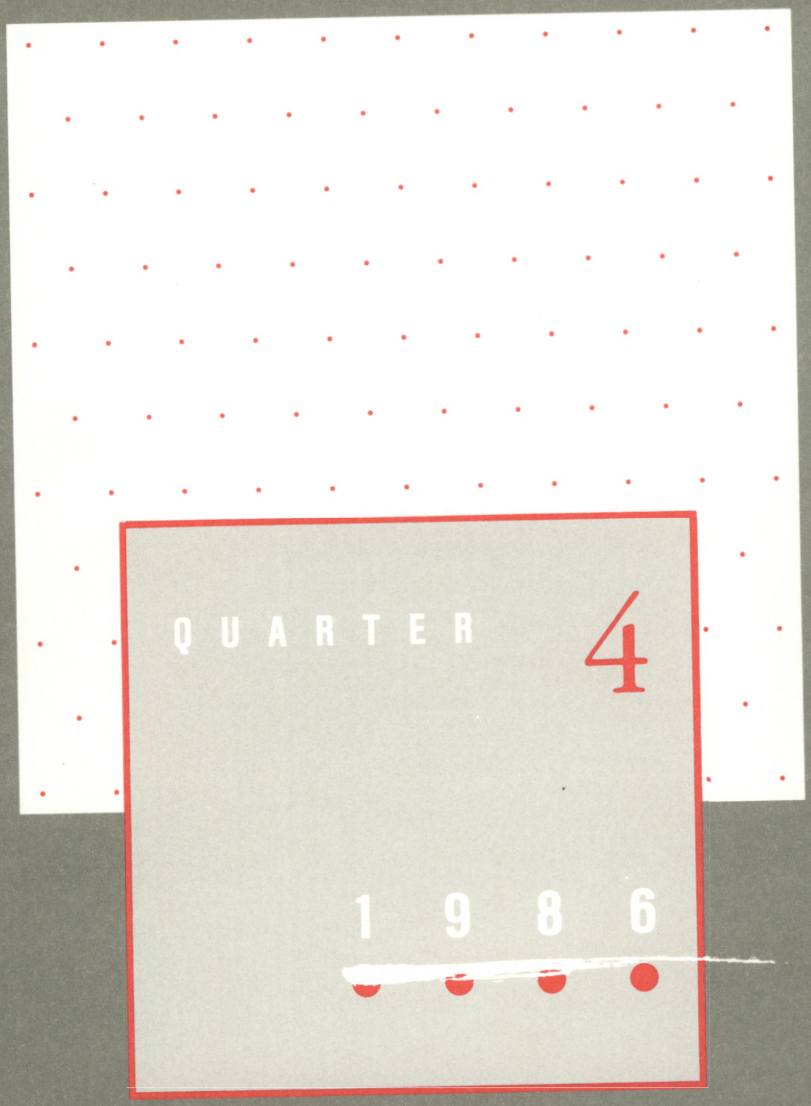
ECONOMIC REVIEW



FEDERAL RESERVE BANK OF
CLEVELAND

Economic Review is published quarterly by the Research Department of the Federal Reserve Bank of Cleveland. Copies of the issues listed here are available through our Public Information Department, 216/579-2047.

Editor: William G. Murmann.
Design: Michael Galka.
Typesetting: Liz Hanna.

ISSN 0013-0281

Disinflation, Equity Valuation, and Investor Rationality

by Jerome S. Fons
and William P. Osterberg

Jerome S. Fons and William P. Osterberg are economists at the Federal Reserve Bank of Cleveland.
The authors would like to thank James Balazsy for research assistance, and William T. Gavin, K.J. Kowalewski, Richard Kopcke and Richard Cohn for their helpful comments.

Introduction

Until the early 1960s, economists largely ignored the effect of inflation on the prices of corporate equities. Since revenues and costs were thought to be proportionately affected by changes in the price level, profits would expand so as to keep pace with inflation. As residual claims to the earnings of corporations, equities were seen as partial, if not complete, hedges against the effects of inflation.

In the late 1960s and early 1970s, this notion was shattered. Despite a 95.2 percent rise in the consumer price index (CPI) from the end of 1966 to the end of 1977, the Standard & Poor's Stock Index rose only 2.4 percent. The 52.5 percent decline in the real value of equities over this period led to the development of many theories to explain the relationship between equity prices and inflation.

Among the most widely received theories was one offered by Franco Modigliani and Richard Cohn (1979). They claimed that investors make valuation errors by ignoring the gains debtors experience from inflation and therefore use the wrong measure of profits in pricing equities. Since inflation implies that the principal of the loan will be paid back in "cheaper" dollars, lenders require an inflation premium in the coupon on the loan. This suggests that a part of the firm's debt service is used to maintain the real value of the firm's debt and should not be treated as an expense. Traditional accounting measures, however, treat the entire debt service as an expense. Modigliani and Cohn claimed that the measure of "true profits" consistent with rational valuation would equal accounting profits, plus the portion of the interest expense attributable to inflation.

They reasoned that a more serious investor error involves the comparison of the discount rate for a pure equity stream with nominal, rather than real, interest rates. In figure 1 we present a time-series plot of the nominal interest rate on Aaa-rated corporate bonds and the earnings/price ratio of stocks in the Standard & Poor's Stock Index. At least since 1960, these two series track one another well. Because long-term nominal interest rates are thought to be largely determined by inflation expectations, this comparison by investors further erodes the level of stock prices in an inflationary environment.

Modigliani and Cohn showed that, in the absence of market imperfections, the real value of the firm should remain unaffected by *anticipated* inflation. Using a statistical model, they found that investors had indeed committed one or both forms of valuation error.

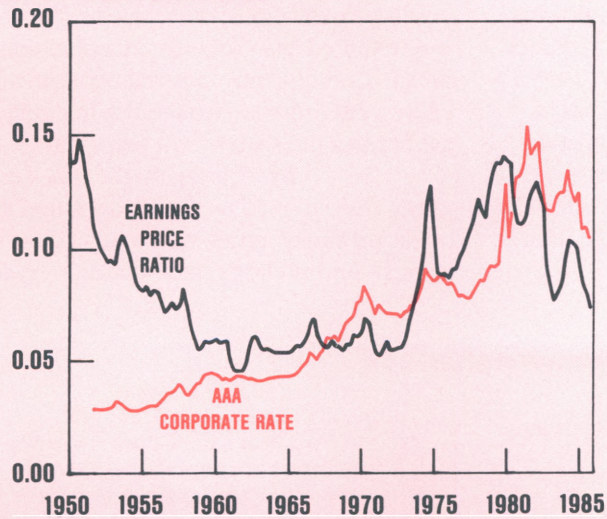
In this paper, we review the model introduced by Modigliani and Cohn and the alternative analyses of other investigators. We then evaluate those analyses by examining the behavior of the rate of return required on equities from 1953 to 1985. Surprisingly, we find little evidence of valuation errors. In particular, we note that when reported earnings are adjusted in the manner prescribed by Modigliani and Cohn, capitalization rates for equities appear to follow real interest rates, though they may also respond to factors related to aggregate risk.

I. A Fundamental Valuation Model

Fundamental equity valuation models assume that the goal of the firm's management is to maximize stockholders' wealth. Projects are accepted only if they increase the market value of the equity, that is, if the present discounted values of the expected net cash flows from new projects are positive. The market value of the firm's equity is found by discounting the cash flows distributed to stockholders at the rate stockholders could earn on alternative investment flows of equivalent risk.[1] The distribution to stockholders, or dividend, equals profits (revenue, less operating expenses and investment expenditures) minus interest payments on the firm's debt.

Following Modigliani and Miller (1958), we make assumptions sufficient to derive an expression for the value of the firm's equity: a) capital markets are frictionless, that is, participants can borrow or lend at the riskless rate of interest and there are no taxes; b) the social costs of bankruptcy are zero; c) all firms are in the same risk class; and d) equity and default-free debt are the only types of claims on firms.

EARNINGS/PRICE RATIO FOR THE S&P STOCK INDEX AND MOODY'S Aaa-CORPORATE RATE



SOURCE: Moody's Investors Service; and Standard & Poor's Corp.

FIGURE 1

The value of an unlevered (all equity) firm at date t, $V^u(t)$, with expected adjusted profits X is found by discounting the firm's expected available net cash flow at the rate that is appropriate for the firm's risk class (ρ).[2] Viewing the firm as an ongoing concern with a perpetual income stream X, its value is given by:

$$(1) \quad V^u(t) = X/\rho$$

Note that if the adjusted profits of the firm are expected to grow continually at a rate g, the firm's value can be represented as:

$$(2) \quad V^u(t) = X/(\rho-g)$$

Given the above assumptions, Modigliani and Miller go on to show that a firm's value is independent of its capital structure. That is, rational investors will ignore the effects of the firm's borrowing and base their valuation on the firm's cash flow from operations. The levered firm's total market value, $VI(t)$, is defined as the sum of the market values of equity, $S(t)$, and debt, $D(t)$:

$$(3) \quad V^l(t) = S(t) + D(t)$$

The adjusted profits available for distribution to the stockholders of a levered firm differ from an unlevered firm's adjusted profits at date t, $X(t)$, by the firm's interest expense, $rD(t)$. The expected rate of return to the levered firm's stockholders, i, is simply:

$$(4) \quad i = [X(t) - rD(t)]/S(t)$$

Combining equations (1) and (3), Modigliani and Miller's Proposition 1 states that:

$$(5) \quad X(t) = \rho[V^l(t)] = \rho[S(t) + D(t)]$$

Substituting (5) into (4) and allowing for earnings growth at rate g gives:

$$(6) \quad i = \rho + (\rho-r)d - g,$$

where $d = D(t)/S(t)$, is the firm's debt-equity ratio. The value of the equity of a levered firm can then be found by discounting the income stream available to stockholders at the appropriate rate (given by equation [6]). That is:

$$(7) \quad S(t) = [X(t) - rD(t)]/[\rho + (\rho-r)d - g]$$

Following Modigliani and Cohn, suppose that at time $t = 0$ there is no inflation and that immediately thereafter fully anticipated inflation begins at the rate p and continues forever. Adjusted prof-

1 This should be distinguished from the so-called book value of equity, found by subtracting the book value of liabilities from the book value of assets.

2 'Adjusted profits' refer to after-tax reported profits adjusted for the effects of inflation on inventory valuation and the value of actual depreciation deductions. In the NIPA these adjustments are referred to as 'IVA' and 'CCadj,' respectively. They are based on corporate tax records and assumptions about asset lives and replacement costs. For a discussion of the NIPA adjustment, see Grimm (1982). A problem with applying this adjustment to the S&P reported earnings index is that the NIPA profits measure is based on "book" profits which vary somewhat from reported earnings, especially after 1981.

its will rise continuously at the rate of inflation so that at any date t, the unlevered firm's profits, $X(t)$ will equal $X(0)e^{pt}$. From equation (1), the value of the unlevered firm at date t, $V^u(t)$, equals $V^u(0)e^{pt}$. In other words, the *real* value of the unlevered firm will not be affected by fully anticipated inflation. Rationally priced equity claims on such a firm are complete inflation hedges.

Conventional accounting measures of a levered firm's profits are distorted by inflation. Accounting profits equal operating income, minus nominal debt expense. Assume that the nominal interest rate (R) is approximately equal to the sum of the real interest rate (r) and the expected inflation rate (p) and that the firm's debt remains fixed in real terms ($D[t]$ equals $D[0]\,e^{pt}$).[3] Also assume that the firm's debt is structured so that it always pays the current rate of interest. The levered firm's accounting profits, Π, can then be written as:

$$\begin{aligned} \Pi(t) &= X(t) - RD(t) \\ &= [X(t) - (r+p)D(t)] \\ &= [X(0) - rD(0)]e^{pt} - pD(0)e^{pt} \end{aligned} \tag{8}$$

The firm's accounting profits have been expressed in this form to illustrate the following essential points. 1) The portion of reported interest expense attributable to inflation, $pD(0)e^{pt}$, should be added back to accounting profits to yield "true" profits. 2) At high enough inflation rates, accounting profits may become negative.[4]

True profits, Π^*, will therefore increase at the fully anticipated inflation rate. That is:

$$\begin{aligned} \Pi^*(t) &= \Pi(t) + pD(0)e^{pt} \\ &= [X(0) - rD(0)]e^{pt} \end{aligned} \tag{9}$$

Substituting the levered firm's true profit stream Π^*, into equation (7), we have:

$$S(t) = \Pi^*(t)/[\rho + (\rho - r)d - g] \tag{10}$$

Equation (10) therefore indicates that the *real* value of a firm's equity is unaffected by inflation. Now substituting for accounting profits and rearranging, equation (10) becomes:

$$[\Pi(t) + pD(t)]/S(t) = \rho + (\rho - r)d - g \tag{11}$$

This expression reduces to:

$$\Pi(t)/S(t) = \rho + (\rho - r)d - pd - g, \tag{12}$$

or,

$$S(t) = \Pi(t)/[\rho + (\rho - r)d - pd - g] \tag{13}$$

Equation (13) shows that, although the real value of a firm's equity should be unaffected by inflation, accounting earnings must be adjusted for inflation's effect.

Modigliani and Cohn hypothesized that investors failed to incorporate inflation in their valuations of equities. They tested this hypothesis by regressing a measure of stock prices on variables that enter either the numerator or the denominator on the right-hand side of expression (13). Their estimate of the coefficient on inflation implied systematic misvaluation. In our attempts to replicate the results of Modigliani and Cohn, however, we found that the results were sensitive to assumptions regarding lag distributions used to construct proxies for *ex-ante,* or expected, values of key variables. In addition, our attempts to replicate the results of Modigliani and Cohn yielded a coefficient on inflation that differed from their estimate (see Appendix). Rather than update their empirical work, we take a different approach to evaluating the performance of Modigliani and Cohn's model.[5]

We utilize observable, *ex-post* observations on each of the relevant variables to simulate the model, calculating implied values for ρ, the required real rate of return of a pure equity stream. To the extent that our measures of g reflect expectations, our estimate of ρ is an *ex-ante* required rate of return on a pure equity stream. Consequently, ρ is analogous to a real interest rate, adjusted for the risk in equity and the fact that the security is a perpetuity.

By focusing on the time-series values of ρ, implied by the model rather than the predicted equity values, we avoid much of the controversy surrounding equity valuation having to

3 If, as finance theory suggests, investors are concerned with after-tax real rates of return, then one could replace $R=r+p$ with $R^*=R(1-\tau)=r+p$, where τ is the marginal tax rate on interest income. Clearly, fixing r implies that the change in R^* due to a change in p is not 1 for 1. This relates to Hendershott's (1981) argument discussed below.

4 An additional factor that is thought to offset the inflation-induced gain from debt service, $pD(0)e^{pt}$, is the possible increase in the firm's pension obligations. This argument requires that inflation be unanticipated and is relevant only for defined-benefit pension plans (currently comprising roughly 75 percent of all pension assets). A defined-benefit pension is one in which contributions are determined by the benefits they will eventually yield. The obligation of the firm to restore underfunded pensions, however, rests in part on the nature of the firm's contract with labor. Feldstein and Morck (1983) find that the stock market appears to react favorably to firms with overfunded pensions and negatively to underfunded pensions. They note, however, that most large, well-managed firms have traditionally had overfunded pensions.

5 An empirical update of Modigliani and Cohn is presented by Townsend (1986).

do with the appropriate form of the discount rate.[6] The advantage of this approach is that we are able to see how ρ varies over time and, in particular, if it is correlated with inflation or real interest rates. This does not contradict the assumption that at any point in time, all variables in the denominator of (13) are expected to remain constant forever. While Modigliani and Cohn assumed that ρ is not affected by inflation, the theories discussed below allow ρ to be related to many factors, including the rate of inflation.

In order to isolate ρ we can rewrite equation (13) as:

$$(14) \quad \rho = \frac{\frac{\Pi(t)}{S(t)} + (r+p)d + g}{1 + d}$$

Using the definition of the nominal interest rate, R, we have:

$$(15) \quad \rho = \frac{\frac{\Pi(t)}{S(t)} + Rd + g}{1 + d}$$

Equation (15) shows the relation between the required real rate of return on a pure equity stream in a given risk class and mostly observable variables. The only unobservable variable is the expected growth rate of reported profits. The variable ρ may be viewed as a modified earnings/price ratio, adjusted for inflation, leverage, and earnings growth.

II. The Determinants of ρ

Below we discuss three theories of the determination of the cost of a pure equity stream. Two of the explanations given for the behavior of ρ focus on a risk premium, while the third considers the relation between ρ and the real rate of return on bonds.

In trying to explain the behavior of the stock market in the mid-1970s, Burton Malkiel (1979) adjusted corporate profits for inflation's effect on corporate debt and found them to be steady in low- and high-inflation periods. He argued that the decline in real stock prices was caused by an increase in the risk premium embodied in the rate of return required by stockholders. The increased risk premium was due to economic developments of the early 1970s that led to a departure from the relative stability of the 1960s. He reasoned that investors thought policymakers could no longer "fine tune away" economic fluctuations and that long-run planning involved greater uncertainty. Although profits rose with the price level, their dispersion across industries also rose, in turn raising business risk. The rising use of debt financing was another source of increased risk for the financial system. Finally, rising government regulation may have been perceived as reducing profitability.

As evidence supporting the perception of increased risk, Malkiel cites the rise in the "risk spread" between anticipated returns on equities and long-term government bonds, as well as between the yields on Baa-rated corporate bonds and government bonds. These widening spreads throughout the 1970s may suggest that investors believed the credit quality of firms was falling. According to Malkiel's findings, we would expect to see a path for ρ that starts out low in the '50s and '60s and then turns higher in the mid-to-late 1970s.

A related theory of the behavior of ρ involves the possibility of a disinflationary distress premium: real required rates of return on pure equity streams rise in a climate of disinflation. Firms may be under greater strain in a disinflationary environment as they are often unable to match declines in revenue with declines in expenses.[7] This is particularly evident following a period of prolonged high inflation. Extreme examples of the upheaval associated with disinflation can be found in the oil and steel industries. Further, corporate defaults have generally been higher in disinflationary periods than in inflationary periods.[8] This hypothesis implies that stockholders will require a premium whenever there are large reductions in inflation in order to compensate them for the increased credit risk. By this hypothesis, ρ should fall with increases in inflation and rise with disinflation.

Hendershott (1981) attributes the valuation error noted by Modigliani and Cohn solely to investors' comparisons of the expected real yield on equities, ρ, with the nominal yield on bonds. He claims that Modigliani and Cohn's

6 The emphasis on ρ is also justified by the implications of work done by Shiller (1981) and others on the volatility of dividends and stock prices. The literature on stock volatility suggests that profits have much lower variances than stock prices. Thus, variation in ρ and other factors influencing the rate at which profits are discounted could be expected to account for much of the variation in stock prices.

7 This may be due to the existence of fixed labor and supply contracts. A simple model introduced by Wadhwani (1986), on the other hand, suggests that the inflation premium in a levered firm's debt service causes nominal debt expense to increase proportionately more than nominal revenue during inflation, forcing the firm to report lower accounting profits. Conversely, this expense will decrease more than proportionately during disinflation, resulting in higher reported, or accounting, profits.

8 Fons (1986) investigates the correlation between "unanticipated" changes in the consumer price index and a measure of expected corporate default rates embodied in yield spreads. Though not statistically significant, the relationship between inflation surprises and an implied default premium on low-rated corporate debt is negative.

model implies that the after-tax real bond yield falls as a result of inflation, while nominal yields remain constant. Since bonds and equities are substitute assets, the fall in the after-tax real yield on debt would lower the rate of return required by stockholders. The decline in the required yield on equity offsets the overpayment of taxes resulting from the inflation-induced understatement of depreciation and inventory costs (see discussion of Feldstein [1980] below), or increased risk premia noted by Malkiel, leaving the nominal value of stocks essentially unchanged. Hendershott claimed that there were other factors responsible for the decline in the real value of equities. First, there was a decline in savings due to lower real after-tax yields. Second, there was a decrease in the productivity of new capital due to higher regulatory costs and higher energy prices. In addition, Hendershott felt that an increase in the realized rates of return on noncorporate assets, such as residential housing, may have induced investors to reduce their holdings of debt and equity.

By Hendershott's reasoning, ρ should decline in inflationary periods and rise with disinflation. Declines in productivity, however, would be reflected in a lower expected growth rate of earnings (g).

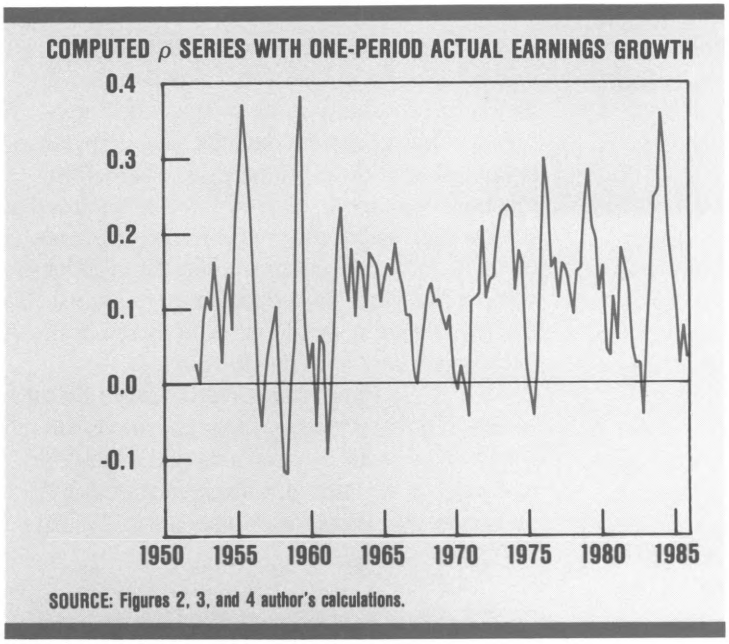


FIGURE 2

III. Data & Methodology.

Quarterly observations for the period covering 1953 through 1985 on each of the following data series were used to construct estimates of ρ: adjusted earnings per share, stock prices, nominal corporate interest rates, aggregate debt-equity ratios, and earnings growth. The values for stock prices and earnings were taken from the Standard & Poor's Index. The price index, based on as many as 500 different equities mostly traded on the New York Stock Exchange, is constructed in such a way that, when divided into the associated earnings index, the unwanted weighting factor cancels. The earnings-per-share index is constructed from the reported earnings over the past four quarters of the firms in the corresponding stock index. We adjust for inflation-caused inventory valuation and depreciation errors by multiplying the earnings index by the ratio of adjusted-to-reported after-tax profits found in the National Income and Product Accounts (NIPA) (see footnote 2).

The interest rate on corporate borrowings is measured as Moody's cross-sectional average yield on single A-rated bonds. This rating corresponded to the average quality rating (in terms of par value) of all publicly traded corporate debt as of December 1985. As was previously discussed, the nominal interest rate embodies inflation expectations. In using this measure, we avoid the problems encountered by Modigliani and Cohn in constructing an econometric proxy for expected inflation.

The debt-equity ratio for nonfinancial corporations, d, was constructed from two sources. Data covering 1953 to 1961 was taken from Von Furstenberg (1977), in which the market value of debt is inferred from a present value relation. The 1961 to 1985 series for the market values of corporate debt and equity were constructed by the Board of Governors of the Federal Reserve System. In this case, the market value of debt is found by pricing all mortgages and long-term bonds at the average price of bonds traded on the New York Stock Exchange, ignoring such nontraded items as deferred taxes, leases, and pension obligations. An attempt was made in the estimation of the market value of equity (the listed values on all exchanges, times the number of corresponding shares outstanding) to avoid the double counting of firm ownership through stock holdings.

The computation of ρ involves assumptions about the process generating the parameter g. One extreme is to let g assume its realized value equal to the annualized growth rate of four-quarter reported earnings for each period. The volatile behavior of g and ρ when g is measured this way can be seen in figure 2. We feel that such erratic movement in g is unreasonable since, in theory, g is the expected perpetual growth rate of earnings. Presumably this precludes g from being negative.

An alternative way to measure g is to utilize a time-series model to construct an in-sample one-period-ahead forecast of earnings growth. We modeled the quarterly growth of four-quarter earnings as following an ARMA(1,1) process. Using the forecast for g at each date in

the calculation of ρ yields the time-series plot of ρ presented in figure 3. This series is only slightly less volatile than the series constructed from actual growth rates.

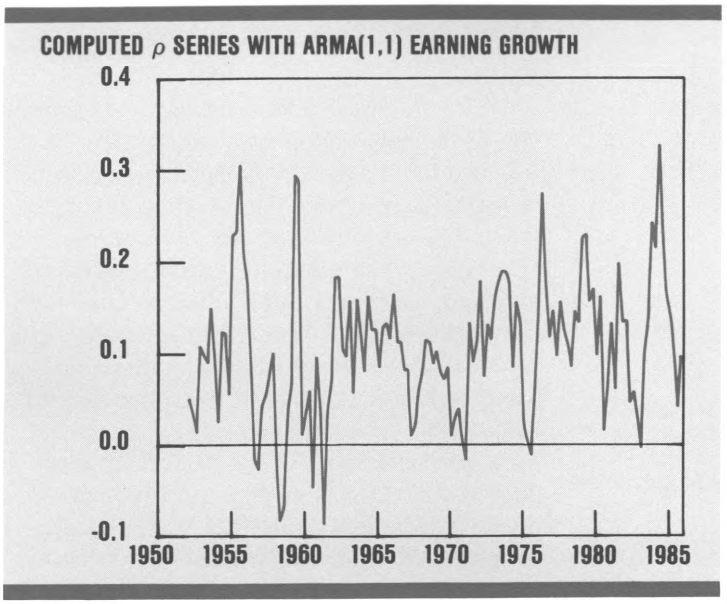


FIGURE 3

A third alternative is to fix the growth rate of earnings at its average value over the entire sample period, 6.4176 percent.[9] This procedure may be justified on the grounds that investors somehow possess perfect foresight of earnings growth and that they ignore short-run fluctuations. The infinite-horizon nature of the estimated model requires an unbiased estimate of perpetual earnings growth. It is possible, with the S&P data, to construct an estimate based on earnings growth as far back as 1926. The inclusion of a persistent recession and a major war, however, would likely result in a less satisfactory estimate of expected earnings.

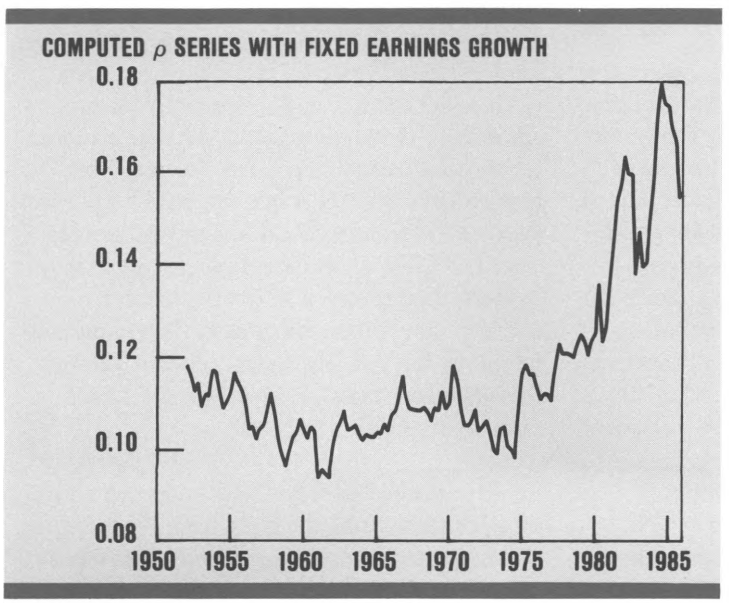


FIGURE 4

A time-series plot of ρ constructed with g fixed at its average value is presented in figure 4. The required real rate of return ranged between 10 and 13 percent from 1952 through 1974, with moderate deviation. At the start of 1975, however, ρ began to rise slowly and then sharply in 1981. It peaked at the end of 1981 and again at the beginning of 1984. For comparison's sake, setting g equal to zero over the entire sample period produces values for ρ ranging between 4.5 and 7.5 percent from 1952 through 1976, topping out at 14.2 percent in mid-1984.

IV. Analysis of Computed ρ Series

In this section, we analyze the behavior of ρ, computed with expected earnings growth fixed at its actual mean value. Our goal is to shed light on this component of equity valuation. By their nature, however, it is not possible to completely separate the implications of the various hypotheses discussed above.

The computed value of ρ appears to support Malkiel's hypothesis that ρ begins to rise in the mid-1970s due to the risk factors cited earlier. In addition, the rapid rise in 1981 could be explained by Bodie, Kane, and McDonald (1986), who concluded that there was a dramatic increase in the risk premium required in long-term bonds in the early 1980s. They attribute this to the switch in operating procedures by the Federal Reserve in late 1979.

The disinflation hypothesis presented earlier suggests that ρ should vary inversely with the level of inflation. In figure 5, we present plots of ρ and the rate of inflation. Note that the major upturns in ρ appear to coincide with the inflationary peaks occurring in 1974 and again in 1981. Smaller, previous inflation spikes do not, however, seem to be accompanied by any significant movement in ρ.

The same figure can be used to examine Hendershott's claims. Conspicuously absent is the hypothesized decline in ρ as inflation rises. The lack of noticeable downward movement in ρ during rising inflation eliminates much of the support for his arguments. His main conclusion, however, that ρ is tied to the real rate of return on debt, can now be addressed.

9 The average annual growth rate of adjusted earnings over the sample period was 17.01 percent. The growth rate of this series since mid-1983 has been so great as to completely dominate this figure. It was felt that the effects of this adjustment could not have been reasonably foreseen over much of the sample period and, in fact, should "wash" over the long run. We therefore chose to use the average annual growth rate of unadjusted earnings in the computation of ρ.

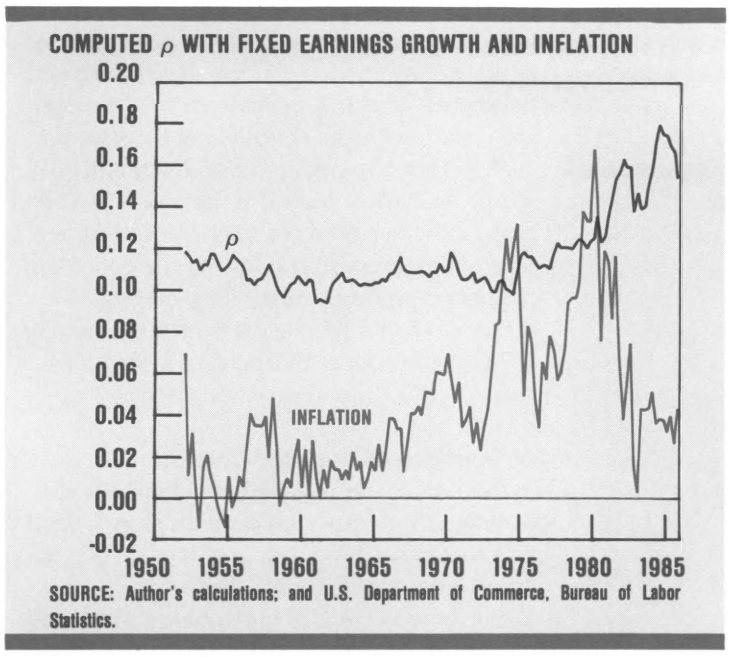


FIGURE 5

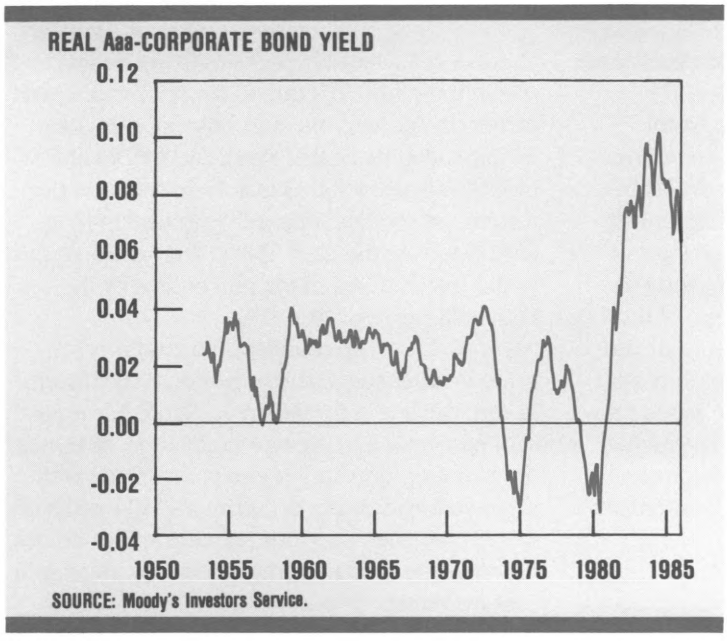


FIGURE 6

A plot of *ex-post* real long-term corporate bond rates is shown in figure 6. This figure was constructed by simply subtracting 1 from the ratio of the gross yield on Aaa-rated corporate bonds to the previous year's gross inflation rate at each date. Note that real required rates of return on fixed income securities reached unprecedented levels in 1981, the same year in which ρ significantly departs from its postwar behavior. Hendershott's hypothesis, therefore, appears to explain the sharp rise in ρ that occurred in 1981. However, it does not shed light on the moderate increase beginning in 1975, but it does help explain the slight decline in ρ that occurs between the end of 1971 and the end of 1974.

Though separate from the risk-related hypotheses, Benjamin Friedman (1986) claimed that an increase in the government deficit, such as that beginning in early 1981, would drive down the realized rate of return on equity relative to either short- or long-term debt, thereby increasing the required rate of return on a pure equity stream. This theory then suggests that the rise in ρ is a function of deficits, thus explaining the sharp rise in 1981.

Had we found no rational explanation for the behavior of ρ, we would have searched for evidence of measurement errors related to corporate earnings. For instance, Feldstein (1980) claimed that biases in the tax system, rather than inflation-induced valuation errors, could explain the poor performance of the stock market. In particular, Feldstein emphasized that corporate capital depreciation deductions are based on historical, rather than current, costs. In inflationary periods, with a rising price of investment goods, this implies that the real value of depreciation deductions declines. This, in turn, implies that taxable profits (net of depreciation deductions) rise, causing real after-tax profits to fall. Feldstein also pointed out that nominal rather than real capital gains are subject to capital gains taxes. This implies that even if the nominal value of equities increased at the inflation rate, the real after-tax yield on equities would decline. In contrast to Modigliani and Cohn, Feldstein viewed the stock market decline as a rational response to inflation.

Modigliani and Cohn, in response to the criticism of Feldstein, discussed the possibility of tax biases due to inflation. They noted that other analyses of the interaction of inflation and taxes have ignored the fact that firms are not taxed on the portion of returns used to depreciate debt. They argue that this offsets the decline in real after-tax profits that results from the decline in real depreciation deductions. They support this by noting that the share of corporate income paid as taxes has remained relatively constant in inflationary periods. In their empirical work, as well as in our construction of ρ, an adjustment factor constructed from the National Income and Product Accounts was used that attempts to correct reported earnings for depreciation and inventory distortions caused by inflation. The NIPA adjustment, however, may misstate the lagged response of tax shelters to inflation. In addition, the analysis is complicated further by the fact that much corporate debt is fixed-rate and thus debt yields do not adjust instantly to inflation expectations.

In figure 7 we present both unadjusted and adjusted reported four-quarter earnings per share using the NIPA data. For the early part of the sample period the two series are virtually identical. They begin to diverge at the end of

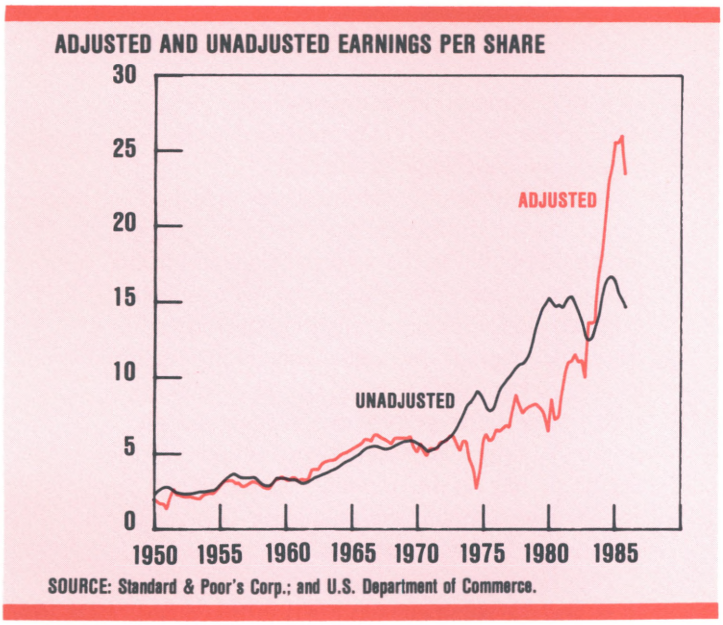


FIGURE 7

1972, with adjusted earnings falling somewhat below unadjusted earnings. The situation reverses dramatically, however, in 1983. At this point, adjusted earnings climb far above unadjusted earnings. Further study may shed light on the sensitivity of our results to the adjustment of earnings, especially for the period following 1981.

V. Conclusion

We conclude that equity prices respond rationally to such factors as real interest rates and risk. When we use the model of Modigliani and Cohn to compute the discount factor applied to a pure equity stream of a levered firm, we find no evidence of inflation-induced valuation errors. The evidence presented, however, is consistent with the hypothesis that *dis*inflation influences the risk premium applied to pure equity streams.

REFERENCES

Bodie, Zvi, Alex Kane and Robert McDonald, "Risk and Required Returns on Debt and Equity," in Benjamin M. Friedman, ed., *Financing Corporate Capital Formation,* Chicago: University of Chicago Press, 1986.

Feldstein, Martin, and Randall Morck. "Pension Funding Decisions, Interest Rate Assumptions, and Share Prices," in Zvi Bodie and John Shoven, eds., *Financial Aspects of the United States Pension System.* Chicago: University of Chicago Press, 1983.

Feldstein, Martin. "Inflation and the Stock Market," *American Economic Review,* vol. 70, no. 5 (December 1980), pp. 839-47.

Fons, Jerome S. "The Default Premium and Corporate Bond Experience," *Working Paper #8604,* Federal Reserve Bank of Cleveland, June 1986.

Friedman, Benjamin M. "Implications of Government Deficits for Interest Rates, Equity Returns, and Corporate Financing," in Benjamin M. Friedman, ed., *Financing Corporate Capital Formation.* Chicago: University of Chicago Press, 1986.

Grimm, Bruce T. "Domestic Nonfinancial Corporate Profits," *Survey of Current Business,* U.S. Department of Commerce, Bureau of Economic Analysis, vol. 62, no. 1 (January 1982) pp. 30-42.

Hendershott, Patric H. "The Decline in Aggregate Share Values: Taxation, Valuation Errors, Risk, and Profitability," *American Economic Review,* vol. 71, no. 5 (December 1981), pp. 909-22.

Malkiel, Burton G., "The Capital Formation Problem in the United States," *Journal of Finance,* vol. 34, no. 2 (May 1979), pp. 291-306.

Modigliani, Franco, and Richard Cohn, "Inflation, Rational Valuation and the Market," *Financial Analysts Journal,* vol. 3, no. 2 (March-April 1979), pp. 24-44.

_____, and Merton Miller. "The Cost of Capital, Corporation Finance, and the Theory of Investment," *American Economic Review,* vol. 1958, no. 3 (June 1958), pp. 261-97.

Shiller, Robert J. "Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?," *American Economic Review,* vol. 71, no. 3 (June 1981), pp. 421-36.

Townsend, Henry. "Another Look at the Modigliani-Cohn Equation," *Financial Analysts Journal,* vol. 42, no. 5 (September-October 1986), pp. 63-66.

Von Furstenberg, George M. "Corporate Investment: Does Market Valuation Matter in the Aggregate?," *Brookings Papers on Economic Activity,* No. 2: 1977, pp. 347-408.

Wadhwani, Sushil B. "Inflation, Bankruptcy, Default Premia and the Stock Market," *Economic Journal,* vol. 96, no. 381 (March 1986), pp. 120-38.

APPENDIX

Reestimation of Modigliani and Cohn's Model

In this section, we describe our attempts to replicate the results of Modigliani and Cohn and then to reestimate their model, extending the sample period through 1984.

Modigliani and Cohn estimated the following regression, which is implied by expression (13), after taking the log of both sides:

$$s(t) = a0 = w1(L)\Pi(t) + w2(L)DIV(t) + a3\ w3(L)[LF/E](t) + a4DVF(t) - \beta w4(L)R(t) + \gamma w5(L)P(t) + u(t)$$

The variable L is the lag operator and the parameters $w1$ through $w5$ represent coefficients on the lagged terms of the five forecasted variables. The distributed lag, $w1(L)\Pi(t)$, embodies the assumption that expected, or *ex-ante,* profits equal a one-sided distributed lag of past profits. Profits were measured as described in the text and in Modigliani and Cohn. Although it is not unusual to view expected dividends as influencing stock prices, Modigliani and Cohn include a distributed lag of dividends, $w2(L)DIV(t)$, on the grounds that dividends provide information about future profits. They then restrict the coefficients of the distributed lag on dividends, so that a change in dividends has no permanent effect on firm value, given the history of profits. Dividends were measured as dividends per share for the issues in the S&P 500, adjusted as described by Modigliani and Cohn. $w3(L)LF/E$, a distributed lag of the ratio of the labor force to employment, is included to provide a cyclical adjustment to the ability of past profits to predict future profits. The term $DVF(t)$ is included as a measure of the risk premium entering the formulation of ρ. Modigliani and Cohn measured DVF as the 15-year moving-average deviation of the unemployment rate from 4 percent. We chose instead to use a 12-quarter moving-average. The distributed lag on the nominal interest rate, $w4(L)R(t)$, and the distributed lag on inflation, $w5(L)P(t)$, are included to measure the real rate, $r(t)$, also a component of ρ. $R(t)$ is measured as the new issue yield on AA corporate bonds. $P(t)$ is measured as the annual percent change in the CPIU.

We used the current value and seven lagged values in each distributed lag. This choice of lag length differed from that of Modigliani and Cohn, but seemed only equally arbitrary. We maintained the following restrictions regarding the form of the distributed lags: a) the coefficients on profits sum to one, b) the coefficients on dividends sum to zero, c) the distributed lag on LF/E is quadratic, d) the distributed lag on dividends is linear, e) the distributed lag on the nominal rate is quadratic, and f) the distributed lag on inflation is quadratic with the endpoints constrained to equal zero.

The parameters to be estimated are $a0$, $a3$, $a4$, β, γ, and the parameters in the distributed lags. The theoretical model of Modigliani and Cohn implies that the coefficient on the distributed lag of inflation, γ, should equal d/K, where d is the debt-equity ratio and K is the capitalization rate. Their estimate of γ, -0.08, differs from a computed value of d/K, 0.05. Thus, an increase in expected inflation reduced market values, although this should not have been the case if investors had been rational. In fact, Modigliani and Cohn calculated that a one percent increase in inflation would reduce the market value of equities by 13 percent. Thus, the market had been drastically undervalued due to inflation-induced valuation errors.

When we attempted to replicate the results of Modigliani and Cohn, over the same sample period, we estimated γ to be .015. When the sample period was extended through 1984, however, the estimate of γ was -0.025. If the misvaluation of equities was being eliminated, the estimate of γ over the longer period would have been closer to the theoretically predicted value (d/K) than for the shorter period. Since our results not only differed from those of Modigliani and Cohn, but indicated worsening misvaluation, we chose to consider a different approach.

The Collapse in Gold Prices: A New Perspective

by Eric Kades

"If all men were rational, all politicians honest and we had a world central bank issuing a single currency that was universally acceptable, then gold would drop to $20 an ounce—and be overvalued at that."

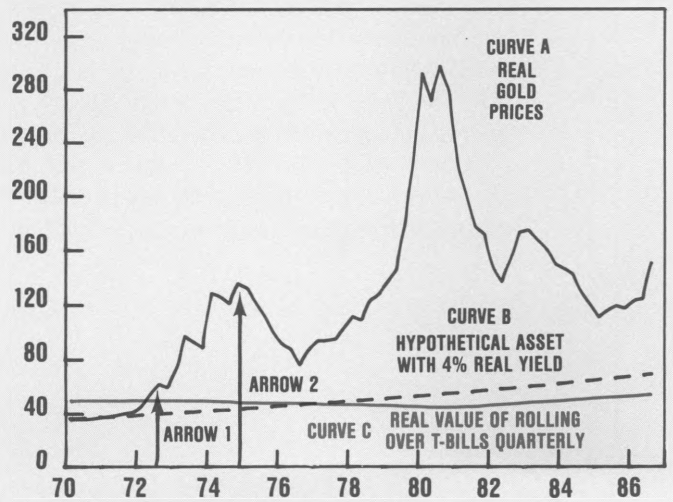
— Andre Sharon, gold analyst, quoted in *Newsweek,* Dec. 16, 1974; as quoted by George Seldes in *Quotable Quotations.*

Eric Kades is an analyst in the Financial Strategies Group at Goldman, Sachs, & Co. He is a former analyst at the Federal Reserve Bank of Cleveland. The author thanks Brian Gendreau of Morgan Guarantee Trust Company, Stephen Salant of the Rand Corporation, and Mark Sniderman of the Federal Reserve Bank of Cleveland for helpful comments and corrections.

Introduction

The daily summaries and analyses of the gold market that appear in most newspapers support Mr. Sharon's assertion. The press invariably attributes gold price movements to political uncertainty, gyrating monetary policies, inflation hedging, and international liquidity concerns. This view implies that the demand for gold is highly volatile, subject to coups, sudden shifts in central bank behavior, oil flow interruptions, and other jolts to the world economy.

GOLD PRICE AND INTEREST RATE TRENDS
Real value of assets, in 1970 dollars



SOURCE: London P.M. Gold Fixing; Consumer Price Index: U.S. Department of Labor, Bureau of Labor Statistics; and Board of Governors of the Federal Reserve System.

FIGURE 1

If Mr. Sharon and the press were right, then economists would have little to contribute to an analysis of even long-term movements of gold prices, or to forecasts of price trends. These activities would be better left to political experts, to central bank analysts, and to other savvy observers in areas that are likely to generate surprises affecting gold prices. There would be no point in statistically estimating a demand function for gold, since demand for gold would be always be fluctuating randomly, not moving systematically.

This conventional explanation of gold price movements is essentially a superficial one. While unexpected political and economic events undoubtedly influence daily gold prices, such events cannot explain long-run trends in gold prices. Before the Bretton Woods international monetary system began to crumble in 1968, the price of gold was fixed at about $35 an ounce. The real price of gold, that is the nominal, or observed price divided by a price index, has followed two distinct trends since 1968 (see curve A in figure 1).[1]

From 1969 to 1981, the real price of gold rose rapidly, except for a few brief, but sharp, price dips, and for one extended slide. From 1981, until this year, the real price of gold fell —

1 The price index in this case is the CPIU. We study the real price to correct for changes in the purchasing power of the dollar.

rapidly at first, then more gradually. This year has witnessed renewed strength in gold prices.

As confusing as these long-run trends might seem, economists can offer an explanation that draws from previous studies and that is rooted in the unique characteristics of the gold market and of gold itself. In this study, we explain gold price movements in terms of a model for exhaustible resources, describing the qualities of gold that account for its spectacular price rise from 1968 to 1980. We show why models offered in previous studies must be modified to explain the *decline* in gold prices since 1981. The recent rise in gold prices can be interpreted as supporting the perspective we present.

I. The Conventional Supply and Demand Analysis

If viewed as any other good, gold price movements would be analyzed in terms of conventional supply and demand. The key factor in supply would be the marginal cost of extracting additional ore from the earth (or the cost of recycling gold). These costs would follow predictable patterns and might be estimated effectively. More analytical difficulty could be expected to arise from unstable demand. While the demand for gold by industries and jewelers is stable, other buyers (inflation hedgers and liquidity seekers) have erratic gold-buying habits.

This demand/supply view of the gold market turns on gold's historical role as a defense against inflation and posits that a predictable relationship between inflation and real gold price movements overwhelms any random influences from other types of gold buyers. International liquidity measures (money supplies and related stocks) and exchange rates also enter conventional gold-pricing models, but these may be thought of as leading and coincident indicators of inflation among nations. But the contemporaneous correlation between inflation and gold prices is weak (see figure 2) and formal statistics show no strong relationship between gold prices and past inflation.

There are more fundamental arbitrage arguments against a *systematic* and reliable relationship between gold prices and inflation. If the real rate of return to gold were *systematically* lower than riskless assets during deflationary times, no rational investor would include gold in his portfolio while prices fell. Since someone must hold existing stocks of gold, the rate of return (i.e., price changes over time) would be bid up to induce agents to hold existing stocks. Similarly, if gold has an extraordinary rate of return during inflations (with little or no risk) then this return would be bid down by investors. Thus, gold cannot yield *systematically* more/less than riskless government bonds during inflations/deflations. Such a regime would not be in equilibrium since people will want to increase/decrease gold holdings and thereby shift the price of gold. Gold may well serve as an effective *hedge* against inflation by *preserving* purchasing power, but this factor cannot explain the long-term trends observed in the high real return to gold.

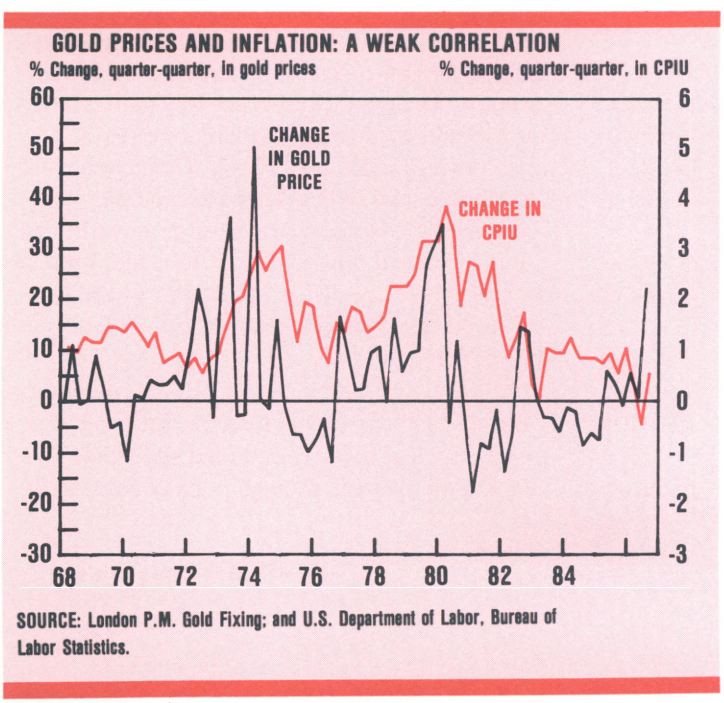


FIGURE 2

II. Gold: An Exhaustible Resource

Gold is literally an *exhaustible* resource. Thomas Wolfe, a former director of Treasury precious metals operations, noted, "[g]old is among the few minerals that could reach a critical supply situation within this century, taking account of available reserves above and below ground."[2] World stocks have declined sharply, on net, since 1968 (see figure 3). Thus, gold's very scarcity requires a unique method of determining its price. Economists do have a valid technique, which has existed since the 1930s, for analyzing prices of exhaustible resources.[3] An ingenious modification

2 See Stephen Salant and Dale Henderson, "Market Anticipations of Government Policies and the Price of Gold," *Journal of Political Economy*, vol. 86. no. 4 (August 1978), pp. 627-48.

3 See Harold Hotelling, "The Economics of Exhaustible Resources," *Journal of Political Economy*, vol. 39. no. 2 (April 1931) pp. 137-75.

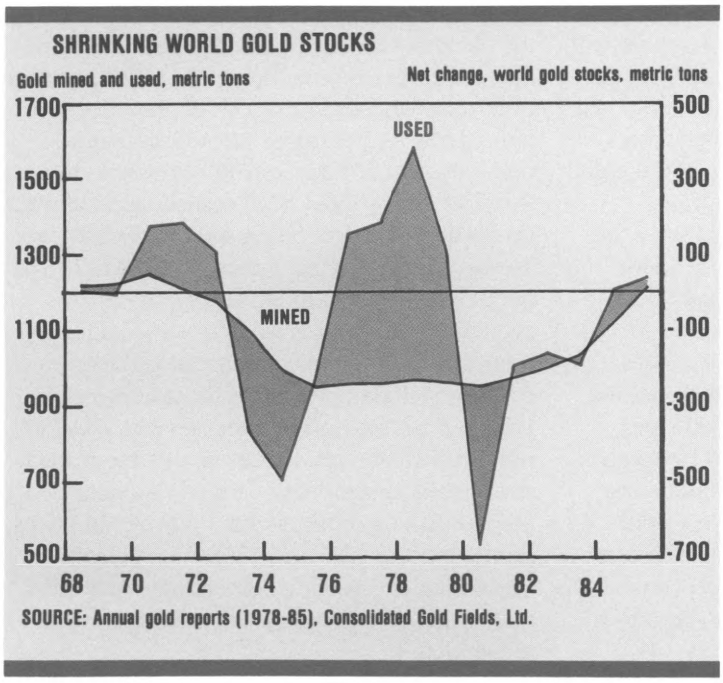


FIGURE 3

of this technique, by Stephen W. Salant and Dale W. Henderson, explains the major gold price movements from 1968 to 1981, but does not fully explain the decline in gold prices from 1981-1986, which is the period that interests us here.[4] We will use a simple allegory to explain how exhaustible resources should be priced.

III. Pricing a Depletable Resource: A Simple Model

A finite amount of "manna" is distributed among some lucky individuals. Once eaten, the manna is gone forever. What direction should manna prices take? Think of manna as an asset, just like stocks and bonds; we are concerned with only the investment characteristics of manna, not with the final consumers. Each asset in the economy has an annual rate of return. We expect the rate of return to be higher for more risky assets. Investors will demand a higher rate of return on risky "junk" bonds of an indebted corporation, for example, than they will require to hold the virtually riskless bonds and bills of the U.S. Treasury bearing a riskless interest rate r_b.

With a fixed, known supply and stable demand, manna is a riskless asset. Therefore it must yield the same rate of return as the riskless interest rate on government securities, r_b. The only way it can yield such a return (since it pays no dividends and earns no interest) is for its price to rise at a rate equal to the riskless interest rate. If the price rose faster, nobody would sell it, since it would be a superior investment over any time horizon. This price path could not constitute an equilibrium, however, since eventually the constantly rising price would cause all demand for manna to be choked off, and yet none would have been consumed. The good's value would be too high for it to be liquidated.

On the other hand, if the price of manna rose more slowly than r_b, agents would not want it in their portfolios; in the rush to sell it, the price would fall precipitously. Then the price would jump from a very low level in the initial period (as everyone rushed to sell) to infinity in the next period, because people would consume the manna, driving the supply toward zero. However, then an investor who bought at the low price in the first period and held on to the manna could charge a very high price in the second period, earning a rate of return above r_b. Only a path with prices increasing at the rate r_b leaves people with no incentives to change their behavior and, thus, is the only equilibrium for pricing a simple exhaustible resource like manna.

Curve B in figure 1 indicates such a price path. It shows how the value of a bond equal in price to an ounce of gold in 1965 would have appreciated at a constant real interest rate of 4 percent (arbitrarily chosen). It is easy to see that from 1965 to 1981 gold yielded a higher return. Curve C shows the *ex post* real rate of return—the real value an investor would have achieved by purchasing U.S. Treasury bills and rolling them over quarterly. Gold even more easily dominated this investment option over the period.

IV. Gold: A Special Case

No resource exactly duplicates the simple price path illustrated in curve B of figure 1. Unexpected events will change the supply of or demand for a good. These jolts, not readily addressed by economic arguments, affect the price of a good randomly, increasing the variance without changing the long-run path.

Also, every commodity has its own peculiarities that alter its price trend systematically. Several such factors operate in the gold market and are often cited as important determinants of price trends.

The costs of mining gold, for example, rose between 1965 and 1981 and contributed to at least part of gold's price increase. However, it can also be argued that mining costs could not be responsible for price increases in a market, like the gold market, where there are speculators.

Rising extraction costs do not prevent speculators from buying gold in one period and selling it in the next. Because speculators do

4 See "Market Anticipations of Government Policies and the Price of Gold?"

not pay extraction costs (and assuming storage costs are constant), competition among them will prevent the rate of gold price increases from exceeding the riskless rate of interest.[5] The gold market unquestionably includes many speculators.

Another salient feature of the gold market is South Africa's dominant, almost monopolistic role in gold production. Since the price of gold began to rise in 1968, the South African share of production has averaged near 75 percent, although it has fallen moderately in recent years. Such hegemony can raise prices above competitive levels, but, like rising production costs, cannot account for observed rapid increases in gold prices. Any attempts by South Africa to raise prices faster than r_b would create arbitrage opportunities that would force prices back down. Speculators would buy gold in one period and then, being willing to accept a rate of return r_b, would undersell the South Africans in the next period.

Salant and Henderson conclude that the only valid special factors in the gold market are the huge stocks governments hold and, particularly, the perceptions of speculators about what buying or selling actions governments will take. This, they argue, causes the price of gold to move systematically at variance with the simple exhaustible resource explanation.

To see how this matters, think about the amount of gold available to satiate demand in a given period. Production levels will be relatively stable, because construction of large mines takes a long time. However, governments hold huge stocks of gold (now about 40 years' worth of current industrial, artistic, and jewelry demand; in 1970, governments held 25 percent more). If they decide to sell a significant amount of gold in a given period, the price will drop sharply.

The "threat" of government sales means that *gold can no longer be considered a riskless asset,* since there is a chance that government actions will have a severe impact on its price. Risky assets must give higher yields, on average, to compensate their owners. Comparing curve A (actual real gold price) with C (actual price trend for real return to three-month Treasury bills), strikingly illustrates that gold did indeed command a return higher than the riskless interest rate from 1968 to 1981.

There were a few exceptions, when the price dropped precipitously for short periods of time. These occasional price dips, however, fit precisely into the scenario that Salant and Henderson present. They are the announcement dates of government sales or news leaks of the likelihood of such sales. These events illustrate the riskiness of holding gold in the presence of government stocks that can depress prices temporarily. For example, arrow 1 in figure 1 marks the first announcement of possible International Monetary Fund (IMF) nation sales; arrow 2 shows the price decline caused by the first U.S. Treasury auction of gold since World War II. The price decline that lasted from 1975 to 1976 occurred while gold's role in the international monetary system was being revised. These changes included provisions for large sales of IMF gold, permission for member nations to sell significant quantities of gold on the free market, and a major de-emphasis of gold's monetary function. All these factors held down gold prices during most of 1975 and 1976. When direct depressing effects ended, prices rose again, and gold achieved superior rates of return—much higher than r_b.

Salant and Henderson's explanation for the trends in gold prices is an elegant and convincing one for the period from 1968 (which marked the end of gold price-fixing) until 1981, but it breaks down after 1981. There has been a striking change in the behavior of gold prices since 1981. They fell, first sharply, then more gradually, with only short-lived reversals. In 1986 they again began to rise sharply. How, if at all, can these trends be reconciled with the relationship between the price of gold and with sales of government supplies of gold described above? The starting price of an exhaustible resource holds the key to our explanation.

V. Initial Price and Expectations of Demand

The *initial* price of a depletable resource plays an important role in its price behavior. The price must increase at the rate r_b (a risky asset like gold, of which governments hold large stocks, increases at a rate higher than r_b. In a "perfect world," the initial price will be set so that the last ounce of gold is used via transactions completed up along a unique price path starting at the initial price and increasing at the set rate.

A low initial price would result in greater demand at every date along the price path and the supply of gold would be depleted at a price that didn't extinguish demand. Conversely, a high initial price would mean less demand for gold in each period; demand would drop toward zero, and gold stocks remain. Profits for owners in both cases would be lower than if the equilibrium price path were to emerge, so market forces tend to seek this unique initial price and price path.

To calculate the correct initial price, it is essential to estimate the demand curve

5 Gold that cost more to extract would not be mined until the price rose sufficiently to justify the expense.

for gold—that is, what demand will be for all prices. Incorrect estimation of the demand curve would lead to incorrect setting of the initial price and would necessitate later adjustment of the price to reflect the true demand.

VI. An Unexpected Price Path

Suppose that, in 1968, market participants estimated a demand curve for gold, based on past demand and existing world stocks. In doing so, they implicitly calculated that if prices began to rise at a rate equal to r_b (plus some risk premium), the world's supply of gold would be exhausted just as the price rose to levels that would choke off demand.

Market participants had many opportunities to observe demand in the price range from $35 to $100 an ounce (in real 1977 dollars) from at least the beginning of the twentieth century until 1978. Although we do not have the data to plot these points precisely, we assume for this example that they fit a linear demand curve fairly well, as shown in figure 4. Based on these observations, gold speculators postulated that the same solid line that approximated

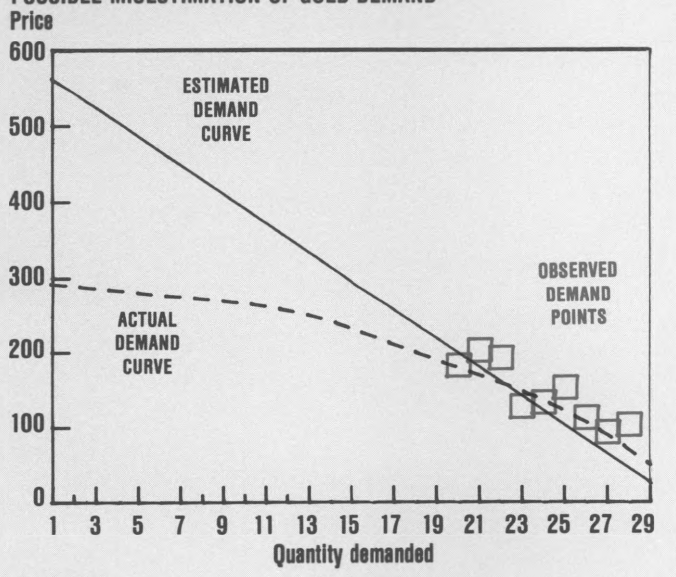


FIGURE 4

demand for prices from $35 to $100 an ounce would also be valid at higher prices. However, if true demand were represented by the broken curve, then it is obvious how this misestimation could produce an unexpected flagging in demand that would, in turn, cause the price decline in gold observed since 1981.

Figure 4 illustrates just one of many ways that agents could have incorrectly estimated the demand curve for gold, given that demand at the very high prices that prevailed in 1980-81 had no precedent. But why has the decline in the real (as well as the nominal) price of gold been so extended? Market participants are actively revising their estimates of demand at the prices where they first began to make serious errors, in the $200 to $400 range. First, the price fell precipitously as all speculators temporarily liquidated stocks in the knowledge that prices would fall. Most speculators were surprised when, after this initial price drop, demand was still too weak to support new price increases (consistent with the exhaustible resource model).

Since 1983, when the price fall moderated, the market may be said to have been groping for a price path that would lead to a depletion of gold just as demand chokes off. Demand was weaker than expected in the intermediate price range, and so the price continued to edge downward. The recent rise in gold prices may indicate that the bottom has been found, and that gold will yield superior returns.

Perhaps a more fundamental question is: why did people misestimate the demand for gold in the first place? Certainly there are many plausible explanations, and this paper does not attempt to establish one as being more correct than another. However, one possible explanation is that their information was inadequate and inappropriate. People had virtually no basis for estimating the entire market demand curve, since the price had been more or less fixed for over 25 years. People did not even have estimates of the average expected demand at higher prices, let alone the variation to be expected about this average. Their estimate of market demand proved correct for prices that were not too far from observed values, but people systematically overestimated demand at higher prices. Oil market analysts undoubtedly had similar difficulties forecasting demand after OPEC suddenly tripled prices in the early 1970s.[6]

6 Salant notes that rising real interest rates, along with incorrect demand forecasting, can help explain why gold prices dropped after 1980. We have implicitly assumed a constant real interest rate. Salant points out that if, for whatever reason, the real interest rate rises, the price of gold would initially fall before increasing at a faster rate. Why is this so? A higher real interest rate implies that gold prices must rise more rapidly. If no price decline occurred when real interest rates rose, the new *higher* gold price path would induce lower demand at every date than the original price path. But the original price path was set such that supply would be depleted just as a sufficiently high price choked off demand. If no price drop occurred when a higher interest rate prevailed, the stock of gold would not be exhausted; some owners would be left holding gold when high prices extinguished demand. This is not an equilibrium; such a prospect forces prices to jump down when the interest rate rises.

Conclusion

Two distinct regimes explain the unique behavior of gold prices since 1968. Between 1968 and 1981, prices increased according to the Salant and Henderson analysis; based on prices actually prevailing during the 1968 to 1981 period (as high as $200 an ounce in real 1977 dollars) estimates of the demand curve for gold were roughly correct. However, incorrect forecasts of gold demand at higher prices meant that the price had to fall. The initial precipitous decline reflects the first reaction to this prediction. The continued mild slide indicated that the market was edging down the demand curve in search of the price that fits the Salant and Henderson explanation of gold price determination. The turnaround in gold prices may well be telling participants that demand has been reestimated with enough confidence to justify a renewed upward trend in gold prices.

"Don't Panic"[1]: A Primer on Airline Deregulation

by Paul W. Bauer

The old dictum says that if the Devil did not exist, the Church would have had to invent him. Similarly, if the regulator didn't exist, the airline industry would have had to invent him—and did in 1938. A current question is what would happen to the industry were it totally deregulated. One thesis is that there would be a rush by existing and new entrants to those routes thought to be profitable. Other routes would be abandoned. Price competition would be destructive. With the essential link between economics and safety there would be an inevitable major air disaster, possibly involving a prominent Member of Congress. Public outcry and congressional responses would lead to the re-establishment of regulation. Since this was the sequence of events in the mid-30's, why re-learn that lesson? This thesis has been challenged, but the lesson of history . . . cannot be totally ignored.

Secor D. Browne, Chairman
Civil Aeronautics Board
(January 1972)[2]

Paul W. Bauer is an economist at the Federal Reserve Bank of Cleveland.

The author would like to thank Randall W. Eberts, Joe A. Stone, and others who provided useful comments on an earlier draft of this paper.

Introduction

Former Civil Aeronautics Board (CAB) Chairman Browne's statement 15 years ago can scarcely be interpreted as an unqualified endorsement of the government's current policy of airline deregulation. It does remind us, however, that the issue of airline regulation has been controversial for quite some time.

The Civil Aeronautics Act (CAA) of 1938, enacted to counteract the alleged conditions of competitive instability of an industry then in its infancy, began 40 years of pervasive government regulation by the now-defunct CAB. With passage of the Airline Deregulation Act (ADA) of 1978, the federal government completed an about-face in policy and reintroduced competitive forces into the market.

For eight years now, the airline industry has been experiencing a great deal of turmoil, as evidenced by the large number of entries, mergers, and bankruptcies. Much of this turmoil, however, is not the result of deregulation, but rather of the fuel price increase in 1979, of the recession in the early 1980s, and of the air traffic controllers' strike in August 1981. Even so, the regulation debate is heating up again as the events predicted by Mr. Browne seem to be unfolding—with such examples as the recent bankruptcy of Frontier Airlines, the financial problems of People Express and Eastern Airlines, and the crash of the Aeromexico airliner in southern California in August 1986.

This paper analyzes the conditions that prevailed under CAB regulation and that led to the Airline Deregulation Act of 1978. These conditions are contrasted with the effects of deregulation observed so far. Finally, an attempt is made to predict the future evolution and performance of the U.S. airline industry under deregulation.

I. The U.S. Airline Industry Under CAB Regulation

Between 1938 and 1978, the CAB maintained strict control over the two most important decisions airlines had to make: where to fly and how much to charge. This meant that airlines could only compete with one another by offering a higher quality of service (primarily more frequent flights

1 Sound general advice from *The Hitchhiker's Guide to the Galaxy* by Douglas Adams.

2 Foreword to R.E.G. Davies' *Airlines of the United States Since 1914*, Putnam & Company Limited, London (1972).

and other amenities). Studies have shown that CAB regulation led to more frequent flights and to lower load factors (the proportion of seats on a flight that are filled by paying passengers) than would be normal in a competitive airline industry.[3]

Since these actions resulted in higher costs for the airlines, and since the CAB was charged with maintaining the financial health of the industry (that is, preventing losses), it follows that fares were higher. In fact, the interstate carriers subject to CAB regulation marked up fares 20 to 95 percent more than the intrastate carriers not subject to CAB regulation for similar routes.[4] The General Accounting Office (GAO) estimated that passengers could save up to $2 billion dollars or more per year with competitive fares.[5]

II. The Theory Behind Deregulation

Given fare markups of these magnitudes, why were the airlines' earnings so mediocre? The answer appears to be that regulated industries do not have sufficient incentives to control costs. Given the CAB's mandate to maintain the health of the industry by raising fares whenever the airlines experienced hard times and the lack of a threat of competitive entry (the CAB had not allowed the formation of a single new trunk airline from 1938 to 1978), a strong prima facie case exists for inadequate cost control. Using data from 1972 to 1978, Bauer (1985) found that, on average, airline costs during that period were 48 percent over the minimum cost of providing the same service.

Another example of the poor incentive structure can be found by analyzing labor costs. Providing a service product—transportation between two points—airlines could not stockpile their output in anticipation of a strike. Any output diverted by one carrier (either to other carriers, or to other transportation modes) as a result of the strike is a permanent loss to that carrier. Further, even when the strike is settled, the airline may lose some of its customers to other carriers. Regulated airlines could not offer large discounts and free flights to lure their customers back, as United Airlines did after a strike in 1979. Under CAB regulation, strikes were very costly to the airlines, but higher labor costs could be absorbed by CAB fare increases or CAB approval to enter some profitable new route. Thus, there was little incentive for airlines to endure strikes.

Given the evidence on fare markups and the suspicions about airline inefficiency, proponents of deregulation became convinced that elimination of CAB regulation, and a move towards more competition in the industry, would be beneficial to travelers and, ultimately, to the industry itself. Two basic tenets drive the model of the industry that proponents of deregulation had in mind: one, that the minimum efficient scale size is reached at a relatively low level of output and, two, that new entry and the threat of new entry into the industry would ensure sufficient competition to hold fares close to marginal cost and only allow firms to earn a normal profit.[6]

Numerous studies performed prior to deregulation, using various data sets from the late 1950s forward, found that larger airlines had no significant unit-cost advantage (measured in passenger miles) over smaller airlines. This research implied that there was plenty of room in the U.S. airline industry for anywhere from 20 to 100 efficiently sized airlines (see White [1979]), and that there was little chance of concentration increasing in the industry if it were deregulated.

The second tenet, that freedom of entry would severely limit any market power that an airline may have, was being strongly supported by the new theory of contestable markets (see Baumol, Panzar, and Willig [1982]). Simply stated, this theory predicts that if market entry and exit involves no irrecoverable costs and can occur quite rapidly, the threat of entry is sufficient to ensure that firms in this market earn no more than a normal profit.

The following illustrates how this result occurs. Suppose the firms in a contestable market decided to collude and to raise their prices. Although the strategy might work in the very short run, soon new firms not party to this agreement would recognize the opportunity for above-normal profits and would enter the industry, driving prices back down. In a contestable market, even a monopolist would thus earn a normal profit, because if it tried to take full advantage of its monopoly power to earn more than a normal profit, another firm would enter and charge the lower price, capturing the entire market for itself.

Clearly, not all industries in the economy can be considered contestable (the auto industry, for example, is definitely not). However, deregulation proponents considered

3 Douglas, George W. and James C. Miller, (1974) *Economic Regulation of Domestic Air Transport: Theory and Policy*, Brookings Institution, Washington, D.C.

4 T. E. Keeler, "Airlines Regulation and Market Performance," *Bell Journal of Economics* 3 (Autumn 1972), pp. 339-434.

5 General Accounting Office, Report to Congress, *Lower Airline Costs per Passenger Are Possible in the United States and Could Result in Lower Fares*, February 1977, p. 11.

6 A normal profit is the minimum return required to keep the firm from shifting resources out of the industry.

the airline industry a good candidate for contestability—once the artificial barriers to entry created by the CAB were eliminated.

The following market characteristics were considered to promote contestability:

- Inputs used by the airline industry are all relatively mobile when compared to most other industries. Labor, energy, and materials can either be employed or let go on fairly short notice, as in most industries, but capital is much more mobile than in almost any other major industry.
- Airlines can quickly shift planes from one route to another as the need arises. Further, since there is a ready secondary market for used aircraft—in fact, many carriers rent a significant portion of their fleets—planes are fairly mobile from one carrier to another.
- Ground facilities are usually rented, making them fairly disposable (acquisition is another matter, and will be discussed later).

These properties are thought to make it relatively easy for incumbent airlines to begin service on new routes, so that if fares are too high on a given route, other airlines will enter those markets at lower passenger fares. These properties are also thought to facilitate the start-up of new airlines if existing lines are making more than a normal profit.

Thus, according to the contestable market view, there was little to fear on the part of consumers from airline deregulation. Even if the industry did evolve into a handful of firms, the contestable market theory predicted that they could only earn a normal profit and fares would be as low as possible.

In summary, the proponents of deregulation predicted sharply lower coach fares, as fare markups would be bid down and airlines would strive to reduce their costs in the face of observed and potential competition. There would be some deterioration in service quality as flight frequencies would be reduced. However, this would in turn lower airline costs (by increasing load factors), thus further lowering fares, and passengers would receive the fare-service mix that they prefer. It was felt that there was no need to worry about increased concentration in the airline industry, because the minimum efficient scale would be small enough to make room for many carriers. Besides, the threat of entry would be sufficient to hold fares down and service quality up, even on routes with few carriers.

III. The Effects of Airline Deregulation

The actual effects of airline deregulation, while being generally beneficial to date, have not materialized precisely as the proponents predicted. This divergence of prediction and reality can be traced to changes in the airlines' operating strategies that were induced by the increased freedom given to them by the elimination of CAB regulation. These changes in strategy occurred in the two areas mentioned earlier: where to fly and how much to charge. Market competition seems to have induced even more innovation than industry experts foresaw, leading to predominately beneficial changes in airline behavior.

Fares

As the CAB's authority over fares was diminished, the airlines gradually developed a more complex fare structure to replace the relatively simple first-class and coach-fare structure that existed under regulation. While an element of price discrimination certainly exists, most of the variation in fares is based on differences in the cost of serving the various classes of passengers.[7] Fares are lower for travel outside the periods of peak demand. Examples include flying on weekends, flying in the middle of the day or late evening, and flying to locations that are out of season. A prime example of fare differences based primarily on cost is found between those who can book and pay for tickets in advance and those who cannot. It is costly for airlines to fly planes with empty seats, yet they intentionally have some slack in their systems so that they can accommodate last-minute travelers—for a higher price.

These pricing strategies have enabled the airlines to increase both traffic and revenue far more than if a uniform pricing policy had been followed. The increase in the industry's revenue passenger miles (RPM) and average load factor are plotted over time in figure 1. Both have increased since deregulation, although the effect of the recession in the early 1980s is clearly evident. Traffic increased 33 percent just from 1977 to 1979.

As a result of this shift in pricing strategy, the average fare that passengers actually paid (adjusted for inflation) has fallen about 20 percent in the last 10 years, even though the standard coach fare has fallen very little. Though this is a far cry from the drop that had been expected given the fare markups and inefficiency that existed under regulation, it does represent a

7 For example, whether one stays over a Saturday night on a round trip has no effect on the airline's cost of providing the service, yet it provides a very useful screening device enabling the airlines to charge higher fares to business travelers (who generally cannot meet this restriction) and lower fares to pleasure travelers (who usually can). Thus the airlines can price discriminate between the two classes of consumers, taking advantage of the business travelers' higher price elasticity of demand (and the leisure travelers' lower elasticity of demand) to increase their revenue and profits.

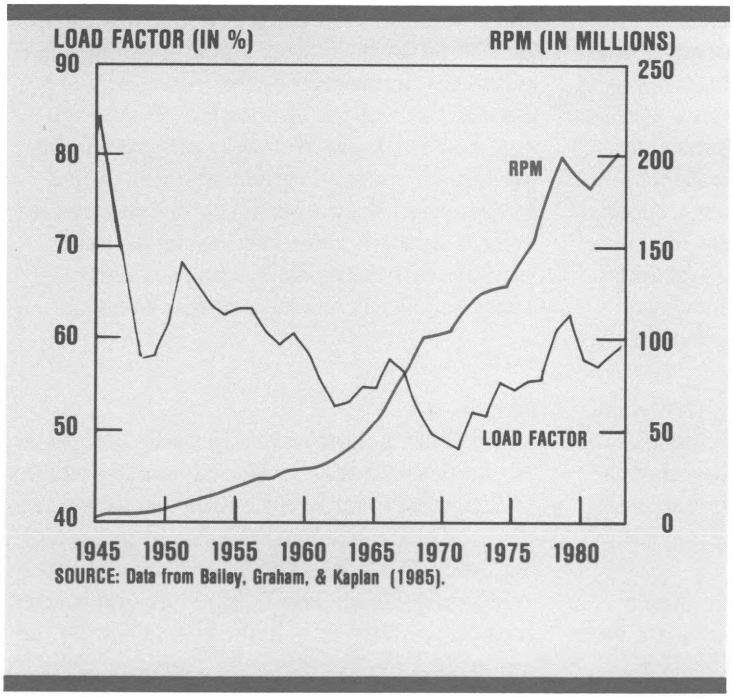


FIGURE 1

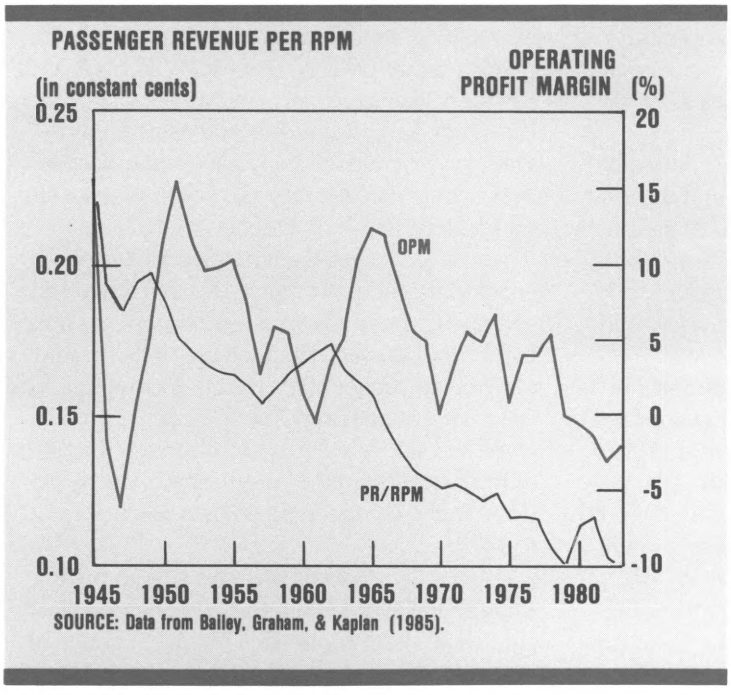


FIGURE 2

considerable savings to travelers. A measure of the average fares paid by travelers, the average passenger revenue per RPM, is plotted along with the average operating profit in figure 2.

All parties benefited to some extent by this new fare structure. The super-low fares enabled many leisure travelers to take trips they would not have considered before; business travelers gained by the increase in flight frequency on most routes (as a result of the increase in traffic) and by the lower fares (for those who could qualify for the discount fares); and the airline industry was able to increase its profits over what they would have been under regulation as the increase in load factors lowered costs.

Routes

The other fundamental change in the airlines' strategies concerns the decision of where to fly. Few people inside or outside the industry foresaw the shift of the airlines to what is now known as a hub-and-spoke system. Since deregulation, instead of serving a hodgepodge of routes as dictated by the CAB, airlines organized their routes so that most of their flights now converge on one or two hubs. These hubs collect traffic from the "rim" cities, then the passengers change planes at the hub to go out on other flights to their final destinations. The potential benefits of this system were demonstrated to a small extent by Delta Airlines, which had a hub in Atlanta even under regulation.[8]

The hub-and-spoke system has enabled airlines to increase their load factors on flights both into and out of the hub, thus lowering their costs and enabling them to lower their fares. An important side benefit is that flights can be scheduled more frequently because of the higher traffic density. Thus, instead of flight frequencies decreasing under deregulation, as was generally predicted, they actually increased. Passengers are also more likely to be able to complete their entire trip on one airline (which is advantageous to the airlines) and to avoid the inconvenience of changing planes at busy airports (which the passengers like). Another benefit is that passengers can fly from almost any city to almost any other city without having to endure multi-stop flights. Usually a one-stop flight can be found, and routes with sufficient traffic density still receive nonstop service.

How much are these innovations worth to consumers? Morrison and Winston (1986) estimated the total benefit of deregulation to consumers to be $5.7 billion a year. For the average passenger, the benefits per trip were $11.08 and came from the following sources: a gain of $4.04 from lower fares, a loss of $0.96 from slightly increased travel time, and a gain of $8.00 from increased flight frequency. Morrison and Winston further estimate that airline profits would have been $2.5 billion higher than they were under regulation. Thus, airline earnings would have

8 The joke then was, "It does not matter whether you are going to heaven or hell; you have to go through Atlanta first."

been even worse than they actually were (as reported in figure 1) had CAB regulation continued. These are substantial aggregate benefits.

Passenger Concerns

Even so, the gains of deregulation have not been shared equally by all travelers and, in fact, some may be worse off. Travelers who do not qualify for the discount fares and who must pay the full coach fare are probably worse off, unless the benefit from the increase in flight frequency is sufficient to offset this effect. Also, due to the oversupply of wide-body jets, which are ideally suited to carrying passengers coast to coast, fares for flights between 2,000 and 2,999 miles have fallen much more than other fares, so that travelers on these routes have benefited proportionately more than travelers on shorter routes. This is a temporary benefit, however, and will last only until the airlines adjust their fleets. Finally, travel time for most flights involving large hubs has increased due to the increase in traffic.

One of the early concerns of opponents and even of some supporters of deregulation centered on the availability of air service to small communities. Provision was made in the ADA for subsidies to help support air service to small communities for a period of up to 10 years, but many communities were not covered by these provisions. However, most small communities, far from losing service, have gained service. In general, hedgehopping, multi-stop flights have been eliminated (lowering travel time), and flight frequencies have been increased. Travel time for trips involving nonhubs has fallen from one to six percent on average.[9] While service by trunk airlines has been replaced with service by commuter airlines in many cases (which is seen as less desirable), most of these commuter lines have their schedules coordinated with a major carrier at the connecting hub. When there is provision for online ticketing, travelers can save approximately 25 percent over the interline fare. The few communities that have lost all service have not had enough traffic to support scheduled carrier service by any class of carrier. In these cases, service could be restored by government subsidies if the affected taxpayers deemed it desirable to do so.

Beyond the basic issues of where to fly and how much to charge, there is the issue of whether the skies have become less safe under deregulation. Generally, the argument is that competition gives airlines an incentive to cut corners on maintenance and to force pilots to fly more hours than is prudent. Under regulation, it was claimed that this was not a problem because the CAB ensured that the airlines were financially healthy so that they would not be as tempted to cut corners.

So far, the safety record of the airlines is as good as ever, but there is the charge by some that the country has simply been lucky. There are two responses to this charge. First, it is bad for an airline's business for its aircraft to be involved in an accident that is shown to be a result of its own negligence. Not only is the public likely to avoid the airline, but the airline would also have lost a plane worth millions of dollars and exposed itself to even greater claims of liability.[10] Second, and more important, one sure way of forcing the airlines to perform proper maintenance is to step up inspections by the Federal Aviation Administration (FAA). There may be a problem in doing this, however. The number of airlines and aircraft in service has risen dramatically since 1978, but the number of FAA inspectors has remained the same due to federal budget constraints.

A related problem is that the number and the level of experience of the nation's air traffic controllers has declined since deregulation as a result of the Professional Air Traffic Controllers' Organization (PATCO) strike in the summer of 1981. Thus, if there is a potential safety problem, it is likely to arise from inadequate attention to inspection and flight control, not from deregulation.

Industry Concerns

As one might have surmised from the earlier discussion of strikes, labor leaders were also concerned about the effects of deregulation. In fact, however, overall employment in the industry is up and compensation has kept pace with inflation. According to data presented by Morrison and Winston (1986), from 1975 to 1984, pilots' average real income fell a modest $500, dropping to $47,720 in 1977 dollars, while that of flight attendants increased $1800 to $14,428, and that of mechanics increased about $500 to $19,775.

Industry employment has increased since the early 1970s. Employment declined from a 1980 peak until 1983 when it rebounded and continued the upward trend it followed from 1971 to 1978 (see Morrison and Winston [1986]). Though the average worker has not suffered

9 An airport is classified as a "nonhub" if its total enplaned revenue passenger miles represents less than 0.05 percent of the total U.S. market.

10 It is assumed, of course, that the idea of preserving life also enters into the issue.

under deregulation, many union workers have been forced to take wage- and work-rule concessions, and some have had their careers interrupted as they have been either laid off or let go by airlines performing poorly in the new competitive environment. Two-tiered labor contracts have also been introduced. All this and the growth of the nonunion sector of the industry among the entering airlines have induced wide, and sometimes surprising, wage differentials between workers for different airlines, so that aggregate data on the welfare of workers is somewhat misleading.[11]

Finally, some firms may not have benefited from deregulation. There have been a number of bankruptcies in the airline industry since deregulation, most notably Braniff Airlines and Continental Airlines, which are both still flying after Chapter 11 reorganizations. Another airline (Frontier) is not flying, but is being acquired by Texas Air. In addition, there have been numerous mergers, particularly in the last year. Currently pending are two large mergers involving Continental-Eastern-People Express-Frontier (by Texas Air) and Delta-Western, that would create the first- and fourth-largest airlines in the U.S., respectively. While business failures impose some costs, such as uncertainty and inconvenience on the part of consumers, the loss of jobs on the part of workers, and the financial loss to creditors and stockholders, failures are a necessary force to ensure that firms operate efficiently in providing the services that consumers desire at a cost they are willing to pay.

IV. Future Evolution of the Industry

The current merger wave could be regarded as a natural process leading toward a competitive airline industry. Travelers prefer to have nonstop or one-stop flights with one carrier, rather than take a flight that would require them to endure two or more stops, or to change airlines at a busy airport. Providing such service requires a national route network with several regional hubs. In addition to the benefits for travelers, there also might be cost advantages to operating such a large hub network. Though the cost studies performed during the regulatory period indicated that there were no scale economies in the airline industry, the cost inefficiencies present in the regulatory era may have distorted these estimates. Bauer (1985) used an econometric procedure that allowed for these inefficiencies and found evidence of substantial returns to scale (contrary to the cost studies that did not allow for inefficiency). This issue aside, there are definitely cost advantages to the extent that large hub-and-spoke systems lead to higher load factors. Currently, only United Airlines and American Airlines operate such networks. However, once the current wave of mergers subsides, there will be anywhere from six to eight such super-airlines, perhaps another four to six medium-sized carriers, and perhaps 10 to 30 regional carriers.

Should the public be concerned about the potential anti-competitive effects of these airline mergers? If the industry were perfectly contestable as discussed earlier, then the answer would be no. Many researchers have tested whether or not the implications of the theory of contestable markets hold exactly; unfortunately, no one has found that they have. Bailey, Graham, and Kaplan (1985), for example, found that on concentrated routes (routes served by only one or two carriers) airlines can raise fares five to 10 percent over what they could charge on nonconcentrated routes.

There are two reasons why actual and potential competition have not lived up to their promise in the airline industry. First, capital—both physical and human capital—may not have fully adjusted to the new deregulated environment. The number of merger proposals recently is evidence that the airline industry is not in a long-run equilibrium with respect to the number and size distribution of carriers. Given that it has been eight years since the formal deregulation process started, it appears that the transition from a regulated to a competitive market equilibrium will take longer than expected.

A second reason for the apparent lack of competition on some routes is that entry into some concentrated markets is not as easy as was first expected. Many airports across the country have severe problems with traffic congestion (for example, airports in Denver and Washington, D.C.); obtaining gates and takeoff and landing slots at these airports is difficult. Since gates and landing rights are "grandfathered" to the airline holding them as long as they are used, the airlines that have these scarce resources can earn monopoly returns from them. This creates a severe barrier to entry for airlines wishing to begin service on these routes. The importance of this problem was highlighted in the recent merger of Continental Airlines with Eastern Airlines. To get approval for the merger, slots at LaGuardia airport had to be sold to Pan-Am so that it could set up a competing shuttle service. Even at relatively uncongested airports, such as Cleveland Hopkins, airlines are reluctant to release unused gate space. Much of the impetus for the current merger wave is that airlines find it is easier to buy other airlines to expand (in an

11 For example, unionized Western Airline workers earn less than Delta's nonunion workers. Also, United's unionized pilots earned 40 percent more than their ill-fated Frontier brethren.

effort to reach the most efficient size) than it is to grow internally (and be forced to try to obtain takeoff and landing slots on their own).[12]

Given that the contestable market theory does not seem to apply on all routes, should consumers worry about the increasing concentration of the industry? Currently, the national four-firm concentration ratio (CR), the sum of the market shares of the largest four firms in an industry, has remained unchanged at 47 from 1975 to 1986. Depending on how the current merger proposals are approved, it is likely that the resulting concentration ratio for the industry will be anywhere from 57 to 61. While this is high enough to cause concern, particularly in light of the fact that some individual city pairs now have even higher concentration ratios, there are reasons not to become alarmed just yet.

First, even though the industry has a fairly small number of firms, and concentration is relatively high, fare and route competition has been intense since deregulation. There have been no accusations that the industry as a whole is earning more than a normal profit. Furthermore, to the extent that only large airlines can provide the national route structure and the potential for nonstop and one-stop service that consumers prefer at the lowest cost, the level of concentration is only a reflection of the fact that there is only room for a limited number of efficiently sized airlines in the market.

If the ultimate effect of deregulation is a national market with six to eight huge airlines, there still would be a great deal of competition in the industry, even if many of the major cities are dominated by as few as two carriers. If one wants to fly from Cleveland to Los Angeles, for example, there may only be one or two airlines to choose from that provide nonstop service. However, one-stop service is a close substitute for nonstop service and, in that case, one would conceivably have six to eight choices depending upon which hub city he or she preferred to change planes. On shorter routes, such as Cleveland to Chicago, the smaller regional carriers would provide additional competition to the major carriers and thereby put a check on fares.[13] On still shorter flights, Cleveland to Columbus for example, surface transportation provides some additional competition even if the market for air travel between those points is concentrated. Given the shortcomings of the contestable market theory as applied to the airline industry, however, the disciplining effect of potential competition may not be enough to ensure competitive behavior. It may still be necessary for the Departments of Transportation and Justice to enforce current antitrust laws.

In summary, at this point, the market for air travel in the U.S. is not perfectly contestable and, on some concentrated routes, airlines are able to charge modest fare markups on the order of between 5 and 10 percent. This situation is likely to continue for the foreseeable future, until steps are taken to alleviate the congestion problems at certain airports. The next few years will probably witness an increase in the concentration in the industry to the point where six to eight large airlines dominate the national market with a host of smaller regional and commuter lines filling a variety of special niches. There will be sufficient competition to ensure that travelers are better off than they were under regulation, but it remains to be seen how closely the industry will conform to the perfectly contestable ideal that was envisaged by proponents of deregulation.

V. Conclusion

Deregulation of the airline industry has been a painful experience for some travelers, workers, and firms. Large fuel price increases, the air traffic controllers' strike, and recessions have made the process even more difficult. On the whole, however, deregulation has been favorable. Far more individuals have benefited than have been hurt. Consumers are receiving better service for lower average fares; employment and compensation in the industry are up; and the airlines are generally earning higher profits than they would have under regulation. Yet, even eight years later, the industry is still adjusting to its new environment, and the final results of deregulation have yet to be determined.

There are several steps that can be taken to ensure that the gains to date are not lost and that the costs of adjustment to deregulation are minimized. First, airport expansion is needed to help reduce one of the few barriers to entry that remain in the industry. Deregulation, by greatly increasing air travel through lower fares, made the congestion worse. The solution, however, is not to reduce air travel, but to expand the system.

The federal government has a $3.5 billion fund that can be spent only on promoting air travel. This fund is financed by an 8 percent tax on air fares, but has become embroiled in the current federal budget problems. The money

12 A further cause of the increased merger activity now is that the Department of Transportation (DOT) has authority over airline mergers for the next two years, at which time the Department of Justice (DOJ) will have that responsibility. The DOT has been much more lenient than the DOJ.

13 If they cannot obtain space at the major airports on the route in question, they have the aircraft that can effectively utilize the smaller regional airports which, in some cases, may be more convenient for passengers.

could be spent to expand airport facilities, to modernize the air traffic control system, and to hire more FAA inspectors. These expenditures would enhance the competitiveness of the system by lessening the incentives for airlines to merge, as well as by improving their safety and reliability.

Second, the U.S. Departments of Transportation and Justice should continue to enforce existing antitrust laws. While the competitive discipline that free-entry into the industry offers should not be ignored, it is important that these agencies not place too much faith in free-entry to the exclusion of other factors, particularly in the short run.

Finally, allowing foreign air carriers into the U.S. market (with reciprocal agreements for entry into their markets) should be considered as a way of further increasing industry competition. These policies would help the U.S. to maintain its position as having the world's foremost airline network.

REFERENCES

Adams, Douglas. *The Hitchhiker's Guide to the Galaxy.* New York, NY: Pocket Books, 1979.

Bailey, Elizabeth E., David R. Graham, and David P. Kaplan. *Deregulating the Airlines.* Cambridge, MA: The MIT Press, 1985.

Bauer, Paul W. "An Analysis of Multiproduct Technology and Efficiency Using the Joint Cost Function and Panel Data: An Application to the U.S. Airline Industry," Ph.D. Dissertation, University of North Carolina at Chapel Hill, 1985.

Davies, R.E.G. *Airlines of the United States since 1914.* London: Putnam & Company Limited, 1972.

Douglas, George W., and James Miller. *Economic Regulation of Domestic Air Transport: Theory and Policy.* Washington, D.C.: Brookings Institution, 1974.

Graham, David R., Daniel P. Kaplan, and David S. Sibley. "Efficiency and Competition in the Airline Industry," *Bell Journal of Economics,* vol. 14, no. 1 (spring 1983), pp. 118-138.

Keeler, Theodore E. "Airline Regulation and Market Performance," *Bell Journal of Economics and Management Science,* vol. 3, no. 2 (autumn 1972), pp. 399-424.

Meyer, John R., and Clinton V. Oster, Jr. *Deregulation and the New Airline Entrepreneurs.* Cambridge, MA: The MIT Press, 1984.

_____. *Airline Deregulation: The Early Experience.* Boston, MA: Auburn House, 1981.

Morrison, Steven, and Clifford Winston. *The Economic Effects of Airline Deregulation.* Washington, D.C.: Brookings Institution, 1986.

U.S. General Accounting Office. "Lower Airline Costs per Passenger are Possible in the United States and Could Result in Lower Fares," Report to Congress, February 1977.

White, Lawrence J. "Economies of Scale and the Question of 'Natural Monopoly' in the Airline Industry," *Journal of Air Law and Commerce,* vol. 44, no. 3 (1979), pp. 545-573.