

Session Number: Second Poster Session
Time: Thursday, August 26, PM

*Paper Prepared for the 30th General Conference of
The International Association for Research in Income and Wealth*

St. Gallen, Switzerland, August 22-28, 2010

**AN ANALYSIS OF ANNUAL HOUSEHOLD INCOME FROM THE
SURVEY OF CONSUMER FINANCES**

Laurence G. Taff

For additional information please contact:

Name: Laurence G. Taff
Affiliation: Taff & No Associates

Full mailing address: 648 Mistland Trail, Ruckersville, VA 22968-6000, U.S.A.

Email addresses: larrytaff@embarqmail.com

This paper is posted on the following website: <http://www.iariw.org>

**AN ANALYSIS OF ANNUAL HOUSEHOLD INCOME FROM THE
SURVEY OF CONSUMER FINANCES**

Laurence G. Taff
Taff & No Associates
648 Mistland Trail
Ruckersville, VA 22968-6000
U.S.A.
larrytaff@embarqmail.com
434 985 8154

ABSTRACT

This paper presents a brief analysis of the frequency and cumulative distributions of Americans' annual household income. In addition, a thorough discussion of the correlations between average annual household income and some of the key demographic variables is included. The data comes from the 2004 Survey of Consumer Finances. One significant advantage of the Survey is its simultaneous, same sample, measure of income and total financial assets. Most interesting, utilizing this data a new, extraordinarily close, power law function, relationship between *current* annual household income and *accumulated* financial assets is demonstrated. Another startling result is that the dependence of mean annual household income on portfolio composition has nearly perfect power law function dependence too. The discovery of the latter was facilitated via a re-interpretation of the Herfindahl Index in the asset mixture context.

1. BACKGROUND

The Survey of Consumer Finances is a triennial effort between the Board of Governors of the Federal Reserve and the Statistics of Income Division of the Internal Revenue Service. It is the best survey of many financial and demographic variables of the U.S.A.'s populace---especially total financial assets, portfolio composition, and annual household income. The Federal Reserve's principal publications regarding the Surveys include some comparisons between the current and earlier editions (e.g., [Bucks, et al. 2006 and 2009]). There are also many other papers by Federal Reserve personnel dealing with technical aspects of the Survey's canvassing methodology.

However, neither Federal Reserve personnel, nor others I could find publishing in the academic or professional literature, have thoroughly discussed the average annual household income distribution using the Survey's first-rate data set. Neither have they made any interesting connections between annual household income and total financial assets or portfolio composition. This paper contains a traditional analysis of the mean annual household income distribution and displays the tight linkages between average annual household income and several important demographic parameters. Most importantly, and most revealingly, is also demonstrated mean annual household income's close relationship with total financial assets on the one hand and asset mixture on the other.

To be more specific, employing the information in the 2004 edition of the Survey¹, this paper investigates the dependence of

average annual household income on the age of the head of household, on the level of education of the head of household, on (pre-sub-prime housing mess) home value, on the total stock of a household's wealth, and on a household's division of funds among investment possibilities. In order to perform the latter task a new quantity, the Portfolio Composition Parameter, had to be invented. It is a root of a quadratic function of the dollar allocations in a financial portfolio. Both total assets and asset allocation have a power law function dependence on annual household income. As a power law function has no scale---that is it is scale invariant whereas money certainly is not---this makes no sense to me. Moreover, a logarithmic function works almost equally well as a trial fitting form deepening this mystery because the logarithm is definitely not scale invariant. Lastly, because of the remarkably good fit to the data with a power law function, this measure provides an especially accurate insight into a household's investment risk avoidance/acceptance tendencies and its correlation with mean household annual income.

¹The 2007 edition of the Survey of Consumer Finances was released as this paper was being prepared. A quick comparison between it and the 2004 edition shows no significant differences for this research. Moreover, the 2007 version is slightly contaminated as the authors' tried to include the early effects of the sub-prime housing debacle.

2. THE DISTRIBUTION OF ANNUAL HOUSEHOLD INCOME

The Frequency Distribution And Basic Statistics

The average values of annual household income from the 2004 Survey of Consumer Finances were placed into 1-percentage point population bins. The resulting plot is shown in Figure 1. The first few moments of the distribution are:

Mean = \$67,100, Median = \$41,700,
Standard Deviation = \$121,000, and Skewness = 0.21.

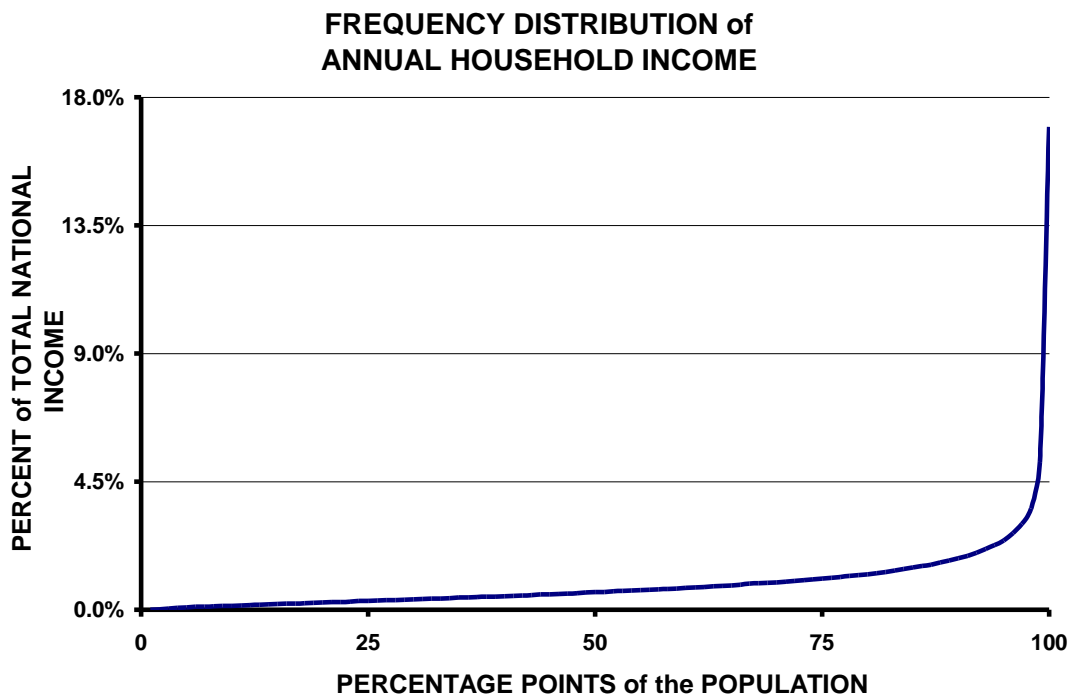


Figure 1. Frequency Distribution of Annual Household Income.

The standard deviation about the mean clearly shows the wide range of average annual household income values. As is common with frequency distributions with a wide range, the average is biased much higher than the median.

The value of skewness above is Pearson's Median Skewness Coefficient. It is defined as:

Pearson's Median Skewness Coefficient

$$= 3 \times (\text{Mean} - \text{Median})/\text{Standard Deviation}.$$

The reason it was utilized is that this quantity is much more informative regarding the extremely asymmetric shape in Figure 1 than is the modern definition of the skewness. (The latter's value is 7.3 in this circumstance.)

The Cumulative Distribution Of Annual Household Income

One way to examine the cumulative distribution of average annual household income is by population sub-sets. Chart 1 shows the decile distribution. The annual household income curve is so steep that the ninth and tenth highest deciles can not be meaningfully separated.

The cumulative distribution of mean annual household income itself is shown in the following plot; Figure 2. Given the sharp turn up at about the 95'th percentile visible in Figure 1, the shape of this

graph should not be surprising. Finally, Chart 2 portrays the same distribution in a different fashion again using population deciles.

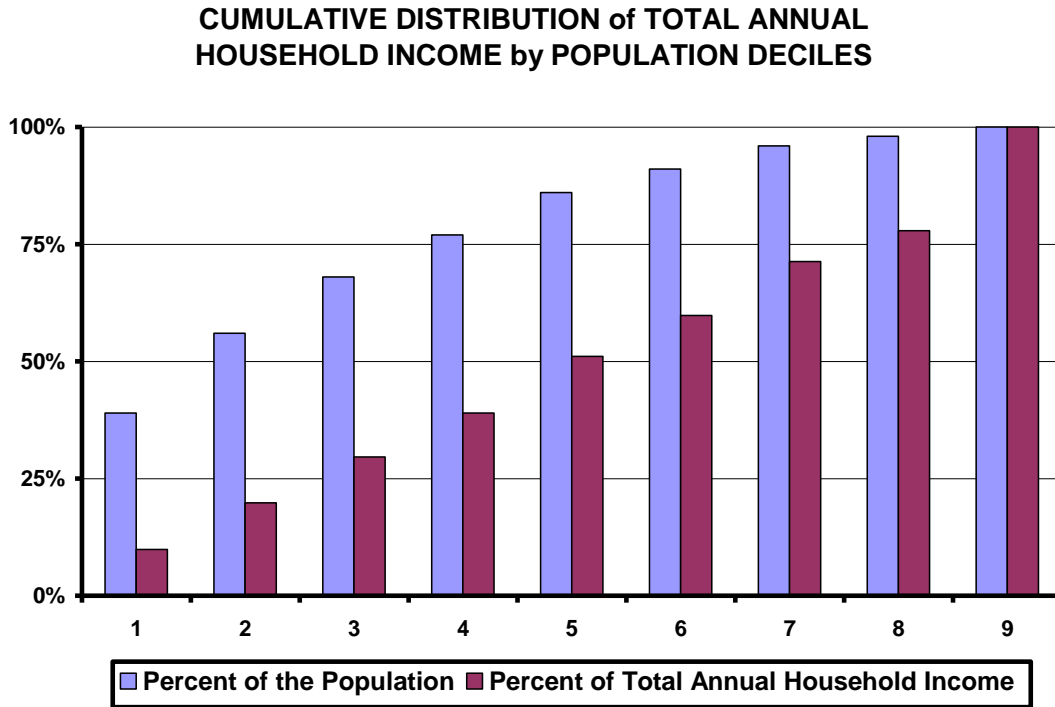


Chart 1. Cumulative Distribution of Annual Household Income by Population Deciles.

CUMULATIVE ANNUAL HOUSEHOLD INCOME DISTRIBUTION

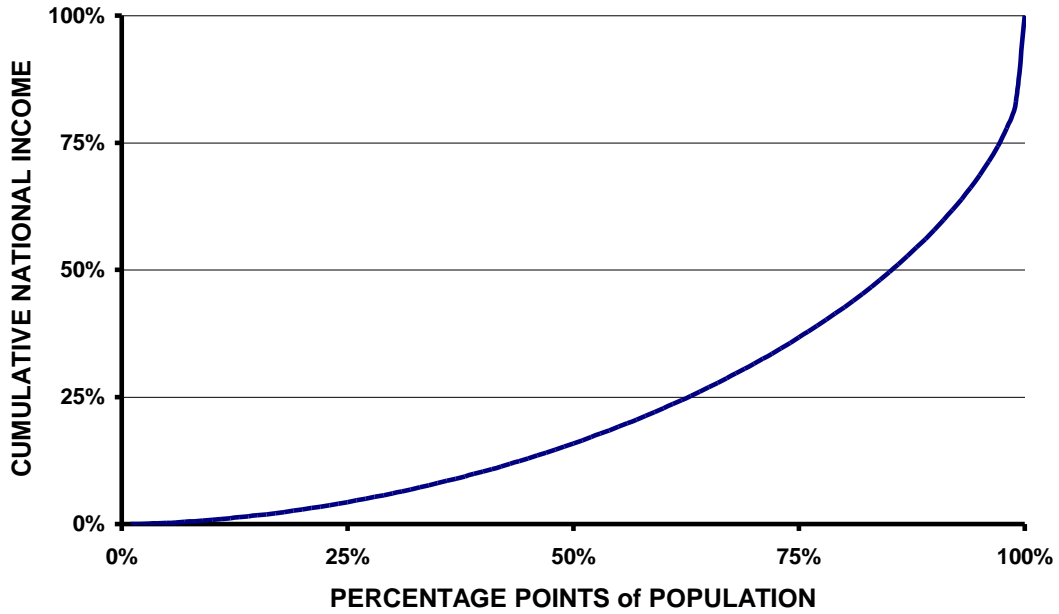


Figure 2. Cumulative Annual Household Income Distribution.

CUMULATIVE DISTRIBUTION of ANNUAL HOUSEHOLD INCOME by DECILES

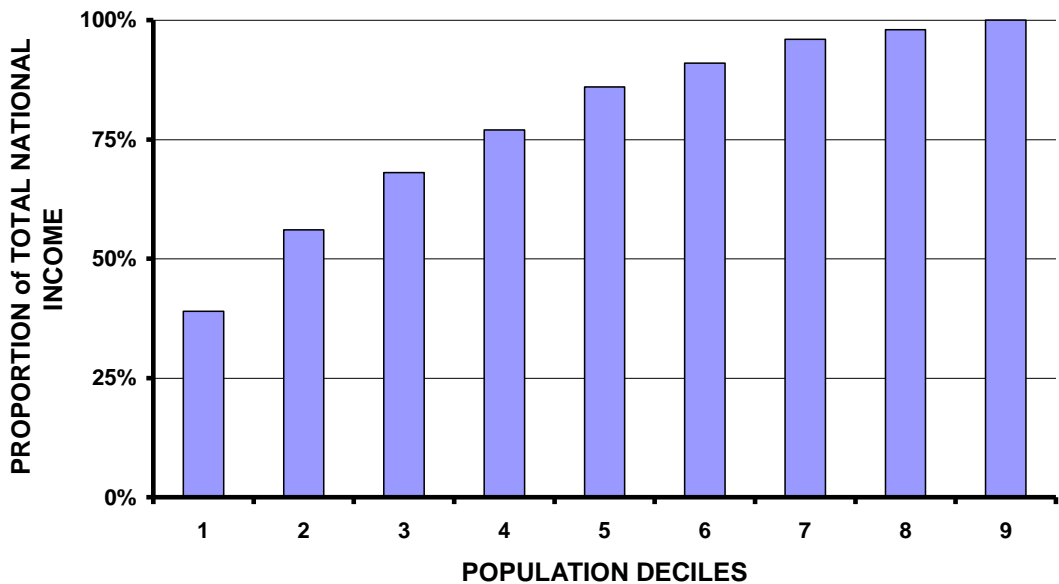


Chart 2. Cumulative Distribution of Annual Household Income by Population Deciles.

The Lorenz Curve and the Gini Coefficient

The Lorenz curve of the average annual household income distribution of Americans is presented in Figure 3. The straight line representing an egalitarian distribution is shown for comparison.

That there is an unequal distribution of annual household income was already clear from the frequency and cumulative distributions and the value of the Pearson Median Skewness Coefficient. This is also reflected in both the large bowing out of the Lorenz Curve from the 45° line egalitarian situation and the value of the Gini Coefficient (0.46). (The Gini Coefficient was computed using the trapezoidal rule and 100 equal population bins.) While the United States is frequently cited as having a more un-equal (in the numerical sense) income distribution than, for example, western Europe, countries run by dictators or authoritarian rulers---and poor countries--have exceptionally large Gini Coefficients.

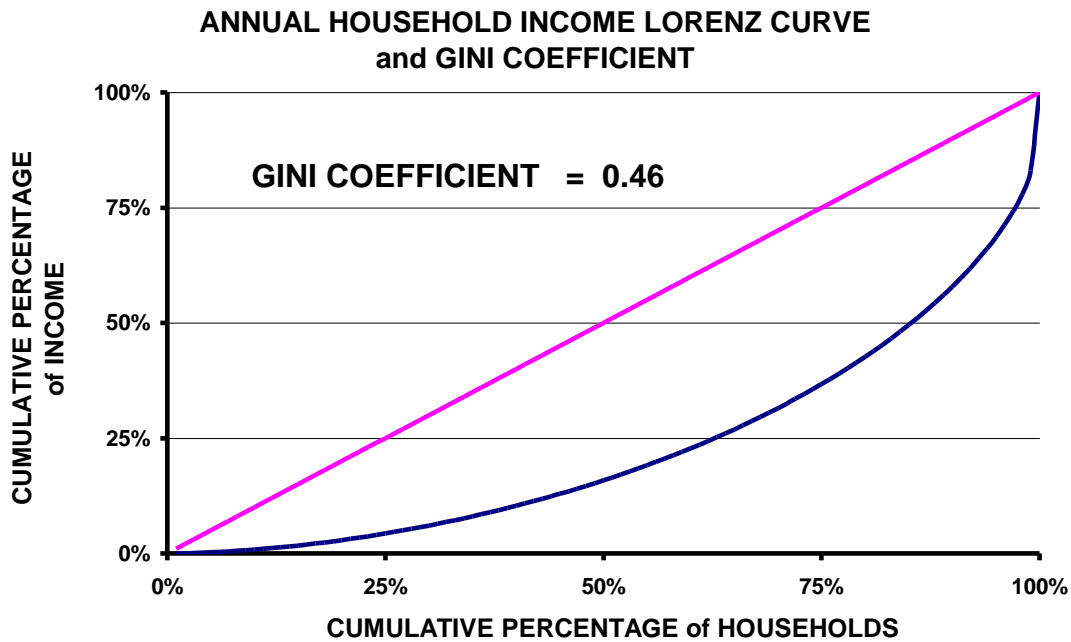


Figure 3. Annual Household Income Lorenz Curve and Gini Coefficient.

3. ANNUAL HOUSEHOLD INCOME VERSUS AGE OF THE HEAD OF HOUSEHOLD

Given the familiar, and common, work/life history of most working American males over the last 50 years, the shape of the graph of the relationship between average annual household income on age of the head of household shown in Figure 4 is to be expected:

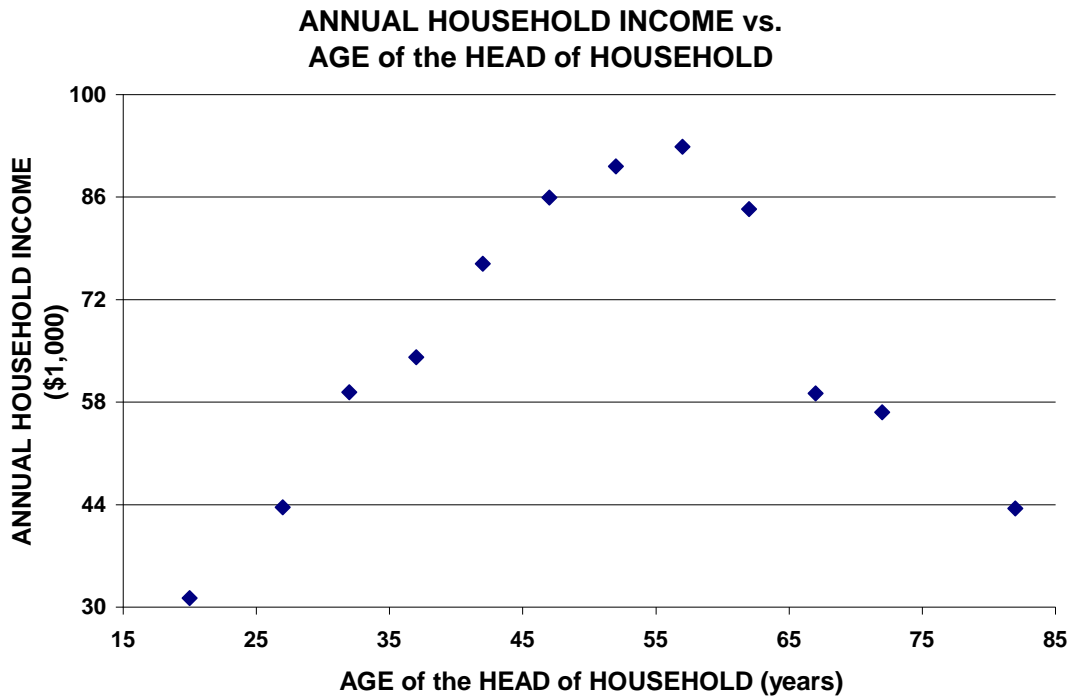


Figure 4. Annual Household Income vs. Age of the Head of Household.

In general, mean annual household income increases most rapidly after education is completed: whether secondary or higher level. Then, as careers start to progress, average annual household income reaches a peak in the years prior to retirement as a consequence of inflation-based increases, merit raises, promotions, or just seniority and its concomitant pay boosts. After that, average annual household income decreases as retirement becomes widespread throughout the previously working population. The often cited rule of thumb---that is in retirement one needs 80% of one's pre-retirement income to maintain the same standard of living---is clearly not enjoyed by the bulk of the U.S.'s population. Social security "benefits", defined benefit and defined contribution pensions, and tax-

protected sources of retirement income [e.g., IRAs and 401(k)s] are not enough to fulfill this general prescription for most people (whose typical savings in this form---plus equity in their principal residence's real estate---amounted to approximately \$50,000 in 2004).

4. ANNUAL HOUSEHOLD INCOME VERSUS LEVEL OF EDUCATION

Table 1 lists the average annual household income versus level, i.e., number of years, of education. The 2000 Census Bureau Current Population Survey data (March, 2000) are in the first and third columns of the table. The years of education I assigned are in the middle column. A professional degree was taken to mean an M.B.A., J.D., M.D., D.D.S., or so forth. From this table Figure 5 was constructed.

The relationship between average annual household income and level of education is amazingly close to linear. The high, positive, slope also confirms the old saw that more education is better in terms of lifetime earnings. Buttressing the conclusion that education beyond secondary school (i.e., High School) did pay over the longer term is the 2005 study by the Organization for Economic Cooperation and Development (OECD). Its annual research publication, "Education At

Table 1. Mean Annual Household Income vs. Years of Education.

CENSUS BUREAU DESCRIPTION of LEVEL of EDUCATION	NUMBER of YEARS of EDUCATION ASSIGNED	CENSUS BUREAU MEAN ANNUAL HOUSEHOLD INCOME ` (2003 Dollars)¹
Less than 9'th Grade	6	27,009
9'th to 12'th Grade and no Diploma	10.5	31,298
High School Graduate Including the GED	12	47,365
Some College, no Degree	14	56,763
Associate Degree	14	62,373
Bachelor's Degree	16	84,864
Bachelor's Degree or More	17	92,568
Master's Degree	18	96,240
Ph. D.	21	117,997
Professional Degree	23	137,138

¹These are the numbers in the actual table. Clearly, the last digit (two digits?) are meaningless in the sense of accuracy.

A Glance”, confirms that obtaining a B.A. or B.S. degree in the U.S.A. had a net present value, using a 5% per annum discount rate, of about \$135,000 or \$82,000 better off that a high school graduate (2005 purchasing power parity figures and other data). This computation includes tuition fees and other education-related expenses, lost earnings, and additional taxes paid on a higher

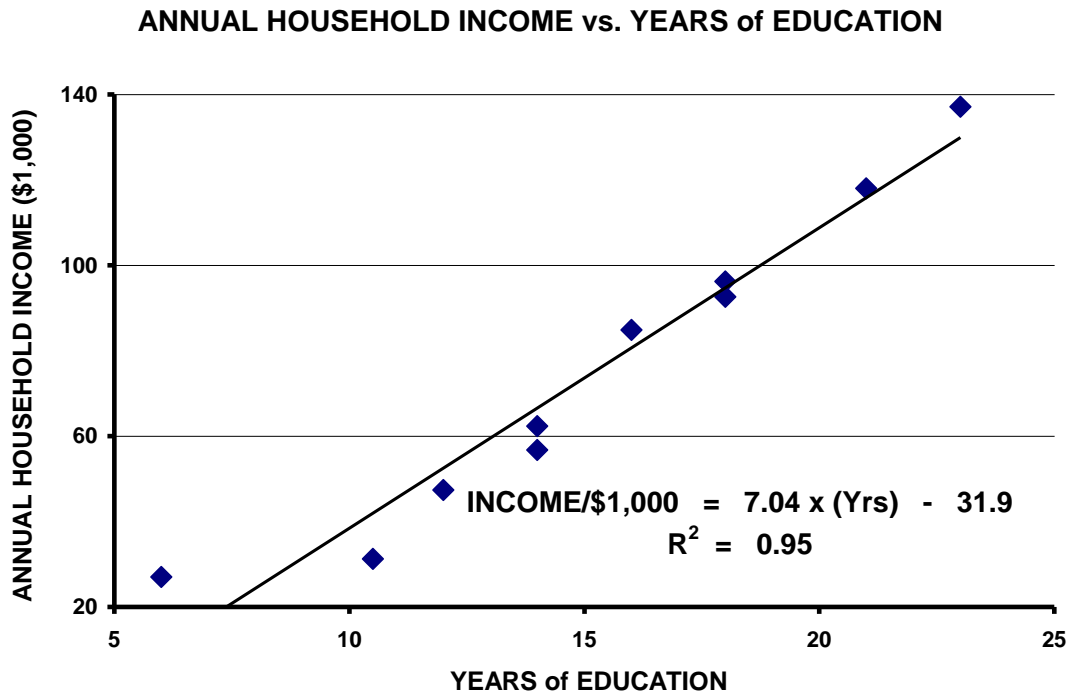


Figure 5. Annual Household Income vs. Years of Education.

income over time. It will be interesting to see if current graduates do as well given the recent, and no doubt continuing, financial woes of the Western world.

5. ANNUAL HOUSEHOLD INCOME VERSUS HOME VALUE

Those knowledgeable about good underwriting criteria for granting mortgage credit for single-family residential real estate purchases would expect a strong association between average annual household income and the cost of the property at the time of purchase. For those of you unfamiliar with the details in the U.S.A., the basic rules used to be (before the 2005-2007 pre-sub-prime

mortgage market fiasco fomented by the approval of mortgage---i.e., home buying---credit for “liar”² and “NINJA”³ loan applicants):

- 1) The front-end debt-to-income ratio should be no more than 28% (33% at most for a prime credit with no other long-term debt). The front-end debt-to-income is defined by the quotient of the monthly housing expenses (i.e., mortgage interest and principal, real estate taxes, hazard insurance, and home owner association dues if any) with respect to gross monthly income. Thus, a potential home owner, with an annual income of say \$80,000, should be able to afford monthly payments of $0.28 \times \$80,000 = \$1,870$.
- 2) The back-end debt-to-income ratio should be no more than 33% (sometimes allowed up to 36% for a prime credit in expensive areas with no other long-term debt). The back-end debt-to-income ratio is defined as the quotient of the total amount of long-term monthly debt payments with respect to gross monthly income. For instance, adding a \$350 monthly car payment to our hypothetical want-to-be home owner means

²A *liar loan* is just what it sounds like. Some of the entries on the loan application form are untrue. The fraud is usually centered on the stated income amount. Sometimes the level of total assets is “modified” too. The polite name for a liar loan is a stated income loan.

³*NINJA* is an acronym for **No Income**, **No Job**, and no **Asset** verifications; the perfect creditworthy loan applicant.

that the sum of \$1,870 and \$350 (= \$2,220) should be less than $0.33 \times \$80,000 = \$2,222$ which it just is. And,

- 3) The loan-to-value ratio should be no more than 80%. In other words, the down payment is expected to be at least 20% of the purchase price.

With a mortgage interest rate of 5.5%/year, the monthly payment on a \$320,000 (a \$400,000 purchase price: 30-year, fixed-rate, monthly-paying, level-payment, fully-amortizing mortgage---that is the standard kind) is \$1,820 or a front-end debt-to-income-ratio of 27%. In this instance the ratio of the average house price to mean annual household income is 5.1:1.

For the pre-sub-prime mortgage disaster median house price of \$200,000, and a family earning the then median income of about \$50,000 per year, the prudent house price to annual household income ratio is 4.0:1 for a mortgage with the above characteristics (a monthly payment of \$908 and a safe front-end debt-to-income ratio of 22%). Reducing the annual household income to \$40,000 raises the house price to annual household income ratio to 5.0:1 and increases the front-end debt-to-income ratio to 27%.

The relationship between average annual household income and home value, for the entire American population, is illustrated in Figure 6. Observe the very rapid turn-up for the higher income people. High annual household income does not indicate a high level of wealth and delinquencies and defaults are common among such households.

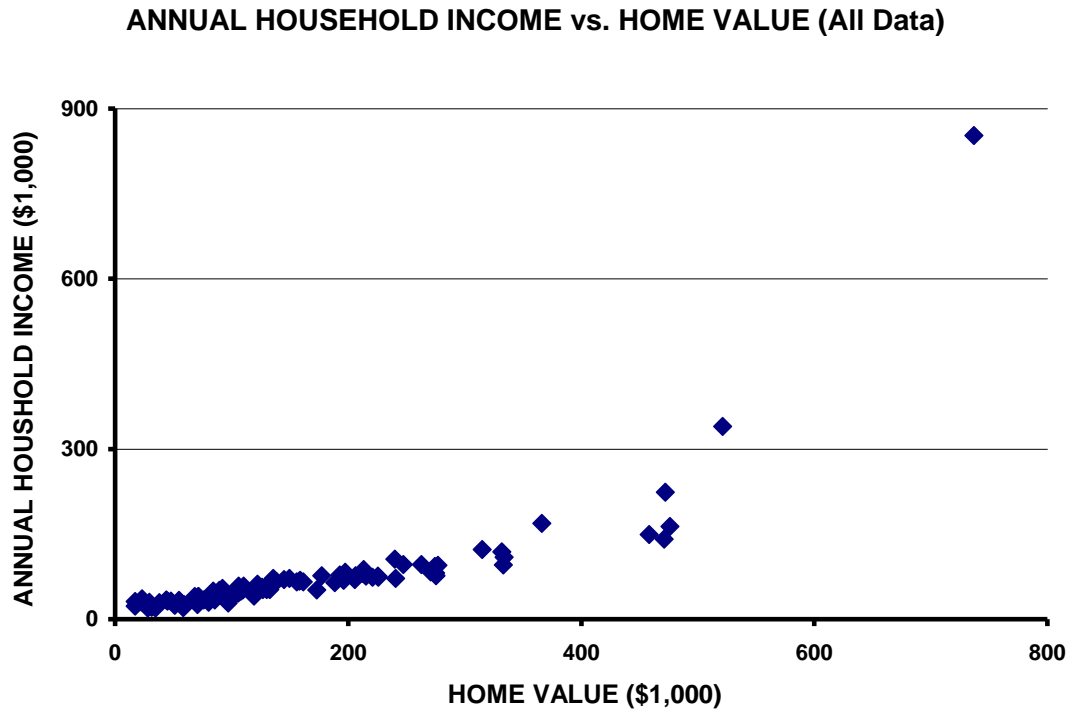


Figure 6. Annual Household Income vs. Home Value (All Data).

[As an aside, there is no method of taking out the effect of real estate appreciation, possible (serial) re-financing activity, alternative mortgage products, the amount of the original down payment, and so on because there is no way to infer the original purchase price or any of these other characteristics. Thus, purchase price and current home value, or market price, are not necessarily highly correlated. That they are can be seen in Figure 7. The explanation follows the picture.]

Keeping the underwriting discussion above in mind, and removing the highest two annual household income points, Figure 7 shows that the relationship between annual household income and home value is linear for the bulk of America's home owners. This might seem surprising especially given the geographic variability within the U.S.A., the assortment of mortgage products, the

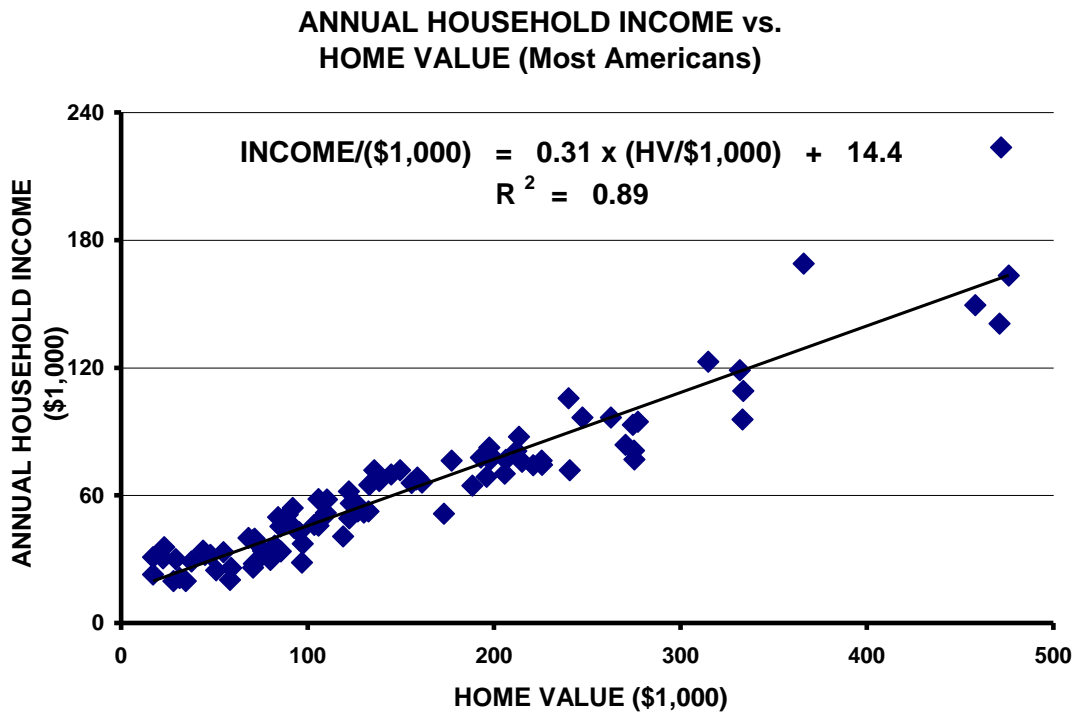


Figure 7. Annual Household Income vs. Home Value (Most Americans).

sometimes rapidly changing availability of funds for home buying, sundry mortgage interest rates over the years, a variety local underwriting practices, the extreme diversity of neighborhood residential real estate markets, housing price inflation, and so on. The slope is also below that for a conforming⁴ home loan (i.e., mean house price to average annual household income ratio of 3.1:1). Why is this the case? Because of the dominance by Fannie Mae and

⁴A conforming mortgage is one that conforms, or meets, the usual underwriting standards of Fannie Mae and Freddie Mac.

Freddie Mac of the secondary mortgage market. Whether or not a retail mortgage origination firm plans to sell their product to one of the Agencies, they usually want the ability to do so in case they need cash for liquidity purposes. Hence, the near universal, but voluntary, accord with the Agencies underwriting standards.

6. ANNUAL HOUSEHOLD INCOME VERSUS TOTAL FINANCIAL ASSETS

The curve portrayed in Figure 8 is very puzzling to me. It illustrates the association between *current* annual household income and a household's *accumulated* stock of wealth. I would expect today's total savings to be related, in a very complicated fashion, to the time histories of annual household income, savings behavior, money allocation in a household's financial portfolio, capital gains and income taxes, investment returns, and "dis-saving" activity. Nonetheless, the power law function fit exhibited in Figure 8 is essentially perfect given that the source is financial data.

Note the scale on the abscissa; this plot includes the wealthiest households. Nonetheless, the pattern, and the exponent of the power law function, are both repeated for ever "poorer" households.

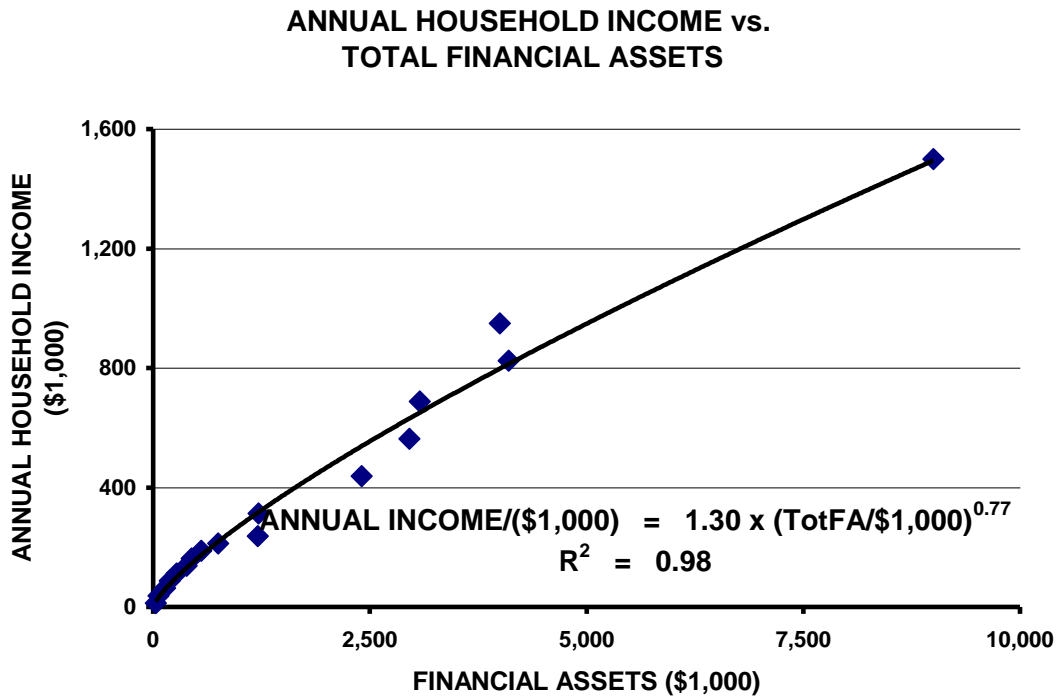


Figure 8. Annual Household Income vs. Total Financial Assets.

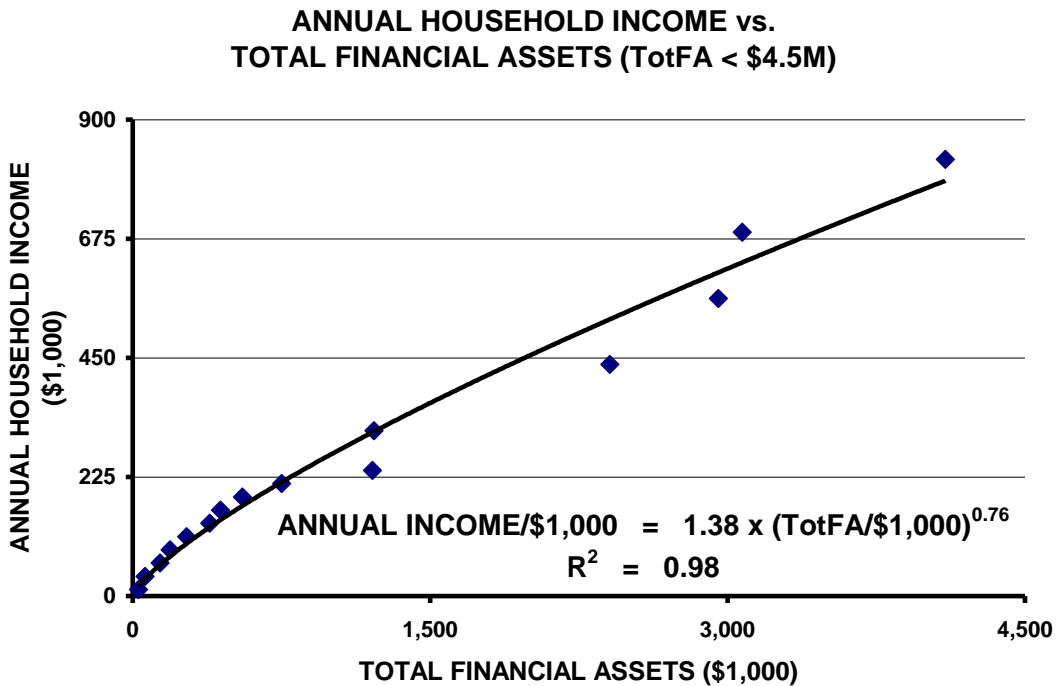


Figure 9. Annual Household Income vs. Total Financial Assets (TotFA < \$4.5M).

and finally,

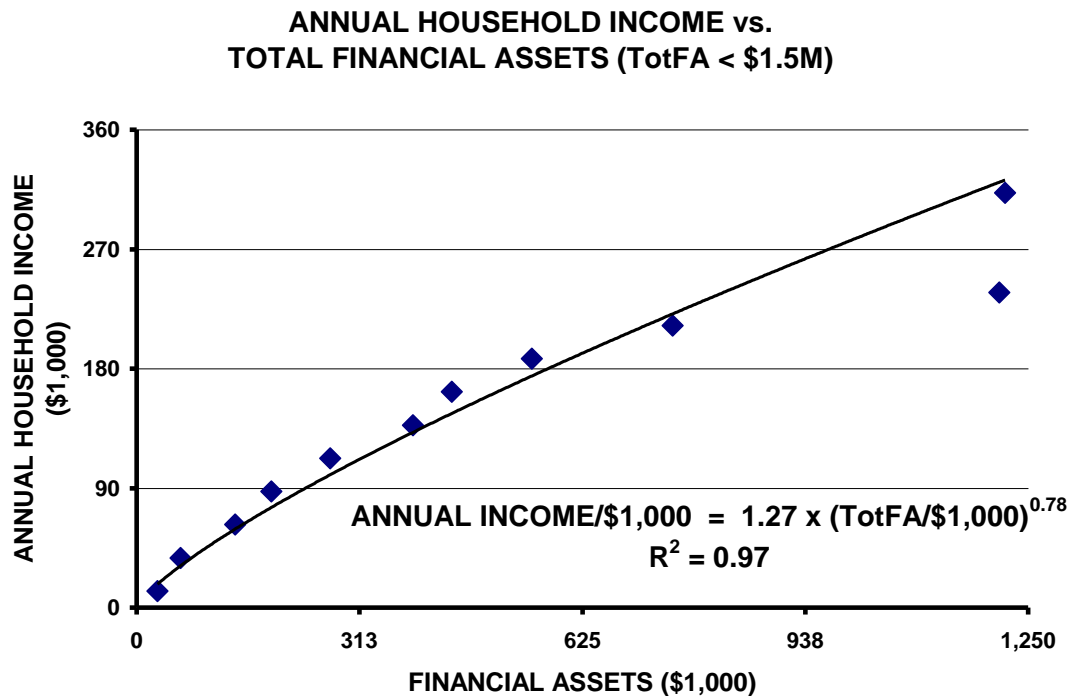


Figure 10. Annual Household Income vs. Total Financial Assets (TotFA < \$1.5M).

7. ANNUAL HOUSEHOLD INCOME VERSUS PORTFOLIO COMPOSITION PARAMETER

The Definition Of The Herfindahl Index

The Herfindahl Index is defined as the sum of the squares of the percentages of the market shares of a set of firms in a specific segment of the economy. The larger the Herfindahl Index, the higher the level of dominance of that market niche by a fewer number of

firms. Conversely, the smaller the Herfindahl Index, the less concentrated the market shares of the companies are. As real life examples, both the U.S. Government and the European Union use the Herfindahl Index as one guide as to whether or not to approve mergers between rival commercial enterprises.

In symbols, suppose that there are $N > 1$ firms in an industry. Number them by $n = 1, 2, \dots, N$. Indicate the market share of each company, in percentage terms, by p_n , $n = 1, 2, \dots, N$; $0 \leq p_n \leq 1$ for $n \in [1, N]$. The definition of the Herfindahl Index (HI) is

$$HI = p_1^2 + p_2^2 + \dots + p_N^2, \quad HI \in [1/N, 1].$$

Note that the Herfindahl Index is computed under a constraint; since the market shares are expressed as percentages, the sum of the market shares must be 100%:

$$\sum_{n=1}^N p_n = 1.$$

An Alternative Interpretation

Instead of considering N companies in the same market sector, consider N assets of a specific household's financial portfolio. Rather than have p_n represent the market share of the n 'th firm, it will now represent the share of the n 'th asset category: percentage or dollar amount?

When computing the Herfindahl Index, given its purpose, it only made sense to consider the $\{p_n\}$ as percentages. However, it is now possible to think about a p_n as being the dollar amount devoted to the n 'th asset type rather than its percentage of the total stock of financial wealth. Which to choose? To decide the issue think about two \$1 million portfolios. One is completely invested in Certificates of Deposit, the other in small capitalization stocks. Using percentages both would have an equal value of Portfolio Composition Parameter (as will be seen below). This would be very mis-leading. On the contrary, if dollar amounts were used---while both would still have the same value of the Portfolio Composition Parameter---the probability of a million dollar portfolio being so conservatively invested is very low. With a more realistic asset allocation commensurate with the two hypothetical households, a dollar-based Portfolio Composition Parameter will, in almost all cases, be more revealing than a percentage-based version could ever be. Hence, the Portfolio Composition Parameter is computed using dollar amounts.

The Definition Of The Portfolio Composition Parameter

The quadratic nature of the Herfindahl Index will produce extremely large numbers when dollar amounts are used. So, to bring the results back into a more reasonable numerical range, namely 1-10, the sum of the squares of the $\{d_n\}$ (d now for dollars) is normalized by a combination of first dividing by a large number and then taking an irrational root---equal to about one tenth---of the quotient.

To make the discussion more concrete, focus on one household. Divide its financial assets into N ($= 7$ as is in the Survey of Consumer Finances) mutually exclusive investment categories. Let d_n be the amount of dollars invested in the n 'th asset type for $n = 1, 2, \dots, N$. Compute the sum of the squares of the $\{d_n\}$, namely

$$d_1^2 + d_2^2 + \dots + d_N^2.$$

Divide the result by a large number, say 10^{10} . Now take a small irrational root of the reduced sum of the squares. The result is the Portfolio Composition Parameter.

The seven asset categories are (my paraphrasing, with minor modifications from the Survey of Consumer Finances groupings, and in order of decreasing risk):

- 1) "Other"; a heterogeneous category of less common financial contracts.
- 2) Individual shares of publicly traded common and preferred stock.
- 3) Individual Federal government, municipal government, Agency, corporate, and foreign bonds.
 - a. This category includes "mortgage-backed bonds" by which is almost certainly meant mortgage-backed securities and not true mortgage-backed bonds.
 - b. I also placed savings bonds in this category.
- 4) Mutual funds; non-money market pooled investment funds.
- 5) Annuities.

- 6) Keogh and IRA accounts and other pension accounts from which withdrawals or loans may be taken, e.g., 401(k) accounts. And,
- 7) Transaction accounts; e.g., savings, checking, and money market deposit accounts, money market mutual funds, and call accounts.
 - a. I also placed Certificates of Deposit in this category.

Then the computation of the Portfolio Composition Parameter first requires the calculation of (the dollar sign indicates the number of dollars allocated to the abbreviated asset class)

$$\begin{aligned}
 &(\$OTHER)^2 + (\$STK)^2 + (\$BNDS)^2 + (\$MF)^2 \\
 &+ (\$ANN)^2 + (\$RET)^2 + (\$TRANS)^2.
 \end{aligned}$$

As indicated above, this sum was divided by a large number because when using dollar amounts it can easily become a very big figure. Finally, to complete the computation and produce reasonable quantities, for example less than 10, the Portfolio Composition Parameter is a small irrational root of the just obtained quotient.

When the computations are completed the next figure, number 11, can be constructed. The result portrayed in it is remarkable. There is an almost perfect, once again power law function, relationship between average annual household income and the asset mixture of a household. Why?

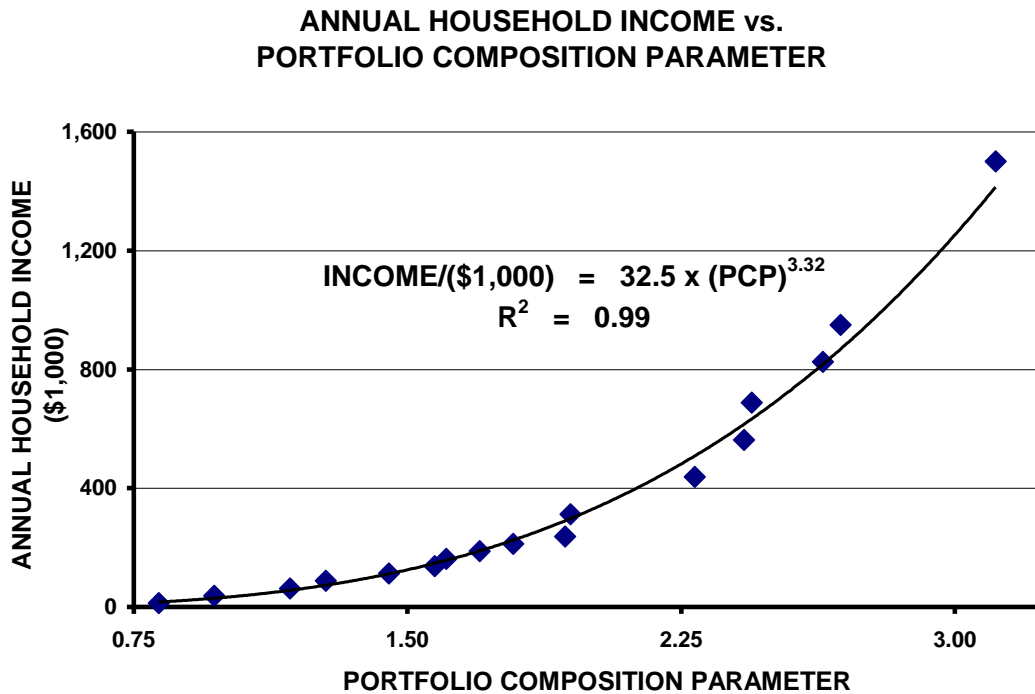


Figure 11. Annual Household Income vs. Portfolio Composition Parameter.

8. SUMMARY

I have provided both an analysis of the frequency and cumulative distributions of American mean annual household income and its relationship with the most important demographic variables. Furthermore, I have demonstrated a surprising relationship between annual household income and portfolio composition: it is a nearly perfect power law. The power law dependence of average annual household income on total financial assets has also been illustrated. These results provide novel insights into the asset allocation of American families and their risk acceptance/rejecting propensities.

REFERENCES

Bucks, B.K., Kennickell, A.B., and Moore, K.B., "Recent Changes in U.S. Family Finances: Evidence from the 2001 and 2004 Survey of Consumer Finances", *Federal Reserve Bulletin*, 92 (February, 2006), pp. A1-A38.

Bucks, B.K., Kennickell, A.B., Mach, T.L., and Moore, K.B., "Changes in U.S. Family Finances from 2004 to 2007: Evidence from the Survey of Consumer Finances", *Federal Reserve Bulletin*, 95 (February, 2009), pp. A1-A56.