relate to the amount of assets held. The regression results indicate, for instance, that higher interest deductions are associated with more total assets. This result makes sense, given that the deductions probably act as a proxy for the value of real estate. One other variable

of note is “propadj.” This variable indicates the proportion of gross income subject to tax that is made up of adjustments and deductions such as those pertaining to Keogh plans, IRAs, and SEPs. Higher deductions can thus be associated with the building up of assets. Consequently, we might expect that, for a given level of gross income, people with a higher “propadj” would have higher total assets; indeed, the regression coefficient on “propadj” is positive.

Segmenting our data helps us craft even better predictions of total assets. Total assets are closely related to the amount of reported dividends. The relationship is even stronger for dividend amounts above a threshold of about \$2000. Total assets are also correlated positively – though less strongly -- with other income components such as taxable and tax-exempt interest.

These relationships suggest categorizing decedents on the basis of dividends received, with special treatment for those who realized very small amounts. Table 4 reports the results of regressions by dividend class. These results indicate that non-corporate and tax-exempt interest income are more important in predicting total assets for people reporting relatively small amounts of dividend income, whereas dividend income matters more in predicting total assets for those receiving larger amounts of dividend income.

The predictions from the set of regressions reported in Table 4 appear quite promising, because they yield estimates of well-being that are much more closely related to total assets than are income measures. Significantly, the Pearson coefficient relating the predicted value to actual total assets is .79, whereas the coefficient for taxable income is .22 and for gross income is .48. Predicted values from regressions pertaining to

decendents indicate that these values are also better than income measures at ranking observations. The Spearman coefficient for the predicted value is .70, as compared to a coefficient of .56 for taxable income and .67 for gross income. The Spearman rank results are not as clear for the regressions pertaining to the living population. Because the way this index is constructed can yield a lower value when rank shifts are more frequent but relatively more minor, however, we give more weight to the Pearson results. What they suggest is that we may have found a useful technique to gross up income components to yield a predicted value of well-being for wealthy people.

CONCLUSION

The evidence shown here indicates that income from capital is in many ways a voluntary event. Realized property income can vary dramatically across wealth and age classes, most likely reflecting tax considerations rather than differences in true economic returns. Indeed, wage income may be doing the same, particularly for executives who earn substantial amounts of non-wage compensation that receive preferential tax treatment. Income alone is no longer a reasonable way to assess individual well-being.

Our research suggests that merged estate and income tax records offer an effective way to gauge individual well-being among the wealthiest portion of the population. In short, we show how various components of income and deductions associated with capital assets can be combined to yield a predicted value for total assets that is highly correlated with actual assets. Applying our methods may therefore help reduce the degree of mismeasurement in man's well-being.

SOURCES

- Bauman, Kurt. Extended Measures of Well-being: Living Conditions in the United States:1998. *Current Population Reports*. Washington, DC, March 2003.
- Eller, Martha Britton; Erard, Brian; and Chih-Chin Ho. Federal Taxation of Wealth Transfers, 1992-1995. *Statistics of Income Bulletin* (Winter 1996-1997): 8-63.
- Johnson, Barry W. and Paul B. McMahon. Using Auxiliary Information to Adjust for Nonresponse in Weighting a Linked Sample of Administrative Records. *Statistics of Income Bulletin* (Winter 2002).
- Johnson, Barry W. and Louise Woodburn. The Estate Multiplier Technique, Recent Improvements for 1989. In Barry Johnson. ed. *Compendium of Federal Estate Tax and Personal Wealth Studies*. Washington, DC: Department of the Treasury, Internal Revenue Service, Pub. 1773, 1994.
- Kennickell, Arthur. Using Income Data to Predict Wealth. Paper presented at the American Economics Association meetings (January 1999). Also found at <http://www.federalreserve.gov/pubs/oss/oss2/papers/wealth.income.7.web.pdf>.
- Kennickell, Arthur. Modeling Wealth with Multiple Observations of Income: Redesign of the Sample for the 2001 Survey of Consumer Finances. (October 2001). <http://www.federalreserve.gov/pubs/oss/oss2/papers/scf2001.list.sample.redesign.9.pdf>.
- Mishra, Ashok K., Hisham S. El-Osta, Mitchell J. Morehart, James D. Johnson, and Jeffrey W. Hopkins. Income, Wealth, and the Economic Well-being of Farm Households. *Agricultural Economics Report No. 812*. Washington, DC: U.S. Department of Agriculture Farm Sector Performance and Well-being Branch,

Resource Economics Division, Economic Research Service, July 2002.

Projector, Dorothy S. and Weiss, Gertrude S. Income-Net Worth Measures of Economic Welfare. *Social Security Bulletin* 32 (November 1969).

Radner, Daniel B. Alternate Estimates of Economic Well-being by Age Using Data on Wealth and Income. *Social Security Administration Office of Policy Office of Research and Statistics Working Paper Number 42* (March 1990).

Ringen, Stein. Direct and Indirect Measures of Poverty. *Journal of Social Policy* 17 (1988): 351-65.

Short, Kathleen and Patricia Ruggles. Accounting for Net Worth in Measuring Poverty. *Proceedings of the 28th General Conference of The International Association for Research in Income and Wealth* (2004). www.iriw.org.

Smith, James. Comment. In Martin David and Timothy Smeeding, eds. *Horizontal Equity, Uncertainty, and Economic Well-being*. Chicago: University of Chicago Press, 1985.

Steuerle, C. Eugene. Relationship Between Realized Income and Wealth, a Report from a Selected Sample of Estates containing Farms or Business. *Statistics of Income Bulletin* (Spring 1983).

_____. Wealth, Realized Income, and the Measure of Well-being. In Martin David and Timothy Smeeding, eds; *Horizontal Equity, Uncertainty, and Economic Well-being*. Chicago, University of Chicago Press, 1985.

Wolff, Edward N., Ajit Zacharias, and Asena Caner. Levy Institute Measure of Economic Well-being: United States, 1989, 1995, 2000 and 2001. Annandale-on-Hudson, N.W.: The Levy Economics Institute, 2004.

FIGURE 1

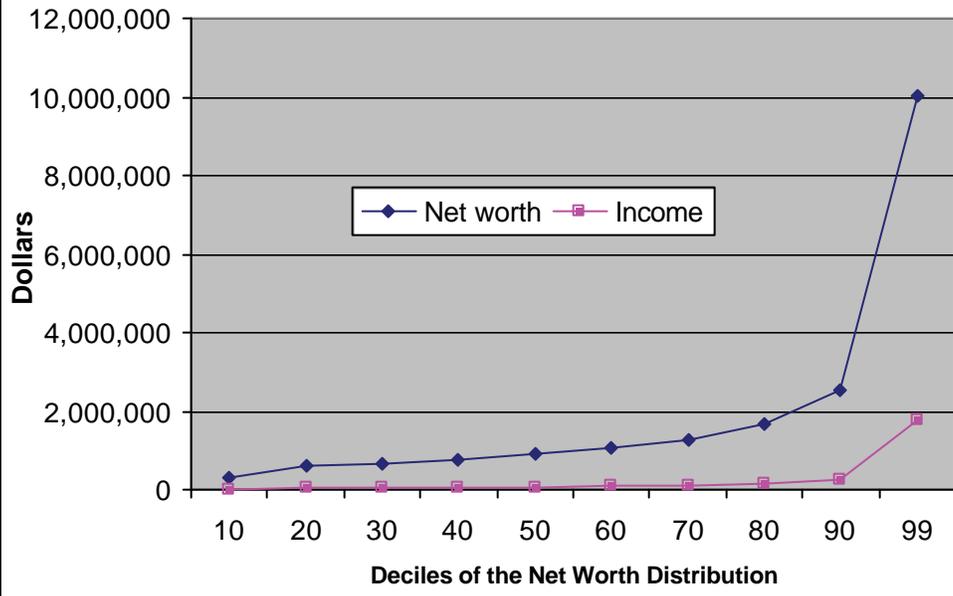


Table 1: Income and Wealth by Source	Estate Tax Decedent Population Estimate			Estimate for Living Population		
	Average amount of item	Item as a percentage of Net income subject to tax	Net Worth	Average amount of item	Item as a percentage of Net income subject to tax	Net Worth
Income by source						
Wages and salaries	16,702	15.74	0.99	62,781	45.61	4.40
Dividends	20,209	19.04	1.20	12,195	8.86	0.86
Taxable interest	28,910	27.24	1.72	22,003	15.99	1.54
Tax-exempt interest	16,591	15.63	0.99	10,069	7.32	0.71
State tax refunds	831	0.78	0.05	1,046	0.76	0.07
Alimony	73	0.07	0.00	324	0.24	0.02
Schedule C income or loss	2,278	2.15	0.14	5,777	4.20	0.41
Net capital gain or loss	13,239	12.47	0.79	15,233	11.07	1.07
Capital gain distribution	65	0.06	0.00	42	0.03	0.00
Supplemental gain or loss	97	0.09	0.01	(1,035)	-0.75	-0.07
Pension & annuity income	7,095	6.68	0.42	8,138	5.91	0.57
Partnership and S-corp income	14,565	13.72	0.87	29,729	21.60	2.08
Rents, royalties, REMIC	6,567	6.19	0.39	8,474	6.16	0.59
Estate and trust income	1,110	1.05	0.07	770	0.56	0.05
Farm income	162	0.15	0.01	(403)	-0.29	-0.03
Reported other income, Social Security, unemployment comp.	4,779	4.50	0.28	248	0.18	0.02
Gross income subject to tax	109,093	102.79	6.49	146,351	106.33	10.26
Total adjustments	657	0.62	0.04	1,757	1.28	0.12
Net income subject to tax	106,135	100.00	6.31	137,633	100.00	9.65
Exemptions	2,946	2.78	0.18	3,813	2.77	0.27
Interest deduction	3,153	2.97	0.19	9,334	6.78	0.65
Other deductions (Standard Ded. or Itemized less mortgage int.)	22,692	21.38	1.35	20,194	14.67	1.42
1040 taxable income	96,045	90.49	5.71	127,247	92.45	8.92
	Average amount of item	Item as a percentage of total wealth		Average amount of item	Item as a percentage of total wealth	
Wealth by source						
Stock	421,610	23.86		254,487	15.81	
Closely held stock	150,026	8.49		199,561	12.40	
Personal residence	141,503	8.01		180,781	11.23	
Real estate	228,478	12.93		285,483	17.73	
Tax-exempt bonds	217,058	12.28		128,061	7.95	
Cash, bonds, notes and mortgages	362,225	20.50		250,740	15.57	
Noncorporate assets	46,305	2.62		69,086	4.29	
Other assets	202,546	11.46		241,424	15.00	
Total wealth	1,766,938	100.00		1,609,940	100.00	15
Debts	86,234	4.88		183,727	11.41	

Rates of Return to Capital, Estimates for the Living Population, 1992					
	\$600,000 under \$1 million	\$1million under \$5 million	\$5 million under \$10 million	\$10 million or more	All
Age					
Under 50	4.1	4.1	3.1	4.6	4.0
50 under 60	2.6	2.9	3.1	5.5	2.9
60 under 70	4.6	5.0	4.3	4.8	4.8
70 and older	8.4	5.1	6.0	5.7	6.9
All ages	5.0	4.2	4.2	5.2	4.6

	\$600,000 under \$1 million	\$1million under \$5 million	\$5 million under \$10 million	\$10 million or more	All
Age					
Under 50	2.6	3.3	2.8	7.3	3.0
50 under 60	3.2	2.6	3.7	5.0	2.8
60 under 70	4.4	5.8	5.1	4.8	5.2
70 and older	7.4	5.4	5.5	5.5	6.5
All ages	6.7	5.2	5.2	5.4	5.9

	Coefficients living populations	Coefficients decedent population	Means	S.D.
total assets			7542341	29912054
married	123755	-266952	0.52	0.5
age	**71893	90987	70.96	16.42
agesqrd	**-.699	-.795		
propadj	**728681	**559004	0.93	0.21
ncorpinc	**1.77	**2.01	46916	570688
divinc	**83.35	**67.51	93373	379337
farminc	-0.35	-0.16	-3042	216106
intinc	**9.37	**13.22	90982	354940
teintinc	**-.5.51	**-.2.39	77143	372086
cginc	-0.61	-1.32	71858	419944
penaninc	2.06	1.52	12620	87503
othinc	**-.70	**-.1.57	2019	367147
nonkinc	**1.20	**5.81	72378	215246
intded	**7.63	**8.94	16541	174756
1/weight	**-.1733084	-.2386298		
Adj Rsqr	0.73	0.68		
N obs	3767	3767		

