

Average debt and equity returns: Puzzling?

By Ellen R. McGrattan and Edward C. Prescott*

Historically, the average return on S&P stocks has far exceeded the average return on short-term U.S. government debt. Rajnish Mehra and Edward Prescott (1985), for example, found that the average difference was 6.2 percent per year in the 1889–1978 period. They tried to account for this difference by assuming it is a premium for bearing nondiversifiable aggregate risk, but found that risk accounted for only a tiny fraction of the difference. They concluded that there is an “equity premium puzzle.”

Here, we re-examine this puzzle, taking into account some factors ignored by Mehra and Prescott — taxes, regulatory constraints, and diversification costs — and focusing on long-term rather than short-term savings instruments. Taxes should not be ignored because individuals have faced different effective tax rates on their interest and dividend income in most years during the past century. Further, the difference in effective tax rates has varied a lot over time because of changes in both the tax code and the regulations governing financial intermediaries. Other regulatory constraints that have mattered are government regulations on households and businesses during World War II. Diversification costs should not be ignored because they have been high and have varied by asset and by period. One final difference in our analysis is the focus on long-term savings instruments, whereby long implies long enough so that assets’ liquidity values are small. Individuals do not hold 90-day U.S. Treasury bills for their retirement.

Unlike Mehra and Prescott, we find that there is no equity premium puzzle. Accounting for taxes, regulations, and costs, the difference between average debt and equity returns during peacetime in the last century is less than 1 percent, with the average real debt return

almost 4 percent, and the average real equity return somewhat under 5 percent. As theory predicts, the real return on debt has been close to the 4 percent average after-tax real return on capital. Similarly, as theory predicts, the real return on equity is equal to the after-tax real return on capital plus a modest premium for bearing nondiversifiable aggregate risk.

I. A Key Condition for Asset Returns

We start with a key first-order condition for a household making savings decisions and choosing among various assets. This condition, which is standard in macroeconomics and finance, is the motivation for our accounting exercise.

Consider a household with a period utility function given by $u(c_t, l_t)$, where c_t is consumption in period t and l_t is leisure in period t . The household can choose to allocate savings to different portfolios. Here we will limit attention to savings in capital, equity assets, and debt assets held long term. In equilibrium, household choices must satisfy

$$0 = \beta E_t \left[\frac{u_c(c_{t+s}, l_{t+s})}{u_c(c_t, l_t)} (r_{t,t+s}^i - r_{t,t+s}^j) \right], \quad (1)$$

for each pair $i, j \in \{e, d, k\}$, where e denotes equity, d denotes debt, k denotes capital, β is the household's discount factor, $r_{t,t+s}^i$ is the *realized* return after taxes and costs on portfolio i between t and $t + s$, and s is large enough so that the assets' liquidity value is small.

Mehra and Prescott (1985) use condition (1) with a small s . They compare theoretical returns of stocks and bonds with the return on a portfolio of S&P 500 stocks and that on a 90-day U.S. T-bill. They find that the difference in average returns on stocks and bills over 1889–1978 is too large to be a premium for bearing nondiversifiable aggregate risk.

Mehra and Prescott's analysis has several problems which we try to avoid. One problem

is interpreting the return on a 90-day T-bill as the rate at which households intertemporally substitute consumption. We do not interpret it as such. Treasury bills provide considerable liquidity services and are a negligible part of individuals' long-term debt holdings.

Another problem with Mehra and Prescott's analysis is abstracting from the costs of creating diversified portfolios. We say that an asset is diversified if holding it involves little idiosyncratic risk. Here we are comparing average returns on diversified assets. For some of these assets, diversification costs are large.

Another expense that is relevant to the household is taxes. Until recently, dividends were taxed much more heavily on average than interest payments. The reason this was the case is that debt assets could be and were held as life insurance and pension fund reserves, thereby escaping most taxation, while equity could not be. Now equity can and is held as pension fund reserves and in tax-deferred retirement accounts, thereby escaping taxes on dividends. In the pre-WWII period important classes of debt, namely, municipal bonds and Treasury securities, were fully or partially tax-exempt.

Finally, Mehra and Prescott compare average asset returns for a period that includes a subperiod when condition (1) is not a first-order condition of households. In particular, during WWII and the Korean war, the government had restrictions on production, consumer credit, and the investments of financial intermediaries. These restrictions led to a large deviation between debt and equity returns that can only be studied if they are taken explicitly into account in the households' optimization problem.

From previous work we know that for (1) to hold approximately, the mean of the second term must be near zero because the covariance between the two terms is small — at least for standard preferences — and the mean of the first term is not zero. Most of the literature

following Mehra and Prescott (1985) has focused on modifications to preferences that are needed for condition (1) to hold if one uses unadjusted S&P 500 and T-bill returns. In the next two sections, we focus on adjustments to returns that are needed before evaluating this condition. We will use the return on capital as a comparison point for both equity and debt returns. An estimate for the return on U.S. reproducible capital can be deduced from the economy's national accounts. As we will see later, this return is about 4 percent on average and varies little over time.

II. Equity Returns

Here we compare returns on U.S. reproducible capital to the returns on a diversified equity portfolio from 1880 to 2002. We find that differences in these average returns are about one percentage point or smaller, once we account for taxes and diversification costs.

To construct a return for U.S. reproducible capital, we use after-tax capital income for the noncorporate sector and divide by the stock of capital generating this income flow. (These data are available from the U.S. Bureau of Economic Analysis, 1929–2002.) Our construction uses the noncorporate sector because investment of corporate intangible capital is large but is not included in national income. (See Ellen McGrattan and Prescott, 2002.)

To construct a return for corporate equities, we must correct for taxes and diversification costs paid. Since 1913, individuals holding corporate equities have been paying taxes on corporate distributions, that is, on dividends and realized capital gains. Income tax rates were quite low before the mid-1930s, but rose dramatically during WWII. High-income households that had net long-term capital gains took advantage of the alternative tax and lower marginal tax rates on dividends. But even these households faced high marginal tax

rates during and after the war.

In Figure 1, we plot estimates of the average marginal tax rate on dividend income for 1913–1999 using data from the U.S. Internal Revenue Service (IRS). For each year, we constructed a weighted average of the marginal rate on an additional dollar of dividend income reported on 1040 returns. We use the fraction of dividend income for the weight of each marginal tax rate class. To account for state and local income taxes, we multiply the 1040 marginal rate by the ratio of total individual income tax revenues to federal individual income tax revenues. To account for the fact that some equities are not taxed or are tax-deferred, we multiply the rate by the fraction that is taxed. As is evident in the figure, taxes during the war and early postwar periods were high, with the peak being 50 percent.

We also want to account for the taxes paid on realized capital gains which are unavoidable in rebalancing an equity portfolio over time. This is hard to do with the data available from the IRS. So here we abstract from these costs and view our estimate of the average return to equity as an upper bound.

In addition to taxes, we need to account for costs of holding a diversified equity portfolio. We use Investment Company Institute (ICI) (2002) numbers to estimate the sum of mutual fund costs and annuitized sales loads relative to the sum of a fund assets. As Figure 2 shows, these costs are large. In fact, they are even larger than the figure shows because brokerage fees associated with buying and selling securities are excluded from them. (For years before 1980, we use the 1980 number.)

One final computation is needed before constructing a time series of equity returns: an adjustment for inflation. In each period, we subtract from returns inflation in the consumer price index, which is available for our entire sample period.

Figure 3 plots the resulting real equity and capital returns after measurable taxes and costs have been accounted for. (Equity returns are available from Cowles Commission, 1939, and Ibbotson Associates, 2002.) We have smoothed equity returns using a 31-year centered moving average, in order to see the pattern of average returns over time. On the figure, we include a line marking 4 percent, which is the average of the NIPA capital return, our comparison point for equity and debt returns. An upper bound for the average equity return — if we ignore capital gains taxes, brokerage costs, and possibly higher pre-1980 diversification costs — is 5.4 percent for 1880–2002.

The largest difference between returns on equity and capital occurs after 1980. McGrattan and Prescott (2002) find that the large decline in the tax rate on dividends generated this large transient capital gain. The average equity return for the period before 1980 is 5.1 percent, which is about one percentage point above the average return to NIPA capital. Since capital gains taxes and unmeasured portfolio costs are not negligible, the difference between equity and capital returns is less than one percentage point.

III. Debt Returns

Now we compare returns on U.S. reproducible capital to the returns on long-term debt assets during 1880–2002. We find that differences in these average returns are small during the gold standard period and after the Korean War. During WWII and the Korean War, regulatory policies led to very low returns on debt assets relative to capital.

A. The Gold Standard and Postwar Periods

During the gold standard period, savers in long-term debt assets realized relatively high returns: close to 4 percent. In Figure 4, we plot yields on long-term, high-grade municipal bonds for 1880–1934. (Rates are available from Sidney Homer and Richard Sylla, 1991, and the Federal Reserve Board of Governors, 1914–2002.) These bonds were tax-exempt, and many were held directly by individuals, so we do not need to adjust for taxes or for costs incurred by financial intermediaries. Since prices were expected to be stable in this period, we also do not adjust for inflationary expectations.

Other evidence on the returns to savers in the early period of our sample shows that long-term debt assets earned high returns. S. H. Nerlove (1932) studied the performance of life insurance companies between 1906 and 1929. Because of state regulations, virtually all of the investments made by these companies then were in debt assets. Nerlove documents that the “excess return on the investments of these companies over the yield on high-grade bonds made it possible for these companies to pay an average rate of interest on instalment settlements and on dividends left to accumulate in many cases as high as, and in some instances higher than, the average yield on high-grade bonds” (p. 163). Thus, even after intermediation costs, savers in insurance companies realized returns at least as high as 4 percent. Furthermore, these investments were, in large part, tax-free. Annuity payments were taxed as ordinary income, but the assets grew untaxed between payments of premiums and receipts of installment payments. Survivors’ benefit payments were not taxed at all. The findings of Nerlove (1932) are consistent with data from the Teachers Insurance and Annuity Association (TIAA). In each of the years 1919 through 1931, the total effective interest rate

on traditional TIAA annuities was 4.5 percent.

In Figure 4, we also plot estimates of returns on long-term, high-grade bonds in the postwar period. The average for 1960–2002 is 3.72 percent, which is close to the 4 percent average return on NIPA capital. In this postwar period, we focus on the highest grade corporate bonds, which were held primarily in tax-deferred pension funds. An implication of the evidence of Nerlove (1932) and the TIAA is that the returns on high-grade bonds are a good proxy for what savings in annuities would have realized in the early period we consider. We assume that high-grade bond returns are a good proxy for the returns that savers can realize on pension funds and annuities in the postwar period.

A major concern of debt asset holders in the postwar period has been inflation. To account for inflationary expectations, we convert nominal yields to an estimate of real yields by subtracting the average inflation rate over the previous 10 years.

The average return on debt for the two periods displayed in Figure 4 is almost 4 percent. In the next section, we show that this is significantly higher than the average return during WWII and the Korean War.

B. World War II and the Korean War

The U.S. Treasury and the Federal Reserve succeeded in keeping nominal yields on Treasury securities at or below 2.5 percent during the wars. To achieve this goal, they had to regulate the behavior of households and business in addition to buying and selling Treasury securities.

One form of regulation was through restrictions on production, which affected the availability of many consumer goods. For example, in January 1942, civilian production of automobiles was halted by government order, and companies converted their manufacturing

capabilities to military purposes.

A second form of regulation was through restrictions on credit. For example, Regulation W, enacted in 1941, prescribed minimum down payments, maximum maturities, and other terms applicable to consumer credit. Credit controls under Regulation W, and later for real estate under Regulation X, were in effect until 1952.

A third form of regulations was through restrictions on household investments that had been in effect before WWII and became important during the war with the policy of fixed yields instituted by the Treasury and the Federal Reserve. Federal and state governments restricted the types of assets that could be held by trust funds, life insurance companies, and savings banks, and investors had few “safe” alternatives to government debt. Thus, the government could tap into the savings of the investors by lending to these financial intermediaries at very low rates.

Figure 5 summarizes our results. There we include the war period and returns for equity and debt and, for reference, the line marking 4 percent. It is clear that debt returns during the war period were low. Returns averaged -0.65 percent for the period 1940–1955 and 3.32 percent for the entire sample period. In normal times, debt returns were slightly below returns to capital, and equity returns were somewhat above.

IV. Future Research

So where’s the equity premium puzzle? There is none. We find that average real returns, after adjusting for taxes and diversification costs, are not puzzling. The equity risk premium is modest as predicted by the standard growth model used in macroeconomics to study growth and fluctuations.

But much work is still needed, especially in three areas. First, while much has been written about recent developments in financial services and the opportunities for long-term savers, little is known about the circumstances during the pre-war and war periods, especially for individuals in lower and middle income classes. Second, with regard to theory, more work is needed to derive predictions for asset returns and savings behavior from models with individuals of different types, for example, by various ages, abilities, and opportunities. Third, although average returns no longer seem puzzling compared to the predictions of theory, the excessive stock price volatility puzzle remains. (See the work of Sanford Grossman and Robert Shiller, 1981, and Stephen LeRoy and Richard Porter, 1981.)

*McGrattan: Federal Reserve Bank of Minneapolis, NBER, and University of Minnesota; Prescott: University of Minnesota, Federal Reserve Bank of Minneapolis, and NBER; both: Research Department, Federal Reserve Bank of Minneapolis, 90 Hennepin Avenue, Minneapolis, MN 55401-2171. We thank Brett Hammond for providing us with data from the TIAA, and Hanno Lustig, Erzo Luttmer, Lee Ohanian, Monika Piazzesi, and Martin Schnieder for their helpful comments. We also thank the NSF for financial support. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System. For a more detailed version of the paper and the data used in this study, see <http://minneapolisfed.org/research/sr/sr313.html>.

References

Cowles Commission for Research in Economics. *Common-stock indexes*. 2nd edition.

Bloomington, IN: Principia Press, 1939.

Federal Reserve Board of Governors. *Banking and Monetary Statistics (Later: Statistical Digest)*. Washington, DC: Board of Governors, 1914–2002.

Grossman, Sanford J. and Shiller, Robert J. “The Determinants of the Variability of Stock Market Prices.” *American Economic Review*, May 1981 (Papers and Proceedings), 71(2), pp. 222–27.

Homer, Sidney and Sylla, Richard. *A history of interest rates*. 3rd edition. New Brunswick, NJ: Rutgers University, 1991.

Ibbotson Associates. *Stocks, bonds, bills, and inflation yearbook*. Chicago, IL: Ibbotson Associates, 2002.

Investment Company Institute. “Total Shareholder Cost of Mutual Funds: An Update.” *Fundamentals: ICI Research in Brief*, September 2002, 11(4), pp. 1–8.

LeRoy, Stephen F. and Porter, Richard D. “The Present-Value Relation: Tests Based on Implied Variance Bounds.” *Econometrica*, May 1981, 49(3), pp. 555–74.

McGrattan, Ellen R. and Prescott, Edward C. “Taxes, Regulations, and Value of U.S. Corporations: A General Equilibrium Analysis.” Staff Report 309, Federal Reserve Bank of Minneapolis, August 2002.

Mehra, Rajnish and Prescott, Edward C. “The Equity Premium: A Puzzle.” *Journal of Monetary Economics*, March 1985, 15(2), pp. 145–161.

Nerlove, S.H. “Should Life Insurance Companies be Permitted to Buy Common Stocks?”

The Journal of Business of the University of Chicago, April 1932, 5(2), pp. 155–174.

U.S. Bureau of Economic Analysis. Department of Commerce. *Survey of Current Business*. Washington, DC: U.S. Government Printing Office, 1929–2002.

U.S. Internal Revenue Service. Department of the Treasury. *Statistics of Income*. Washington, DC: U.S. Government Printing Office, 1916–2002.

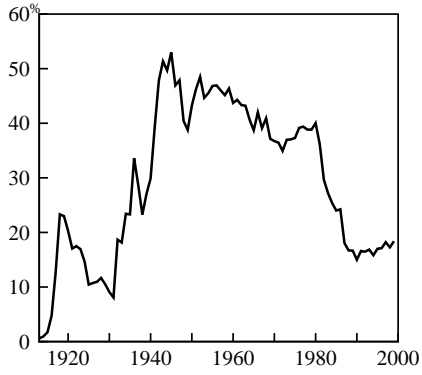


Figure 1. Average Marginal Tax Rate on U.S. Dividend Income, 1913–99

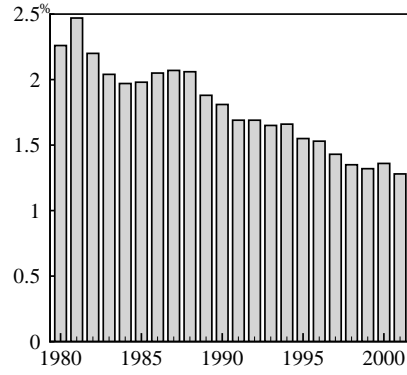


Figure 2. Equity Mutual Fund Costs (% of Assets), 1980–2001

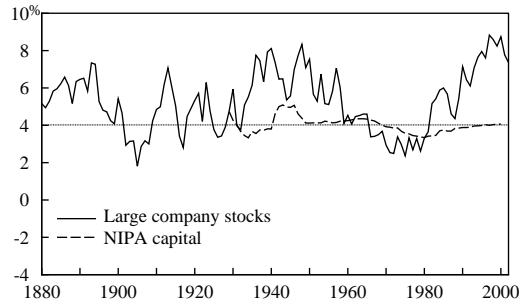


Figure 3. Equity and NIPA Capital Returns, 1880–2002

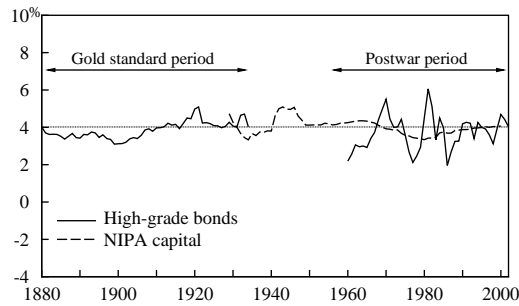


Figure 4. Debt and NIPA Capital Returns, 1880–2002

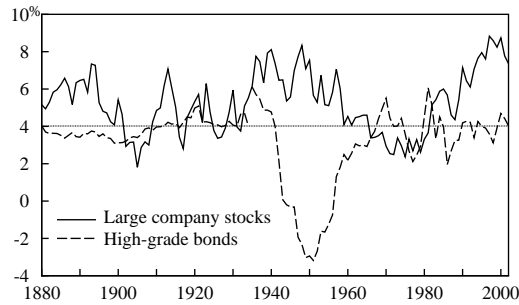


Figure 5. Equity and Debt returns, 1880–2002