Measuring financial integration via idiosyncratic risk: what effects are we really picking up?

David C. Parsley
Owen Graduate School of Management
Vanderbilt University
David.parsley@vanderbilt.edu

Christian Schlag
Johann Wolfgang Goethe-University
Frankfurt am Main
schlag@finance.uni-frankfurt.de

March 2006

ABSTRACT

We study the method proposed by Flood and Rose (FR, 2004, 2005) for checking for financial integration by estimating the risk-free rate using the idiosyncratic component of individual stock returns. Performing simulations with data with a known return generation process, we find that the FR methodology produces poor estimates of the risk-free rate, and hence the FR method fails to accept integration when true. We then show analytically that the FR method actually provides an estimate of the market return, and conclude the FR methodology would also falsely accept integration as long as the market returns in the two markets do not differ widely.

We thank Charles Engel and Hans Stoll for helpful comments on an earlier version.

1. Introduction

Flood and Rose (2005, hereafter "FR") propose a new methodology to measure financial integration based on an estimate of the risk-free rate of return (or equivalently, the expected intertemporal marginal rate of substitution, the pricing kernel, etc.).[1] A necessary condition for asset market integration is that there is only one risk-free rate of interest. A key insight made by FR is that idiosyncratic risk contains information on the risk-free rate. Given multiple estimates of idiosyncratic risk (e.g., from individual stocks, or portfolios), FR empirically construct confidence intervals and conduct tests for equality of the common risk-free rate across asset markets in the same country, e.g., NYSE and the Treasury Bill market, NASDAQ and NYSE (FR, 2004), or across markets in different countries, e.g., NYSE and TSE.

FR argue that in most empirical asset pricing applications, a common risk-free rate is *assumed* rather than *tested*; hence they propose a methodology to a test an important, underlying necessary condition for market integration (the equivalence of the risk-free rates prevailing in the two markets). In the case when the two markets are in the same country (like NASDAQ and NYSE – as in FR 2004), the results of this test are especially easy to interpret. Specifically, a rejection of the null implies the existence arbitrage opportunities. Given that their methodology can be implemented using commonly available data and statistical tools, their method may be repeated in many different contexts relatively easily.

The application of their methodology however, yields some surprising (at least to us) results. FR (2005) find that the NYSE and the U.S. bond markets are not integrated, since their regression methodology delivers significantly different estimates for the risk-free rate across these two markets. They also conclude that the NYSE and the Canadian stock market (TSE) are not integrated. Finally, in the NBER working paper version, FR (2004), they report that NASDAQ and NYSE are also not integrated. Since these markets are well developed and transaction costs are low, we find these results surprising.

The purpose of this paper is to explain what is driving these results. We proceed by applying FR's methodology to simulated data where we know with certainty the underlying data generation process. Our motivation derives from the observation that minimum variance

[1] FR take considerable care in motivating the importance of the subject to macroeconomics and international finance, from both a theoretical, and an empirical perspective, hence we do not repeat this exercise here (see also Chen and Knez, 1995).

is not the only desirable feature of an estimator. That is, it is reasonable to require that their procedure yield statistically accurate estimates of the risk-free rate embedded into our simulated data. For example, one could imagine an experiment where we took 1,000,000 measurements of an object of interest using a faulty recording device. Obviously reporting the mean and standard error of our individual estimates would not necessarily guarantee veracity.

Hence, as a check on FR's procedure, we perform their estimation (using a SUR approach) on a sample of synthetic individual stocks where we specify the data generation process (DGP). In our simulations, we find that the FR method performs poorly, since it rejects integration even on data created with perfect integration assumed in the DGP.

We explore possible reasons for these results and find that applying the FR methodology to our simulated data produces estimates of the risk-free rate which are highly correlated with one key parameter of our assumed DGP, i.e., the overall market return.[2] Of course, such negative findings from simulations cannot *explain* why the FR method performs poorly; for that, we must turn to a theoretical discussion of the FR method. Thus, in section 3, we provide an analytical resolution to the question in this paper's title. In particular, we show that their regression methodology estimates the risk-free rate with an equally weighted average of the returns of the individual returns being studied.

Moreover, based on this theoretical result it becomes immediately obvious that the FR method would also produce a "Type-II error". In the case where two markets actually cannot be integrated, because there are arbitrage opportunities due to differences in the risk-free rates, the FR approach can falsely fail to reject the null of integration, depending on whether the overall market returns are sufficiently close to each other.

2. Simulation: how close is the FR estimate to the 'true' risk-free rate?

This section describes our simulation exercise designed to check the FR methodology in a model economy. The motivation for this exercise is that, in an economy with perfectly

[2] FR (2005, p. 964), also note a "high correlation between $\{\hat{\delta}_t\}$ and the market return", but do not explore this in their paper. Marshall (2005) also notes the high correlation indicates econometric problems with the FR approach. He suggests the inclusion of an intercept term in the FR regressions to mitigate the problem of a high correlation between the FR estimates and the market return. We discuss this approach in the context of our simulation exercise presented in the next section.

integrated stock and bond markets, the estimation approach must yield a result that does not allow rejection of the null of integration (at standard confidence levels).

The setup of the simulation is rather simple. We assume an economy in which excess stock returns are generated by a one-factor market model, i.e. the excess return on a stock is given be the excess return on the market times the beta factor, plus an idiosyncratic term. The aggregate market return is normally distributed with an expected value of 10% p.a. and a volatility of 15% p.a. We analyze a sample period of 100 months, and a selection of 100 individual stocks. The risk-free rate is assumed to be 3% p.a. Individual stocks are characterized by their beta coefficients and the standard deviation of the idiosyncratic risk term. The beta coefficients for the 100 are sampled randomly from a uniform distribution over the interval (0.4;2.0), and the residual standard deviation is sampled randomly from a uniform distribution over the interval from 7.5% to 17.5% p.a.[3]

The simulation of the FR approach then proceeds in two steps. First, we generate the market return and the 100 individual stock returns for the 100 months in the sample. Then, we perform the FR regression via a SUR approach. Given our independence assumptions, this approach is equivalent to running 100 separate OLS regressions. This produces 100 independent estimates of the risk-free rate, which we test for equality with the "true" risk-free rate (3% as specified above).

The results can be summarized quickly. The mean value of the regression coefficients (corresponding to the average of the $\hat{\delta}_t$ in FR 2005) is 1.008319 per month, or 1.104525 annualized. So, for the model-economy we obtain an estimate that is far away from the risk-free rate of 3% or 1.03 annualized. The F-statistic for joint equality of the regression coefficients and the risk-free rate is significant at a p-level far below 0.1 percent. What is more striking however is that the average slope coefficient in the regression is so close to the market return. The correlation of the time series of our regression estimates with the simulated market return is a breath-taking 99.8%, so that we are forced to conclude that our estimator basically is the market return or, at least, an almost deterministic linear function of

[3] Sensitivity analyses have shown that the results of our simulation exercise are not particularly sensitive to the choice of these quantities.

the market return, an issue we will return to below.[4] As an aside, note that the results for the annualized slope estimate are almost identical to those in the FR paper, where the authors also comment on the (obviously surprisingly) high correlation of the estimates with the market return (see FR, 2005, p. 964).

As indicated in the introduction, Marshall (2005) argues that the FR approach might suffer from some fundamental problem due to the empirically observed high correlation of the estimates with the market return. In his comments, Marshall focuses his attention on the econometric aspects of the FR technique, and suggests the inclusion of an intercept term in the FR regression equation. As a check on this econometric approach, we re-estimated our artificial economy including individual intercepts. We find that the inclusion of the intercept does indeed reduce the correlation between the estimates and the market return to the insignificant value of 0.07. However, the inclusion of the intercept nevertheless does not correct the procedure. In particular, the parameter of interest, i.e., the estimate of the slope coefficient, comes out nowhere near the values it should have if it was to represent some sort of interest rate or compounding factor (or the assumed value of $1.03^{1/12}$). The mean is 2.1537, which amounts to a value of 9959.16 annualized. This result convincingly shows that just adding the intercept does nothing to solve the more fundamental problem inherent in the FR method.[5]

While these simulation results give a first indication that the FR methodology yields an estimate of something other than the quantity for which it was designed, simulation evidence clearly is not enough to show that the approach suffers from fundamental problems. We now turn to a theoretical analysis of the FR regression procedure.

3. Theoretical discussion

In this section, we derive an exact analytical expression for the key parameter of interest (δ_t) in the FR methodology, i.e., the risk-free rate of return. We adhere strictly to the

[4] Of course in any empirical application the correlation is likely to be lower since, e.g., there will not be a perfect correspondence between the sample of stocks chosen and those comprising the market index.

[5] Given the simulation results, we performed the theoretical analysis in the next section only for the model without intercept. Additionally, this is the model that FR (2005) discuss in detail in the empirical section of their paper.

FR (2005) set-up, and make no additional assumptions.[6] We begin with FR's regression equation 17 (page 960). In every month $t+1$ $(t=1,\ldots,T-1)$ in their sample period they estimate the following model:[7]

$$\frac{p_{t+1}^j}{\hat{p}_t^j} = \delta_t \left(\frac{p_t^j}{\hat{p}_t^j} \right) + u_{t+1}^j . \tag{1}$$

Here p_t^j stands for the price of asset j $(j=1,\ldots,J)$ in month t. Furthermore, we use, like FR, the following definitions:

$$\begin{aligned} p_{t+1}^j &= p_t^j \exp\left(r_{t+1}^j\right) \\ &= p_t^j \exp\left(r_f + \beta_j \left(r_{t+1}^m - r_f\right) + v_{t+1}^j\right). \\ p_t^j &= \hat{p}_t^j \exp\left(v_t^j\right) \end{aligned} \tag{2}$$

Note that the individual stock return $r_{t+1}^j \equiv r_f + \beta_j \left(r_{t+1}^m - r_f\right) + v_{t+1}^j$ is assumed to follow a standard market model. This specification is chosen only to simplify the exposition; any other factor model as the return generator would be just as suitable. Indeed, FR themselves stress that the choice of the return generating mechanism is immaterial for consistent estimation of δ_t (FR, 2005, p. 959).

With these preliminaries, the left hand side of equation (1) can be written as

$$\frac{p_{t+1}^j}{\hat{p}_t^j} = \exp(r_{t+1}^j + v_t^j) , \tag{3}$$

and the right hand side, as stated in FR, as

$$\frac{p_t^j}{\hat{p}_t^j} = \exp(v_t^j) . \tag{4}$$

So the regression equation (1) in month $t+1$ ultimately becomes

[6] There are several approaches one could take to point out the error of the FR (2005) approach. Charles Engel has suggested to us that another way to motivate our derivation is to point out that the expectation of the error term in (FR 2005) equation 14 is non-zero. That is, the FR procedure of forcing the covariance of the pricing kernel and the synthetic return into the error term is not as innocuous as FR claim.

[7] FR go to some detail to justify their choice of estimation method (OLS, GMM, and in the working paper version, also IV). This will turn out to be a redundant exercise.

$$\exp(r_{t+1}^j + v_t^j) = \delta_t \exp(v_t^j) + u_{t+1}^j \qquad j = 1,\ldots J, \tag{5}$$

or, after dividing through by $\exp(v_t^j)$,

$$\exp(r_{t+1}^j) = \delta_t + \tilde{u}_{t+1}^j \qquad j = 1,\ldots J, \tag{6}$$

with the new error term $\tilde{u}_{t+1}^j \equiv \exp(-v_t^j) u_{t+1}^j$.

Thus, the regression suggested by FR amounts to regressing individual stock returns on a constant and an error term. This implies for the OLS estimate[8]

$$\hat{\delta}_t = \frac{1}{J} \sum_{j=1}^{J} \exp(r_{t+1}^j), \tag{8}$$

i.e., the coefficient is equal to the average gross return of the stocks in the sample for the given month. This explains precisely, why FR not only do not obtain an estimate equal to the risk-free rate, but also observe that the time series of their monthly estimates are highly correlated with the market return. It is immediately clear that an estimate like the one presented in Eq. (8) is very close to the market return. If one samples the complete universe of stocks in an index, then the only (small) differences between the estimate in Eq. (8) and the market return come from the fact that, in general, the stocks in an index are not equally weighted and from a (usually very small) convexity correction, since $E[\exp(X)] \geq \exp[E(X)]$.

Finally, as pointed out in the introduction, these results also imply another problem with the FR approach. Assume that there are indeed arbitrage opportunities between two markets in that the respective risk-free rates are different. It should now be clear that the FR methodology would not detect this obvious case of non-integrated markets, as long as the market returns are sufficiently close.

4. Conclusions

Flood and Rose (2005) propose a methodology for checking for market integration by estimating the risk-free rate using the idiosyncratic component of individual stock returns. They apply their methodology to examine internal integration of the NYSE and the TSE, integration between the NYSE and the U.S. bond market, and integration between the TSE

[8] We illustrate the point using OLS for simplicity. However given the apparently trivial differences between estimators as reported in FR (2005), derivation for other estimation techniques seems unnecessary.

and the NYSE. They conclude these asset markets are internally integrated, but reject integration between the NYSE and the U.S. bond markets, and between the NYSE and the TSE. We find some of these results surprising; hence we check their methodology on simulated data.

We find that the FR methodology rejects integration in the simulated data, i.e., we statistically reject equality of the estimated risk-free rate and the "true" risk-free rate built into the data. Our simulations replicate another finding of FR (2005), namely that the estimates of the risk-free rate are highly correlated with the market return. Similarly, we point out that the FR methodology would falsely accept integration in data constructed with different risk-free rates. Thus we find the FR methodology subject to both Type I, and Type II errors.

We demonstrate this conclusively by deriving an exact analytic expression for the key parameter of interest (δ_t) in the FR methodology. We show that their regression methodology estimates the risk-free rate with an equally weighted average of the returns of the individual stocks in the sample, thus producing an estimate very close to the market (as opposed to the risk-free) return.

References

Chen, Z., Knez P., 1995, "Measurement of Market Integration and Arbitrage", Review of Financial Studies, v8, p. 287-325.

Flood, R., Rose, A., 2004, "Estimating the expected marginal rate of substitution: Exploiting idiosyncratic risk", National Bureau of Economic Research, working paper # 10805.

Flood, R., Rose, A., 2005, "Estimating the expected marginal rate of substitution: A systematic exploitation of idiosyncratic risk", Journal of Monetary Economics, v52, p. 951-969.

Marshall, D.A., 2005, "Comment on: Estimating the expected marginal rate of substitution: A systematic exploitation of idiosyncratic risk", Journal of Monetary Economics, v52, p. 971-979.