On Model Testing in Financial Economics

October 1, 2009

Abstract

This note discusses the two different contradicting philosophies for testing models in financial economics (asset pricing, corporate finance, and market-microstructure) using linear regression. We synthesize these two contradicting approaches, document the errors that may occur in the existing estimation methodologies, and suggest a modified procedure that avoids these errors.

1 Introduction

In financial economics, linear regression is used in two contradicting ways to test the validity of a financial model.[1] Both methods can be illustrated with the following simple structure. Let $Z_i = f(x_i)$ be a model of a financial variable Z_i as a function of x_i. In addition, let y_i be another financial variable that potentially impacts Z_i, but that is excluded from the model.

To test the model we run the linear regression

$$Z_i = \alpha + \beta f(x_i) + \gamma y_i + \epsilon_i$$

where α, β, γ are constants, and ϵ_i are i.i.d. measurement errors.

The first approach accepts the model if $\alpha = 0$, $\beta = 1$, $\gamma = 0$ and rejects it otherwise. This approach is most often used in empirical asset pricing. For example, consider a simple test of the CAPM or the APT. Here, Z_i is the excess return on a stock at time i, $f(x_i)$ is the excess return on a factor portfolio x_i, and y_i is the standard deviation of the stock's return. The model is rejected unless the standard deviation has no influence on the stock's excess return and the alpha is zero, i.e. unless $\alpha = 0$, $\beta = 1$, and $\gamma = 0$ (for more realistic examples see Black, Jensen and Scholes [2], Roll and Ross [9]).

In contrast, the second approach accepts the model if $\beta = 1$ for any (α, γ), and rejects it otherwise.[2] This second approach is most often used in corporate

[1] Although this paper emphasizes the use of linear regression in financial economics, this same issue applies more generally to model testing in economics.

[2] Sometimes the model is accepted in this alternative approach if $\beta > 0$, but for comparison purposes, we only discuss this variant in footnotes.

finance and market-microstructure. For example, in corporate finance, consider a simple test for whether debt impacts firm value. Here, Z_i is the value of firm i, $f(x_i)$ is the value of the firm as a function of it's current debt level x_i, and y_i is the dividend payout. In this approach, y_i is viewed as a "control variable." The model is only rejected if $\beta \neq 1$ (for more realistic examples see Baker and Wurgler [1], Fama and French [6]). Similarly, in market microstructure, consider a simple test for the determination of the bid-ask spread. Here, Z_i is the bid-ask spread at time i, $f(x_i)$ is the bid-ask spread as a function of the market price x_i, and y_i is the stock's volume. Here, y_i is viewed as a "control variable." The model is only rejected if $\beta \neq 1$ (for more realistic examples see Easley, Kiefer, O'Hara and Paperman [5]).

As contrasted, these two approaches are in contradiction. The first rejects the model if $\alpha \neq 0$ or $\gamma \neq 0$. The second accepts the model under these conditions, if $\beta = 1$. The choice between these two approaches is often dictated by the complexity of the model's structure relative to the market's structure. More complex models tend to use the first approach, while simpler models use the second. Indeed, this is consistent with the current state of modeling in asset pricing, corporate finance, and market-microstructure relative to the complexity of the issues involved. Asset pricing models are often applied to well-functioning and competitive markets with few market frictions, hence, the existing models can more easily capture the complexity of the market structure. In contrast, the same is not true for many corporate finance and market-microstructure models. These models explicitly deal with frictions and game theoretic behaviors, which are more difficult to model, except in simplified forms.

As illustrated by our simple construct, these two linear regression based testing approaches are in contradiction. Yet, both are frequently used. Is one approach correct and the other false? Or, can both be correct, but under different circumstances? The purpose of this paper is to answer these questions and to reconcile this apparent contradiction.

We do this by providing a supermodel that subsumes these two testing approaches as special cases. As such, we synthesize these two approaches, document the errors that may occur in the existing estimation methodologies, and suggest a modified procedure that avoids these errors.

An outline for this paper is as follows. Section 2 presents the structure for the analysis. Section 3 presents the true regression model, while section 4 presents the estimated regression and the reconciliation of the two approaches. Finally, section 5 provides an illustrative example from corporate finance.

2 The Set-up

Let $Z_i \in \mathbb{R}$ be some quantity of interest in a financial model, e.g. the return on a common stock at time i (asset pricing), or the value of firm i (corporate finance), or the bid-ask spread (market-microstructure). The index i could be *time* for a single firm (time series), or it could be a *firm* for a single time (cross-sectional), or it could be both a time and firm index (joint time series, cross-sectional).

We define a *financial supermodel* to be a mapping from a space of primary variables (changing with the index i) $x_i, y_i \in \mathbb{R}$ and parameters (constants) $\theta \in \mathbb{R}^n$ into the real line, i.e. $Z_i = f(x_i, y_i, \theta) \in \mathbb{R}$.

For simplicity we only study two variables but we could have an arbitrary, but finite number. For example in an asset pricing model, letting Z_i be the stock return at time i, the variables might be the returns at time i on factor portfolios and the parameters the factor betas. While in a corporate finance model, letting Z_i be value of firm i, the variables might be the debt equity ratio and dividend payout, the parameters the relative impact that these variables have on firm value. Lastly, in a market-microstructure model, letting Z_i be the bid-ask spread at time i, the variables might be the market price and volume, the parameters being their relative impact on the bid-ask spread.

The crucial distinction between variables and parameters is that the parameters θ do not depend on index i. Because θ does not depend on the index, without loss of generality, we can omit it from the formula, so that a supermodel is denoted $f(x_i, y_i) \in \mathbb{R}$. We assume that f is twice continuously differentiable.

We view a *financial model* as a restricted formulation of the supermodel where some of the variables take on a constant value, i.e. a model is a restriction $f(x_i, \alpha) \in \mathbb{R}$ where y_i is fixed at a constant $\alpha \in \mathbb{R}$.

The idea is that due to the complexity in modeling the phenomena of interest, the modeler simplifies the construct, and only studies the influence of the variable x_i, holding all else, y_i, constant. For example, in asset pricing and market-microstructure models transaction costs are often assumed to be zero, while in corporate finance we often assume the analogous issuance costs are zero. Surely these variables potentially influence the economic setting, but they are often excluded to make the model tractable.

3 True Regression

This section presents the true regression model based on the supermodel. When we go to the data, the true model is

$$Z_i = f(x_i, y_i) + \epsilon_i \tag{1}$$

where ϵ_i are i.i.d. and they represent measurement error in Z_i.

Using a Taylor series expansion on $f(x_i, y_i)$ for y_i, we obtain

$$Z_i = f(x_i, \alpha) + \frac{\partial f}{\partial y}(x_i, \alpha)(y_i - \alpha) + \frac{\partial^2 f}{2\partial y^2}(x_i, \xi(y_i))(y_i - \alpha)^2 + \epsilon_i$$

where $\alpha \leq \xi(y_i) \leq y_i$. Next, we apply a Taylor series expansion to $\frac{\partial f}{\partial y}(x_i, \alpha)(y_i - \alpha)$ in x_i:

$$\begin{aligned} Z_i &= f(x_i, \alpha) + \frac{\partial f}{\partial y}(\beta, \alpha)(y_i - \alpha) + \frac{\partial^2 f}{\partial x \partial y}(\eta(x_i), \alpha)(y_i - \alpha)(x_i - \beta) \\ &\quad + \frac{\partial^2 f}{2\partial y^2}(x_i, \xi(y_i))(y_i - \alpha)^2 + \epsilon_i \end{aligned} \tag{2}$$

where $\beta \leq \eta(x_i) \leq x_i$. Of course, the Taylor series could be expanded to higher orders as necessary.[3] This Taylor series expansion represents the true model. It was formulated to express the true model in terms $f(x_i, \alpha)$, y_i, and higher order terms.

4 Estimated Regression

Due to complexity of modeling, it is often the case that the modeler simplifies the structure and assumes that some potentially important financial variable y_i is a constant. Hence, the financial model is represented by the restriction $f(x_i, \alpha)$. To test the restricted model, the linear regression estimated is

$$Z_i = c_0 + c_1 f(x_i, \alpha) + c_2 y_i + \varepsilon_i. \tag{3}$$

The interpretation is that $f(x_i, \alpha)$ is the model to be tested and y_i is the control variable. We see here that the empirical expression (3) is valid as a special case of the supermodel if and only if

$$c_0 = -\frac{\partial f}{\partial y}(\beta, \alpha)\alpha,$$

$$c_1 = 1,$$

$$c_2 = \frac{\partial f}{\partial y}(\beta, \alpha), \quad \textit{and}$$

$$\varepsilon_i = +\frac{\partial^2 f}{\partial x \partial y}(\eta(x_i), \alpha)(y_i - \alpha)(x_i - \beta) + \frac{\partial^2 f}{2\partial y^2}(x_i, \xi(y_i))(y_i - \alpha)^2 + \epsilon_i.$$

For hypothesis testing, an important observation is that the new error terms ε_i are no longer i.i.d. They depend on the interaction terms involving $x_i y_i$ and y_i^2. If not addressed in the estimation, this can invalidate any inferences drawn. This can be viewed as a standard omitted variables problem. We note a trivial corollary when this omitted variables problem creates no difficulties.

Corollary 1 *The linear regression model is properly specified if and only if* $\frac{\partial^2 f}{\partial x \partial y} = 0$ *and* $\frac{\partial^2 f}{\partial y^2} = 0$ *for all* (x, y) *in a neighborhood of* (α, β).

[3] For illustration purposes, a higher order expansion might be:

$$\begin{aligned} Z_i &= f(x_i, \alpha) + \frac{\partial f}{\partial y}(\beta, \alpha)(y_i - \alpha) + \frac{\partial^2 f}{\partial x \partial y}(\beta, \alpha)(y_i - \alpha)(x_i - \beta) \\ &\quad + \frac{\partial^3 f}{2\partial x^2 \partial y}(x_i, \alpha)(y_i - \alpha)(x_i - \beta)^2 \\ &\quad + \frac{\partial^2 f}{2\partial y^2}(\beta, \alpha)(y_i - \alpha)^2 + \frac{\partial^3 f}{2\partial x \partial y^2}(\zeta(x_i), \alpha)(y_i - \alpha)^2(x_i - \beta) \\ &\quad + \frac{\partial^3 f}{6\partial y^3}(x_i, \xi(y_i))(y_i - \alpha)^3 + \epsilon_i \end{aligned}$$

where $\beta \leq \zeta(x_i) \leq x_i$.

Returning to the two testing approaches in financial economics, the model is accepted if and only if:

$$\begin{array}{cc} \textit{approach } 1 & \textit{approach } 2 \\ c_0 = 0,\ c_1 = 1,\ c_2 = 0 & c_1 = 1 \\ \varepsilon_i \ \ i.i.d. & \varepsilon_i \ \ i.i.d. \end{array}$$

We can now analyze these two approaches in more detail.[4]

4.1 Approach 1

Approach 1 takes the "purest" perspective, that the supermodel is the model itself, i.e. that $\frac{\partial f}{\partial y}(\beta, \alpha) = 0$. Under this perspective $c_0 = 0$, $c_2 = 0$ and the error term ϵ_i is i.i.d. There is no omitted variables problem. Although correct, this is the most narrow interpretation of the model's validity.

4.2 Approach 2

An alternative interpretation is that the model is a restricted version of a more general model. Here, the model is accepted if and only if

$$c_1 = 1 \ \ \textit{and} \ \ \varepsilon_i \ \ i.i.d.$$

The intuition is that under this restriction, the model tested captures an important economic factor where for the argument under consideration, the omitted factors are orthogonal. If the omitted variables are orthogonal, then for the purposes of the model under consideration, they can be viewed as only introducing noise. This is approach 2.

In the empirical testing, the orthogonality condition for the model's error terms needs to be explicitly considered to avoid the omitted variables problem. If not explicitly considered, the estimates could be biased with invalid inferences drawn. To avoid incorrect inferences, among others, the following testing procedure could be employed.

Pragmatic Testing Procedure

- Step 1. Run the regression

  $$Z_i = c_0 + c_1 f(x_i, \alpha) + c_2 y_i + \varepsilon_i.$$

  Test for model misspecification by testing the residuals ε_i for correlation with either x_i or y_i.

  If no, the model is properly specified. Test if $c_1 = 1$. Done.
  If yes, the model is misspecified.

[4] In the second approach, accepting the model if $c_1 > 0$ is inconsistent with the model structure and will lead to false acceptances with a high probability. This is the reason this characterization of approach 2 is not emphasized in this comparison.

- Step 2. Take a higher order expansion in the true model. Add the next higher order terms: $\frac{\partial^2 f}{\partial x \partial y}(\beta, \alpha)(y_i - \alpha)(x_i - \beta)$ and $\frac{\partial^2 f}{2\partial y^2}(\beta, \alpha)(y_i - \alpha)^2$. The new regression model is:

$$Z_i = c_0 + c_1 f(x_i, \alpha) + c_2 y_i + c_3 x_i y_i + c_4 y_i^2 + \widehat{\varepsilon}_i.$$

  Note that all of the coefficients c_k change, except c_1. The error terms change as well. Test for model misspecification by testing the new residuals $\widehat{\varepsilon}_i$ to see if they are correlated with either x_i or y_i.

  If no, the model is properly specified. Test if $c_1 = 1$. Done.
  If yes, the model is misspecified.

- Step 3. Continue with higher order expansions until the residuals are white noise, and finally test the model to accept or reject the restricted version.

5 Corporate Finance Example

This section illustrates the error that could arise if one uses approach 2, but without explicitly considering the omitted variables problem. This example is hypothetical, selected only to illustrate the issues involved. Let's consider the determination of firm value, as discussed in Ross, Westerfield, and Jaffe [10], p. 487, problem 16.12.

$$V_L = V_U + \tau B - C(B) \quad where$$

V_L = value of levered firm,
V_U = value of unlevered firm,
T_C = personal tax rate of the marginal bondholder,
B = debt level of the firm,
$C(B)$ = present value of the costs of financial distress for the firm as a function of its debt level,

$$\tau = \left[\frac{1 - (1 - T_C)}{(1 - T_B)}\right] > 0,$$

and to make the model concrete, let us assume that

$$C(B) = cB + k(cB)^2$$

with k a constant.

For the sake of argument, assume that the modeler has developed a theory for capital structure, ignoring bankruptcy costs. The theory developed is

$$V_L = V_U + \tau B.$$

The scientist wants to test this theory by running a regression and holding constant for the omitted variables.

5.1 Empirical Model

Assume (for the sake of argument) that observable are V_L, V_U, τB and an estimate of bankruptcy costs, cB. We view V_L as measured with error. The scientist recognizes that bankruptcy costs are important, so he runs the regression attempting to control for bankruptcy costs:

$$V_L = \beta_1 V_U + \beta_2(\tau B) + \beta_3(cB) + \epsilon \tag{4}$$

and tests to see if $\beta_1 = 1$, $\beta_2 = 1$.

The scientist adds the third linear term cB to incorporate bankruptcy costs.

5.2 True Model

However, the true model is really

$$V_L = \beta_1 V_U + \beta_2(\tau B) + \beta_3(cB) + \beta_4(cB)^2 + \epsilon \tag{5}$$

with $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = -1$, $\beta_4 = -k$.

5.3 Analysis of Estimation Bias

To understand the bias in the estimates obtained, we see that this example is an omitted variables problem, see Maddala [8], p. 161 or Theil [11] p. 549. The estimated coefficients $\widehat{\beta}_j$ are biased, and the bias is

$$\widehat{\beta}_j = \beta_j + p_j\beta_4 = \beta_j - p_j k \tag{6}$$

where p_j are the coefficients obtained from running the regression

$$(cB)^2 = p_1 V_U + p_2(\tau B) + p_3(cB) + \eta.$$

One would expect that all of $p_j > 0$. That is, the variables are all positively correlated. For large sample sizes, one would certainly reject $\beta_1 = 1$, $\beta_2 = 1$, although the restricted version of the model is in fact true!

6 Conclusion

This paper relates two conflicting regression approaches for testing models in financial economics. The first approach, more often used in asset pricing, takes a restricted view of the model structure and rejects the model if omitted variables are significant in the estimated regression. The second approach, more often used in corporate finance and market microstructure, introduces omitted variables into the regression to control for the model's simplicity. The model can be accepted even if the omitted variables are significant in the estimated regression. Recent examples of this second approach include Hoberg [7] and Dittmar and Thakor [4].

We relate these two approaches by nesting them in a more general model structure. Within this more general structure, we show that the second approach can lead to a misspecified regression and an "omitted variables" problem, unless the regression errors are orthogonal to the independent variables included in the regression. A recent paper showing that this omitted variables problem can significantly bias the conclusions is Butler, Grullon, Weston [3], where the omitted variable is a structural shift in the economy. We propose a simple testing procedure that can identify and remove any model specification in this second approach. Its application, however, awaits subsequent research.

References

[1] M. Baker and J. Wurgler, 2002, Market timing and capital structure, *Journal of Finance*, 57, 1, 1 - 32.

[2] F. Black, M. Jensen and M. Scholes, 1972, The capital asset pricing model: some empirical tests, in *Studies in the Theory of Capital Markets*, Michael Jensen, ed. New York: Praeger, 79 - 121.

[3] A. Butler, G. Grullon, and J. Weston, 2006, Can managers successfully time the maturity structure of their debt issues?, *Journal of Finance*, 61, 4, 1731 - 1758.

[4] A. Dittmar and A. Thakor, 2007, Why do firms issue equity?, *Journal of Finance*, 62, 1, 1 - 54.

[5] D. Easley, N. Kiefer, M. O'Hara and J. Paperman, 1996, Liquidity, information and infrequently traded stocks, *Journal of Finance*, 60, 4, 1405 - 1436.

[6] E. Fama and K. French, 2002, Testing trade off and pecking order predictions about dividends and debt, *Review of Financial Studies*, 15, 1, 1 - 33.

[7] G. Hoberg, 2007, The underwriter persistence phenomenon, 62, 3, 1169 - 1206.

[8] G. S. Maddala, 1992, *Introduction to Econometrics*, 2nd edition, Prentice-Hall, N.J.

[9] R. Roll and S. Ross, 1980, An empirical investigation of the arbitrage pricing theory, *Journal of Finance*, 35, 5, 1073 - 1103.

[10] S. Ross, R. Westerfield, and J. Jaffe, 2002, *Corporate Finance*, 6th edition, McGraw-Hill Irwin, New York.

[11] H. Theil, 1971, *Principles of Econometrics*, John Wiley, New York.