ARTICLE IN PRESS
EMPFIN-0870; No of Pages 19
Journal of Empirical Finance xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Journal of Empirical Finance

journal homepage: www.elsevier.com/locate/jempfin

The usefulness of cross-sectional dispersion for forecasting aggregate stock price volatility☆

Sung Je Byun

Financial Industry Studies Department, Federal Reserve Bank of Dallas, 2200 N. Pearl Street, Dallas, TX, USA

ARTICLE INFO

Article history:
Received 31 May 2015
Received in revised form 23 January 2016
Accepted 24 January 2016
Available online xxxx

JEL classification:
C55 Large Data Sets
Modeling and Analysis C58 Financial
Econometrics G12 Asset Pricing
Trading Volume
Bond Interest Rates G17 Financial
Forecasting and Simulation

Keywords:
Forecasting S&P 500 volatility
Cross-sectional dispersion
Aggregate idiosyncratic volatility
Large panel data model

ABSTRACT

Does cross-sectional dispersion in the returns of different stocks help forecast volatility of the S&P 500 index? This paper develops a model of stock returns where dispersion in returns across different stocks is modeled jointly with aggregate volatility. Although specifications that allow for feedback from cross-sectional dispersion to aggregate volatility have a better fit in sample, they prove not to be robust for purposes of out-of-sample forecasting. Using a full cross-section of stock returns jointly, however, I find that use of cross-sectional dispersion can help improve parameter estimates of a GARCH process for aggregate volatility to generate better forecasts both in sample and out of sample. Given this evidence, I conclude that cross-sectional information helps predict market volatility indirectly rather than directly entering in the data-generating process.

1. Introduction

Modeling and forecasting volatility is an important task and a popular research agenda in financial markets. Volatility models play key roles in the academic literature for testing the fundamental tradeoff between risk and return of financial assets and for investigating the causes and consequences of the volatility dynamics in the economy. Volatility forecasts have many practical applications as well. For example, volatility forecasts are used for market timing decisions, portfolio selection, risk management and pricing of financial derivatives such as options and forward contracts since risks are measured by the volatility of the financial asset returns.

Given its importance, there is a growing number of models and approaches for forecasting volatility in financial assets. Since Engle (1982) and Bollerslev (1986), Autoregressive Conditional Heteroskedasticity (ARCH) models are the most popular, formulating the volatility forecasts of a return as a function of known variables. Adopting the specific functional form and/or alternative explanatory variables in the volatility forecasts, there have been numerous extensions of ARCH models focusing on highlighted characteristics such as volatility persistence, asymmetry, long memory properties and a leptokurtic distribution of

☆ The views expressed are those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Dallas or the Federal Reserve System.
E-mail address: SungJe.Byun@dal.frb.org.

http://dx.doi.org/10.1016/j.jempfin.2016.01.013
0927-5398/

Please cite this article as: S. Byun, The usefulness of cross-sectional dispersion for forecasting aggregate stock price volatility, Journal of Empirical Finance (2016), http://dx.doi.org/10.1016/j.jempfin.2016.01.013

financial asset returns. While earlier researchers use variation over time in variables of interests, cross-sectional information began to be recognized as an important source for improving volatility forecasts. As a measure of dispersion in returns across individual stocks, Campbell et al. (2001) and Connor et al. (2006) suggested cross-sectional dispersion as an important source of individual stock volatility.[1] In a similar context, Hwang and Satchell (2005) tested whether cross-sectional dispersion helps forecast volatility of individual stock returns using GARCH-X models, where X refers to “cross-sectional dispersion.” Although better specified, they found a trivial improvement from GARCH-X for out-of-sample volatility forecasts, concluding that GARCH-X models do not necessarily outperform GARCH models in forecasting individual stock volatility.

In this paper, I investigate potential channels through which cross-sectional information might help predict aggregate volatility. I develop a model of individual returns that could be applied to the study of volatility in any financial assets, though the interest in this paper is in volatility of the stock market index such as the S&P 500 Composite Index (henceforth, S&P 500 index). The approach includes a GARCH-X model as a special case in which measures of cross-sectional dispersion appear in the equation for predicting aggregate volatility, an approach previously investigated by Hwang and Satchell (2005). Using returns of individual stocks in the S&P 500 index, I first confirm earlier researchers' findings that cross-sectional dispersion is time-varying, counter-cyclical and highly persistent with a few clusterings. Then, I study the direct role of cross-sectional dispersion in predicting the volatility of the S&P 500 index return using cross-sectional dispersion as an additional explanatory variable in the GARCH process. I find that although such GARCH-X specifications can improve in-sample forecasting accuracy, cross-sectional dispersion does not appear to be useful for out-of-sample forecasting.

Next, I investigate another channel in which cross-sectional information helps predict aggregate volatility by providing accurate parameter estimates. After jointly modeling the full cross-section of individual stock returns, I estimate population parameters in the bivariate GARCH process for aggregate volatility and cross-sectional dispersion. While sharing the same GARCH process with univariate GARCH, I find improved forecasting accuracies that are statistically significant, both in sample and out of sample, for mean squared error (MSE) and mean absolute error (MAE) loss criteria. Using White's (2006) conditional predictive ability tests, I show that by jointly utilizing cross-sectional information, it also provides more accurate out-of-sample volatility forecasts in times of recessions as well as during bear markets. I conclude that cross-sectional information helps predict market volatility indirectly insofar as it helps obtain accurate parameter estimates for volatility forecasts.

This paper is organized as follows. The following section describes a model of stock returns and clarifies the relation between models of aggregate stock volatility and the volatility of individual returns. After briefly introducing the data in Section 3, I describe the empirical properties of the cross-sectional volatility in the S&P 500 index return in Section 4. In Section 5, I investigate two potential channels whereby cross-sectional information could improve volatility forecasts of the S&P 500 index return: 1) cross-sectional dispersion as an additional explanatory variable as in GARCH-X, 2) cross-sectional dispersion as an aid in parameter estimation when individual stock returns are jointly modeled. Section 6 provides robustness checks by using an alternative volatility proxy, alternative measures for cross-sectional dispersion including Hwang and Satchell's (2005) cross-sectional market volatility and an alternative model specification using cross-sectional factors of Fama and French (1993) and Jegadeesh and Titman (1993). I also consider potential effects from ordinary cash dividends, extreme stock market events, asymmetric effects of volatility and individual stocks' sensitivities to stock return characteristics when predicting the volatility of the S&P 500 index return. All robustness checks confirm results provided in Section 5. Lastly, Section 7 concludes.

2. A simple model of stock returns

Let $r_{i,t}$ denote the monthly return on individual stock i measured in percent. For example, $r_{i,t} = -1.5$ means that stock i fell 1.5% from month $t-1$ to t. My interest is in characterizing aggregate market return as measured by some weighted average of individual returns[2]

$$r_t = \sum_{i=1}^{N} w_{i,t-1} r_{i,t},$$

where $w_{i,t-1}$ is a predetermined weight of a stock i's return in the evolution of the stock index return in period t. For example, $w_{i,t-1} = N^{-1}$ for an equal-weighted index and $w_{i,t-1} = p_{i,t-1}s_{i,t-1}/\sum_{j=1}^{N} p_{j,t-1}s_{j,t-1}$ for a value-weighted index where $p_{i,t-1}$ is stock i's price and $s_{i,t-1}$ is its number of outstanding shares at time $t-1$.

Since Engle (1982) and Bollerslev (1986), Autoregressive Conditional Heteroskedasticity (ARCH) models are the most popular, formulating the volatility forecasts of a stock index return as a function of known variables. One approach for improving volatilty forecasts would be to fit an univariate GARCH-X(1,1) model to the aggregate return:

$$r_t = \phi^0 + \phi^1 r_{t-1} + u_t, \tag{1}$$

[1] The contribution of cross-sectional dispersion is referred to as firm-specific (idiosyncratic) volatility in Campbell et al. (2001) and common heteroscedasticity in asset-specific returns in Connor et al. (2006).

[2] When a stock i is newly added in the stock index in period t, there is no contribution of a stock i's return $r_{i,t}$ to the stock index return r_t.

Please cite this article as: S. Byun, The usefulness of cross-sectional dispersion for forecasting aggregate stock price volatility, Journal of Empirical Finance (2016), http://dx.doi.org/10.1016/j.jempfin.2016.01.013