A Typology of Asset Pricing Models

R. D. Coleman

31 October 1997

TITLE PAGE

Title:	A Typology of Asset Pricing Models
First Submitted:	13 June 1995
Current Draft:	31 October 1997

Author for Correspondence:	R. D. Coleman 3715 Bryn Mawr Drive Dallas, Texas 75225 (214) 363-7424 numeraire @ earthlink.net
Requests for Reprints:	R. D. Coleman 3715 Bryn Mawr Drive Dallas, Texas 75225 numeraire @ earthlink.net

ABSTRACT

A typology of asset pricing models is developed as a two-part classification: first, the motivating theory behind the model, and second, the potential for logical circularity between the explained and explanatory variables. The typology is designed to quickly identify and segregate potentially spurious models for closer scrutiny and thus serves as a cautionary check for empirical studies of asset pricing models formulated to explain expected total return.

CIRCULAR FORMS AND TYPES

Where logical circularity is present in an asset pricing model, it is manifested in the form of either mathematical identities or autoregressions. Mathematical identities are tautologous and thus are not valid independent explanations. Autoregressions that are not theoretically-motivated but rather data-instigated are fallacious on both theoretical and methodological grounds. The ad hoc alleged explanatory variables of such autoregressive models, i.e., price, dividends, and number of shares outstanding, are market-generated and are entailed in the definition of total return,.

The forms of the efficient markets hypothesis (EMH) happen to correspond to the forms of asset pricing circularity which in turn, by design, correspond to the types of circularity. We first consider the EMH forms, then the circular forms, and finally the circular types.

Forms of Efficient Markets Hypothesis

One of the most fundamental concepts in neoclassical economics is that rates of return tend to equality, *ceteris paribus*. This idea can be traced to Adam Smith (1937:99):

> The whole of the advantages and disadvantages of the different employments of labor and stock must, in the same neighborhood, be either perfectly equal or continually tending to equality. If in the same neighborhood, there was any employment evidently either more or less advantageous than the rest, so many people would crowd into it in the one case, and so many would desert it in the other, that its advantages would soon return to the level of other employments. This at least would be the case in a society where things were left to follow their natural course, where there was perfect liberty and where every man was perfectly free ...

Of course, the operative phrases are "perfect liberty" and "perfectly free" in the above concept of the efficiency of competitive markets. The mobility of factors and financial capital is essential to the productivity and growth of a country's wealth.

According to the efficient market hypothesis (EMH), stock prices already reflect all data about the future prospects of the firms whose common stocks are traded. The narrow, semi-broad, and broad versions of the EMH interpret the scope of "all data" differently. One key to profitable securities investing is the distinction between data and information. The processing of data to produce useful information for decision-making takes times, costs money, and requires knowledgeable interpretation. Thus data is considered a free public good, but information is a private commodity.

According to the narrow version of the EMH, stock prices already reflect all information that can be derived by examining market-generated data such as share price, share trading volume, and short interest for each firm. For our purposes, we would also include dividends distributed per share and number of shares outstanding. These are the explanatory variables that are entailed in return, the explained variable in the conventional capital asset pricing model (CAPM). Thus, so-called technical analysis is unprofitable according to this interpretation. Technical analysis involves a search for persistent and predictable patterns in stock prices that result from a lag in the response of price signals to new information about the future prospects of firms and the analysis of this information. This predictability must have not only statistical significance but also practical economic significance in terms of profitability. This condition is exactly opposite to the requirement of a perfectly efficient market that reflects all information instantly and completely. Practitioners of technical analysis or "charting" imply that they have paranormal powers of "graphicity" which enable them to divine future movements in stock prices from visual observation of their charted time-sequences of data.

In contrast to the visual chartists are the quantitative analysts or mathematical traders who develop proprietary algorithms which are the computational formulas for complicated trading questions. The danger for quantitative analysis is that markets can change unexpectedly and computer models thereby lose their power to analyze risk and return. Thus there is a continuing emphasis on new financial research to take advantage of market inefficiencies that manifest as emerging trends that offer transient rewards to whoever identifies them and acts on them first. A hedging strategy is market neutral in the sense of being neither long nor short. A price movement in a stock or bond has value relative to a portfolio of hundreds of securities, each with its own particular mathematical characterization. The computer tries to create a risk-neutral *n*-dimensional matrix of a portfolio of securities.

According to the semi-broad version of the EMH, stock prices already reflect all publicly available data pertaining to the future prospects of the firms whose common stocks are traded in the market. This information includes so-called fundamental data on each firm regarding product lines, management, external accounting reports, earnings forecasts, and monopoly positions held through patents, copyrights, or government regulations.

According to the broad version of the EMH, stock prices already reflect all relevant data pertaining to the firm. In particular, this includes information available only to company insiders and not to the general public at large. This is the most extreme version of the EMH and thus serves as a benchmark for comparison.

The two forms of circularity, identity and autoregression, coincide with the forms of the EMH. Circular identities are associated with market-generated data and thus fit the narrow form of the EMH. Circular autoregressions are associated with any explanatory factor based on available data and thus fit the semi-broad and broad forms of the EMH. The broad form is assumed in testing models because it alone provides a necessary fixed reference for comparison. This assumption is normative, not positive. It is good methodology, although not good economics.

Forms of Circularity

The operational form of the estimating equation can be a pure identity, a pure autoregression, or a mixture of identity and autoregression.

The inverse relationship between return and size defined as market value of equity that is found in empirical studies is not explained by these forms if the true relationship is linear in either size or natural logarithm of size. There is no theoretical reason for the multivariate asset pricing model to be linear either in all variables or in their transforms. Furthermore, it has been found (Brown *et al.*, 1983) that the "size effect" reverses itself when averaged over all months. The direction of causality, if not simultaneous by definition, goes not only from price, market value of equity, or dividends to total return but also from total return to price, market value of equity, or dividends. Thus the problem of "reverse causality" is present in any model that includes price, market value of equity, or dividends as explanatory variables. Any one of the variables total return, share price, market value of equity, or dividends can be isolated on the LHS with equal justification as the explained variable with one or more of the remaining of these variables on the RHS as explanatory variables.

Many researchers eschew the use of the term "cause" but do not hesitate to use the term "effect" , for example, in referring to the alleged size or ME effect, price effect, P/BE effect, P/E effect, BE/ME effect, dividend yield or D/P effect, and January effect. A common usage is to express test results in terms of one variable "statistically explaining" another variable or of the "statistical coordination", "statistical association", or "statistical comovement" of two variables.

Share price is entailed in market value of equity, the price of all of a firm's shares outstanding, and can reduce the apparent return-market equity relationship to a tautology. Even if there may be a price effect or a number-of-shares effect in the determination of return, either of these effects might be difficult to isolate given a meaningful concept of return.

A firm's balance sheet has total assets on the credit side and total claims on assets on the debit side. Claims include both debt and equity. Referring to the market value of equity as "size" can add to the obfuscation because market value of equity as a firm-specific scaling variable does not measure a stable scale of operations but rather the unstable magnitude of the product of the number of shares of common stock outstanding and share price from trade to trade in the market. A more stable measure of financial "size" determined by the capital markets is the magnitude of the sum of the products of unit price and number of units of all classes of financial instruments outstanding. The prices of these financial claims on the firm's assets may be determined in either private or public markets for debt and equity claims. Thus, the choice of a measure for “size” is arbitrary, and in the case of market value of equity, both relatively most unstable and logically circular.

Another form of circularity occurs when using deflating variables (Madansky, 1964). Deflators can lead to spurious correlation between ratios that have a common denominator. We here refer to certain usages of scale deflators and not price-level deflators because price-level deflators are not common denominators. A simple example will illustrate this phenomenon:

$$X = \mathrm{a} + \mathrm{b}Y + \mathrm{e}(Z) \qquad (1)$$

is the estimating equation where X, Y, Z and e are contemporaneous random variables. A scaling variable is used to transform the equation:

$$X/Z = \mathrm{a}/Z + \mathrm{b}Y/Z + e, \qquad (2)$$

or rearranging terms,

$$(1/Z)X = (1/Z)\mathrm{a} + (1/Z)\mathrm{b}Y + e, \qquad (3)$$

which has the form

$$AD = AB + AC + E, \tag{4}$$

or rearranging terms to isolate A on the left-hand side (LHS),

$$A = A(B/D) + A(C/D) + E(1/D), \tag{5}$$

which shows the presence of A on both the LHS and the right-hand side (RHS).

Spurious effects due to logical circularity can occur on both the theoretical and empirical levels of an asset pricing model. Spuriousness at the theoretical level can occur when using a market index proxy that is sample-specific or calculated from the particular sample data. For example, a sample of ten common stocks is equivalent to a sample grouped into decile portfolios. The expected return of any decile portfolio is subsumed in the expected return of a market index proxy calculated from the expected returns of all ten decile portfolios. If the individual stock is removed from a sample-specific index, this will eliminate the double-counting. But with many small-size samples, this may involve the use of a less broad and thus less representative proxy for the complete capital asset market.

Spuriousness at the empirical level can occur if the market value of equity variable or its entailed variables, share price and number of shares of common stock outstanding, appear on both the LHS and the RHS of the estimating equation contemporaneously. In regression equations, the explained variable and the explanatory variable are given asymmetrical treatment (Davis, 1985:9-10). But the mathematics of estimation cannot distinguish between different designated roles for the specified variables. With entailed scaling deflator variables, the spurious effect may not be as problematic as with logical identities, but that is an empirical question that is sample-specific.

Including the contemporaneous share price on both sides of the CAPM equation with total return as the explained variable can effectively reduce the equation to a model of dividend yield or individual firm dividend policy and practice. Likewise, specifying contemporaneous dividends per share on both sides of the CAPM equation with total return as the explained variable can effectively reduce the model of total return to a model of capital gain or price appreciation return.

Expected *ex ante* and realized *ex post* returns are mathematically indistinguishable. What is different are the names for the anterior and posterior points in time that demarcate the beginning and ending of the holding period of return. Since *ex ante* expected returns are not observable and can be expressed in terms of historical returns, we will name these two points "prior" and "current" times as opposed to "current" and "future" times, respectively, for beginning of period and end of period.

Conceptually, capital return per share is the first-difference in share price scaled by the prior share price. Furthermore, it is useful for our purposes to think of share price as market value scaled by the number of shares of common stock outstanding. Thus, capital return is the first-difference in market value per share scaled by prior market value per share, i.e., a unit firm wealth or value relative minus one. Total return is the sum of capital gains return and dividend return. Dividend return or dividend yield is dividends divided by prior share price.

Thus in symbolic notation we find the following implicit syllogism behind this and similar studies where P is share price, D is dividends per share, N is number of shares outstanding of a security issue, and MKEQ is total firm market value of equity, all without time and security indexing subscripts:

Premise 1: $R = f(P*N, D*N)$ (6.a)
Premise 2: $MKEQ = P*N$ (6.b)
Conclusion: $R = f(MKEQ, D*N)$ (6.c)

The second premise can be either an embedded partial identity or an implicit autoregressive function. Such circularities may be camouflaged by a model of return (R) with the explanatory factors of "size" (MKEQ), book-to-market equity (BKEQ/MKEQ), price-earnings ratio (P/E), dividend yield (D/P), and share turnover (trading volume/N). The combination of these five factors would include (1) market value of equity in both positive and inverse forms, (2) share price in both positive and inverse forms, (3) both the capital gains and dividend income components of total return, and (4) number of shares outstanding in inverse form. Such results, even when significant, are misleading and of questionable scientific validity.

Tautologies have pedagogical value. The only successful competition for an indirect tautology is a less indirect tautology. As listed in the Table, several studies have investigated the effects of price/earnings, earnings/price, dividends/price, dividends, market value/book value, book-to-market equity, market value/sales. The market value of common stock equity capitalization goes by several shorter names including market value of equity, market equity, and market capitalization. Likewise, the accounting value of common stock equity goes by the names of book value of equity and book equity. All of these "explanatory" variables entail the share price, dividends per share, and or the number of shares outstanding which appear on both sides of the identity function equation. Return is independent of number of shares purchased except in the important special case of the market for corporate control where "greenmailers" and raiders seek to receive preferential treatment for their accumulated block of shares. So-called greenmail is authorized to be paid by executives of the target firm in order to save their jobs and entrench their positions as agents of the stockholder owners. In this special case, return is not independent of the number of shares outstanding which changes as a result of stock splits, stock dividends, and the exercise of stock options. Thus number of shares outstanding can serve as a scaling deflator for market value of equity and for dividends.

In addition, logarithms of variables are not independent of the variables used as arguments of the logarithmic function. For example ln(price) is not independent of price. Likewise, ln(MKEQ) and ln(B/M) are not independent of MKEQ or M and thus of price.

Although tautologous and not independently valid in itself, market value of equity can be specified in a separate model as an empirical benchmark for the explanatory power of models. In contrast, the January seasonal is not a tautology in relation to returns but rather may be partly an income tax-driven anomaly or a liquidity preference effect. The January seasonal effect and the day-of-the-week effect are the only known non-circular robust economic factor contributing to the explanation of returns, and they are temporal rather cross-sectional.

Typology of Asset Pricing Models

The types of circularity in the typology are designed to closely follow the forms of circularity presented above. We offer a two-part typology of asset pricing models paralleling the genus and species of the Linnéan taxonomy and his compound nomenclature for particular varieties. The first part is based on the motivating theory behind the model, and the second part is based on the potential for circularity between the explained and explanatory variables. Each part of the classificatory scheme is the same for the theoretical model, the corresponding empirical equation, and its operationalization.

In the first part of the typology, there are three categories based on the criterion of motivating theory behind the asset pricing model: (1) utility, (2) arbitrage, and (3) *ad hoc.* Asset pricing models based on utility theory or parameter preference lead to explanatory variables that are parameters of the distribution of return such as mean, standard deviation, variance, semi-variance, skewness, and kurtosis, i.e. distribution moments or descriptive statistics. In contrast, asset pricing models based on arbitrage theory or state preference lead to explanatory variables that are macroeconomic economy-wide factors, microeconomic firm-specific factors, and intermediate sector-specific factors. The arbitrage-theory models can be usefully divided between those with predetermined observable factors and those with so-called "blind" factors. Asset pricing models based on no explicit pricing theory are purely empirical and can lead to a potpourri of *ad hoc* factors that have no *a priori* reason to have a relationship with return. As Koopmans (1947:164) states:

> There is no systematic discussion of the reasons for selecting these particular variables as most worthy of study. ... The choices made may have been the best possible ones. But 'good' choices means relevant choices. What is relevant can only be determined with the help of some notions as to the [phenomenon of interest]. ... The choices as to what variables to study cannot be settled by a brief reference to 'theoretical studies' ... These issues call for a systematic argument to show that the best use has been made of available data in relation to the most important aspects of the phenomena studied.

In the second part of the typology, there are four dichotomized categories based on different criteria of circularity: (1) scale-deflated or not scale-deflated, (2) logical identity or not logical identity, whether full or partial identities, (3) autoregressive or not autoregressive, whether staggered or non-staggered autoregressions, and (4) narrow or not narrow (semi-broad or broad) form of the efficient market hypothesis (EMH). We do not imply that autoregressive models are not valid, a frequent fallacy. Rather we seek to emphasize the distinction between autoregressive models and identity models. A logical identity may be either a strict multiplicative identity ($A = AB$; thus $B=1$, and A is superfluous), a strict additive identity ($A = A+B$; thus $B=0$, and A is superfluous), or a combination of multiplicative and additive identities ($A = AB + C$; thus $A = C/(1-B)$). With any type of logical identity, the explained variable is being used to explain itself, i.e., it serves simultaneously in the roles of both explanatory and explained variable. The only model with more alleged explanatory power than an identity function in this sense may be a lower-degree and thus less-indirect identity function. The autoregressive nature and the identity nature of some models may be unintentional and obscurely embedded. The narrow-EMH form includes models with explanatory variables that are entailed in the explanatory variable, total return. The broad form includes models that include variables other than those of the narrow form. Of course, these are pure types and hybrids or combinations of these types are possible.

Any model designed to explain return that specifies share price, dividends, number of shares outstanding, or a variable such as market value of equity that entails one or more of these variables as explanatory factors either will have a deflating variable that may cause spurious relationships and create statistical artifacts, will be an autoregression, or will be a logical identity. This is a matter of mathematics and does not concern the author's story about the intended roles of different variables in the model. Explicit statements of such embedded relationships can help to avoid misleading interpretations. Market value of equity, book equity/market value of equity, market value of equity/replacement cost or Tobin's *q*, price/earnings, and dividend/price would be classified as either *ad hoc* autoregressive if non-contemporaneous (lagged with returns) or *ad hoc* identity function if contemporaneous (not lagged with returns).

That a practical need exists for such a typology to guard against circular models can be seen from the record of published academic journal papers alone. The Table presents a convenience sampling of such models to demonstrate that irrational behavior and “extraordinary delusions and the madness of crowds” is not limited to common stock investors.

Table. Models of R regressed on P, D and N. Full citations appear in the references.

		Basic Forms of Explanatory Variables									
		Simple						Compound			
	Model	Direct			Inverse			Dir.	Inv.		
	Variable(s)	P	D	N	1/P	1/D	1/N	ME	1/ME	Year	Author
	Models with Simple Variables Only										
1	P	X								1973	Blume
2	P/E	X								1977	Basu
3	BE/P				X					1985	Rosenberg
4	E/P,D/P		X		X					1978	Ball
5	D/P		X		X					1985	Keim
6	D/TA		X							1968	Nerlove
7	V/N						X			1968	Nerlove
	Models with Compound Variables										
8	ME	X		X				X		1981	Banz
9	ME *	X		X				X		1985	Chan
10	ME,E/P	X		X	X			X		1981	Reinganum
11	MC/RC or q				X		X		X	1991	Servaes
12	BD/ME,D		X		X		X		X	1991	Chan
13	ME,MD/ME	X		X	X		X	X	X	1982	Christi
14	ME,BD/ME	X		X	X		X	X	X	1988	Bandari
15	ME,BE/ME	X		X	X		X	X	X	1993	Fama
16	ME,BE/ME	X		X	X		X	X	X	1995	Berk
17	ME,BE/ME,D	X	X	X	X		X	X	X	1994	He
18	BE/TA									1994	Opler

* portfolio formation variable

Legend to Table

BD	book debt
BE	book equity
D	dividends per share
E	earnings per share
MC	market value of debt and equity claims
MD	market value of debt
ME	market value of common stock equity
MP	market value of preferred stock equity
N	number of shares outstanding
P	share price
R	total returns
RC	replacement cost of reproducible assets
TA	total assets
V	trading volume number of shares

REFERENCES

Ball, R., 1978, "Anomalies in Relationships between Securities' Yields and Yield-Surrogates," *Journal of Financial Economics*, 6, July/September, 103-126.

Banz, R. W., 1981, "The Relationship between Return and Market Value of Common Stocks," *Journal of Financial Economics*, 9, March, 3-18.

Basu, S., 1977, "Investment Performance of Common Stocks in Relation to Their Price-Earnings Ratios: A Test of the Efficient Market Hypothesis," *Journal of Finance*, 32, June, 663-682.

Bhandari, L. C., 1988, "Debt/Equity Ratio and Expected Common Stock Returns: Empirical Evidence," *Journal of Finance*, 43, 507-528.

Blume, M. E. and Husic, F., 1973, "Price, Beta and Exchange Listing," *Journal of Finance*, 28, 1973, 283-299.

Brown, P., Kleidon, A. W., and Marsh, T. A., 1983, "New Evidence on the Nature of Size-Related Anomalies in Stock Prices," *Journal of Financial Economics*, 12, June, 33-56.

Chan, K. C., Chen, N., and Hsieh, D., 1985, "An Exploratory Investigation of the Firm Size Effect," *Journal of Financial Economics*, 14, 1467-1484.

Chan, K. C. and Chen, N., 1991, "Structural and Return Characteristics of Small and Large Firms," *Journal of Finance*, 46, 4, September, 1467-1484.

Christie, A. A., 1982, "The Stochastic Behavior of Common Stock Variances," *Journal of Financial Economics*, 10, 407-432.

Davis, J. A., 1985, *The Logic of Causal Order*, Sage Publications, Newbury Park, CA.

Fama, E. F. and French, K. R., 1993, "Common Risk Factors in the Returns on Bonds and Stocks," *Journal of Financial Economics*, 33, 3-56.

He, J. and Ng, L. K., 1994, "Economic Forces, Fundamental Variables, and Equity Returns," *Journal of Business*, 67(4), 599-609.

Keim, D. B., 1985, "Dividend Yields and Stock Returns: Implications of Abnormal January Returns," *Journal of Financial Economics*, 14, 473-489.

Koopmans, T. C., 1947, "Measurement without Theory," *Review of Economics and Statistics*, 29, August, 161-72.

Madansky, A., 1964, "Spurious Correlation Due to Deflating Variables," *Econometrica*, 32, October, 652-655.

Nerlove, M., 1968, "Factors Affecting Differences among Rates of Return on Investments in Individual Common Stocks," *Review of Economics and Statistics*, 50(3), August, 312-331.
Opler, T. C. and Titman, S., 1994, "Financial Distress and Corporate Performance," *Journal of Finance*, 49(3), July, 1015-1040.

Reinganum, M. R., 1981, "Misspecification of Capital Asset Pricing: Empirical Anomalies Based on Earnings Yields and Market Values," *Journal of Financial Economics*, 9, 19-46.

Rosenberg, B., Reid, K. and Lanstein, R., 1985, "Persuasive Evidence of Market Inefficiency," *Journal of Portfolio Management*, 11, 9-17.

Servaes, H., 1991, "Tobin's Q and the gains from takeovers," *Journal of Finance,* 46 (March): 409-419.

Smith, A., 1937, c1776, *An Inquiry into the Nature and Causes of the Wealth of Nations*, E. Cannan (ed.), Modern Library, New York.