Spring 5-15-2015

Two Essays in Financial Economics

Kyle J. Putnam
University of New Orleans, kjputnam@uno.edu

Two Essays in Financial Economics

A Dissertation

Submitted to the Graduate Faculty of the
University of New Orleans
In partial fulfillment of the
Requirements for the degree of

Doctor of Philosophy
in
Financial Economics

By

Kyle J. Putnam

B.A. George Fox University, 2008
M.S. University of New Orleans, 2013

May, 2015

Dedication

To my parents, Jerry and Gale, who instilled in me the value of education and a good work ethic. Thank you for your selfless and loving support of me over the years as I have pursued my goals.

To my darling wife, Julie, who without her love and support none of this would have been possible. Words are not adequate enough to thank you for your endless encouragement, love, and sacrifice while I pursed my Ph.D.; I have earned this degree for you.

Acknowledgement

Above all I am eternally grateful for the opportunities and grace which God has bestowed upon me during my time in the Ph.D. program. Of course, completing my Ph.D. would not have been possible without the support and assistance of the many wonderful people in my life, all of whom I would like to extend my sincerest and deepest appreciation towards. I am grateful to my family and friends who have been unfailing in their love, support, and kindness. Special thanks go to my dissertation committee co-chairs, Dr. Astuyuki Naka and Dr. Mohammad K. Hassan, who have supervised my work and been very helpful and supportive throughout the process. I would also like to extend my gratitude to all the faculty members of the Department of Economics and Finance who have helped to mold me into a better researcher and teacher. Lastly, a special acknowledgement goes to my Ph.D. colleague and good friend Ramesh Adhikari, whose helpfulness, kindness, and friendship cannot go unmentioned.

Table of Contents

List of Figures

List of Tables

Abstract

The following dissertation contains two distinct empirical essays which contribute to the overall field of Financial Economics. Chapter 1, entitled "The Determinants of Dynamic Dependence: An Analysis of Commodity Futures and Equity Markets," examines the determinants of the dynamic equity-commodity return correlations between five commodity futures sub-sectors (energy, foods and fibers, grains and oilseeds, livestock, and precious metals) and a value-weighted equity market index (S&P 500). The study utilizes the traditional DCC model, as well as three time-varying copulas: (i) the normal copula, (ii) the student's t copula, and (iii) the rotated-gumbel copula as dependence measures. Subsequently, the determinants of these various dependence measures are explored by analyzing several macroeconomic, financial, and speculation variables over different sample periods. Results indicate that the dynamic equity-commodity correlations for the energy, grains and oilseeds, precious metals, and to a lesser extent the foods and fibers, sub-sectors have become increasingly explainable by broad macroeconomic and financial market indicators, particularly after May 2003. Furthermore, these variables exhibit heterogeneous effects in terms of both magnitude and sign on each sub-sectors' equity-commodity correlation structure. Interestingly, the effects of increased financial market speculation are found to be extremely varied among the five sub-sectors. These results have important implications for portfolio selection, price formation, and risk management. Chapter 2, entitled, "US Community Bank Failure: An Empirical Investigation," examines the declining, but still pivotal role, of the US community banking industry. The study utilizes survival analysis to determine which accounting and macroeconomic variables help to predict community bank failure. Federal Deposit Insurance Corporation and Federal Reserve Bank data are utilized to compare 452 community banks which failed between 2000 and 2013, relative to a sample of surviving community banks. Empirical results indicate that smaller banks are less likely to fail than their larger community bank counterparts. Additionally, several unique bank-specific indicators of failure emerge which relate to asset quality and liquidity, as well as earnings ratios. Moreover, results show that the use of the macroeconomic indicator of liquidity, the TED spread, provides a substantial improvement in modeling predictive community bank failure.

Keywords: Dynamic Dependence, Commodity Futures, DCC, Copulas, Failure Risk, US Community Banks, Survival Analysis

Chapter 1

The Determinants of Dynamic Dependence: An Analysis of Commodity Futures and Equity Markets

1. INTRODUCTION

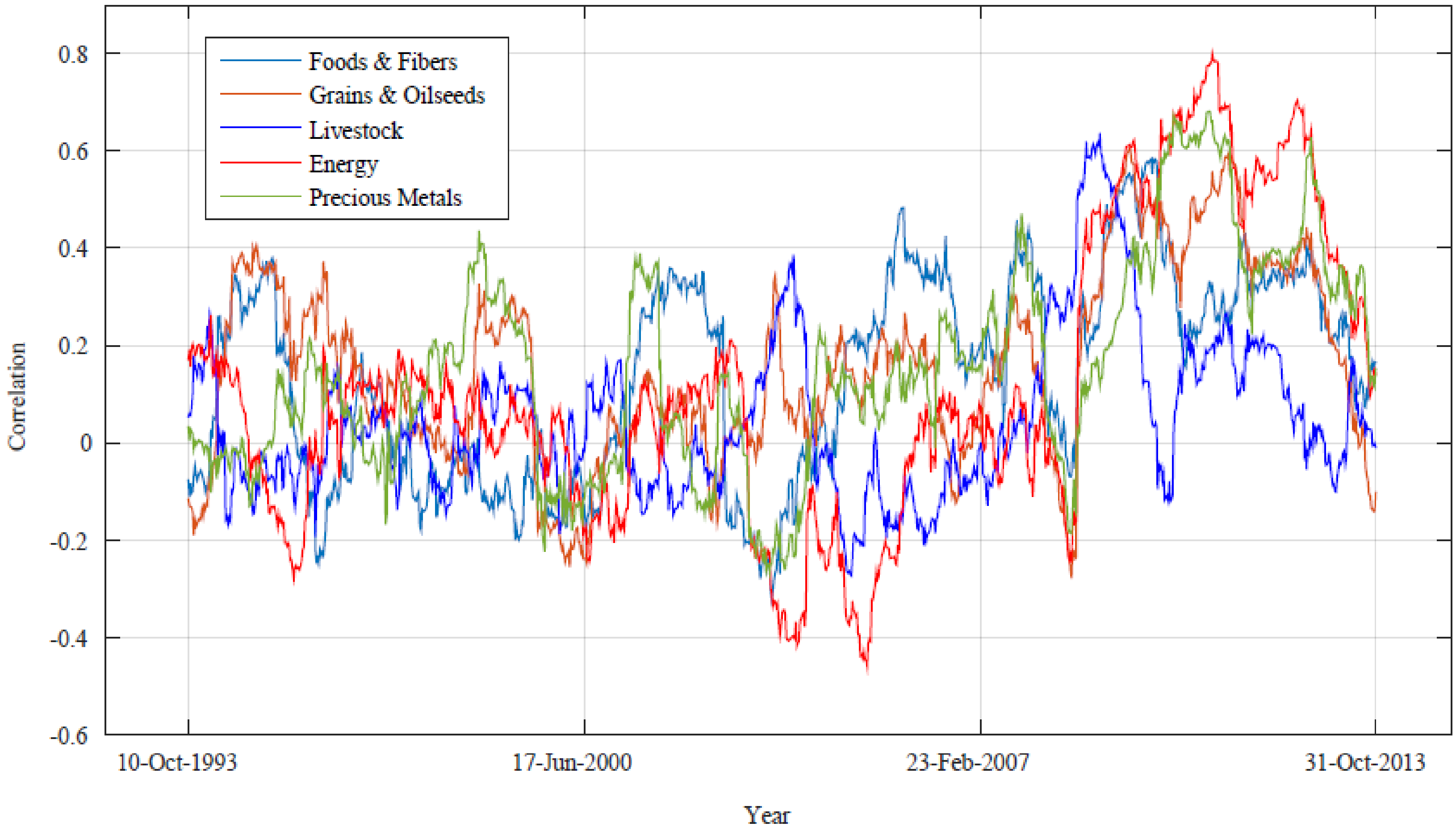
Numerous strands of literature have emerged over the last decade which have touted commodity futures as useful additions to investor portfolios for diversification, inflation hedging, and risk management purposes (see Gorton and Rouwenhorst, 2006; Buyuksahin et al., 2010; Conover et al., 2010; Jensen et al., 2000). Moreover, as documented by Erb and Harvey (2006), among others, investment in commodity futures can provide "equity like" returns through a "tactical" rebalancing strategy. These attractive investment benefits stem from the theoretical motives that commodities, and in turn commodity futures, form an alternative asset class to that of equity and bond markets. Thus, the financially transformed fungible raw materials, in theory, are expected to exhibit little (or even negative) correlation with the more traditional asset classes. The reason for this low correlation is that the underlying factors which drive the commodity futures prices, such as weather, supply and demand constraints, geopolitical conditions, and event risk, are very different, if not completely segmented, from those factors which drive the value of the equity and bond markets (see Symeonidis et al., 2012).

However, an ever-growing strand of literature posits that the financialization of the commodity markets—the process whereby the raw materials have been transformed from mere goods into widely (popular) tradable financial instruments—occurring in the early 2000's, has resulted in increased integration with more traditional asset classes (see Buyuksahin et al., 2009; Daskalaki and Skiadopoulos, 2011; Tang and Xiong, 2012; Silvennoinen and Thorp, 2013). Moreover, this financialization is a result of increased investor participation over the last decade (see Buyuksahin and Robe, 2013) spurred by the potential benefits of commodity-related investment. According to the Commodity Futures Trading Commission (CFTC), the total value of different commodity index-related instruments purchased by institutional investors, including pension funds, endowments, trusts, and banks, increased from $10 billion in 2000 to a staggering $256 billion by mid-2011. In other words, increased investor interest in commodity futures, particularly by speculators, which is motivated by the belief that the unique asset class offers steadfast diversification and hedging opportunities in market downturns, has weakened the potential advantages of commodity futures investment as the shocks from the conventional asset markets enter the commodity futures price dynamics through the increased dependence of the return structure.

However, not all commodities, and hence commodity futures, are created equal—some commodities are storable goods while other are not, and what is more, some commodities serve as intermediate goods while others are merely input goods. Hence, due to these fundamental differences, the factors which drive the return dependence between commodity futures and the other asset classes are likely heterogeneous across different sub-sectors. In fact, such differences can be readily viewed by looking at simple 52-week rolling correlations between equity and commodity returns. Figure 1 provides correlations between the S&P 500 and five different commodity sub-sectors. We note that none of the sub-sectors' correlations move in lock-step with each other (some groups actually move in opposite directions), and the magnitude of the change in dynamic correlations varies greatly by group. Together, these observations motivate our choice to examine the issue of time-varying dependence and its determinants between the commodity futures markets and the equity markets at the sub-sector level. An understanding of the determinants of the various correlations has important implications for portfolio selection,

Figure 1

52-week Rolling Correlations between S&P 500 and Commodity Sub-sectors



price formation, risk management, as well as aiding in the comprehension of the overall futures market, particularly for non-index commodity futures investors. Furthermore, knowledge of how the equity-commodity determinants have evolved over time, the sensitivity of these factors to different forms of dependence measures, and the nature of the return relationship between commodity futures and equity markets (i.e. symmetrical vs. asymmetrical) will provide investors, particularly those involved in commodity futures a more detailed level of understanding of the overall market. This research also provides policymakers information on how different markets react to shocks, hence providing empirical evidence for any potential reforms of the US financial system.

Recent studies on the determinants of equity-commodity return correlations generally utilize a single commodity index composed of futures returns from numerous different sub-sectors (see Buyuksahin and Robe, 2013; Bhardwaj and Dunsby, 2013; Delatte and Lopez, 2013). Moreover, these indices, such as the well-known Goldman Sachs Commodity Index (SP-GSCI) and the Dow Jones UBS Commodity Index (DJ-UBS), tend to put more weight on certain commodity futures (such as energy or agriculture) and less weight on others, hence shifting, and effectively decreasing, the importance of other sub-sectors of the futures market. While studies which implement such broad commodity futures indices in their analysis have uncovered both interesting and valuable contributions to the literature sect on commodities as investments, we feel that the heterogeneous nature of commodities, in general, provides sufficient motivation to further investigate the futures at a more disaggregated (i.e. sub-sector) level in an effort to reveal the distinct determinants of the dynamic equity-commodity return correlations. For instance, the factors which explain the equity-commodity correlations for the energy sub-sector may have a differential impact than the factors which explain the correlations for say, livestock. A commodity futures index makes it virtually impossible to detect and disentangle such effects, but an analysis of the various sub-sectors highlights such relevant information, providing active traders in the commodity futures market, particularly those who do not merely invest in index-related products, invaluable information regarding the futures return behavior and the potential for diversification benefits and/or speculation profits.

However, determining the factors which explain the dependence of the equity-commodity return correlations for the various sub-sectors is complicated by the issue of identifying the appropriate nature of dependence among the two asset classes. It is well-documented that asset classes are not normally distributed (see Longin and Solnik, 2001), thus simple correlation coefficients are not sufficient to properly measure the true relationship between returns. Further, many empirical studies tend to impose the, somewhat unrealistic, assumption of time-stability on asset relationships. In order to account for these problematic issues, a recent study by Buyuksahin and Robe (2013) implement the popular time-varying dynamic conditional correlation (DCC) dependence measure (see Engle, 2002), the likes of which they use to show that the correlation between rates of return on broad market investible commodity and equity indices have increased as a result of greater participation by speculative hedge funds. However, the DCC model imposes the assumption of common dynamics among all assets used (see Billio et al., 2006). This particular restriction may or may not be true, but the imposition that the correlations of commodity futures are identical to US equity indices seems somewhat impractical. In order to overcome these previously ascribed pitfalls and assumptions associated with estimating asset correlations we appeal to the alternative copula approach which provides a dynamic measure of financial market comovements. This approach disentangles the unique characteristics of each return series from the dependence structure which links them together and allows for a range of models which capture different forms of dependence between variables. The dependence structure estimated via copula is more robust in the sense that the approach separates the dependence structure from the choice of marginal distributions. Moreover, the copula approach does not require elliptically distributed returns and is invariant with respect to increasing and continuous transformations of the marginals.[1]

In this paper, we calculate the dynamic dependence structure between the returns of five different commodity futures sub-sectors (energy, foods and fibers, grains and oilseeds, livestock, and precious metals) and a well-known value-weighted equity market index (S&P 500).[2] We employ the DCC model, as in Buyuksahin and Robe (2013), as a baseline approach to our investigation of the determinants of equity-commodity correlations, as well as three time-varying copulas. In particular, we analyze (i) the normal copula—a symmetrical and frequent dependence structure which has no tail dependence, (ii) the student's t copula—a symmetrical but non-zero tail dependence structure which nests the normal copula, and (iii) the rotated-gumbel copula—a left tail, non-linear, asymmetrical dependence structure, which is mostly present during extreme negative events. Practically speaking, these copulas represent the most relevant shapes for finance and are frequently used in empirical papers (see Embrechts et al., 2002; Patton, 2004; Rosenberg and Schuermann, 2006; Patton, 2009; Chollete et al., 2011; Aloui et al., 2011; Delatte and Lopez, 2013). We then explore the determinants of these various dependence measures by analyzing several comprehensive macroeconomic, financial market, and speculation variables over the period October 1992 to October 2013, via sub-sample and dummy variable regression analysis.

[1] Manner and Reznikova (2012) provide an extensive survey on time-varying copulas and their properties.

[2] We also calculate results using the value-weighted Russell 3000 equity index and find very similar univariate and regression results. Therefore, for both brevity and space we only discuss and report the equity-commodity results regarding the S&P 500.

Our examination finds that while copulas offer a more robust measure of time-varying dependence there are numerous similarities between the DCC model and the copula dependence measures. We document that the dynamic equity-commodity return correlations for the energy, grains and oilseeds, precious metals, and to a lesser extent the foods and fibers sub-sectors have become increasingly explainable by broad macroeconomic and financial market indicators, particularly after the period May 2003. This evolution of explanatory variables coincides with the financialization of the commodities market, whereby commodity futures prices, and hence returns, behave in a manner more strongly associated with traditional asset class returns. Contrastingly, the livestock sub-sector return correlations seem to be much less explainable using the broad market indicators. Additionally, we document that increased participation by financial market speculators is not a primary determinant for all sub-sectors' equity-commodity return correlations, as posited by previous literature. Though sensitive to the dependence measure, the energy and foods and fibers sub-sectors generally exhibit a positive statistically significant speculation coefficient. This suggests, in line with prior literature, that an increase in speculation, caused by non-commercial participants, increases dynamic equity-commodity correlations. Contrastingly, the grains and oilseeds sub-sector generally exhibits a negative and significant speculation coefficient; this means that an increase in speculation actually causes a decrease in dynamic equity-commodity correlations. Interestingly, both livestock and precious metals generally yield insignificant speculation coefficients. These findings are of particular importance given that prior work has acknowledged that increased participation, particularly by speculators, helps to predict observed long-run fluctuations in dynamic commodity-equity correlations.

The rest of this paper is organized as follows. Section 2 provides a pertinent review of the literature on the commodity-equity dependence relationship, as well as our contribution. Section 3 focuses on the methodology and dataset. Section 4 describes our empirical regression analysis and provides the results over all sample periods. Section 5 offers concluding remarks.

2. LITERATURE REVIEW

In a flight-to-quality argument, Chong and Miffre (2010) document that over the period 1981-2006, the correlations between equities and individual commodity futures tend to fall both over time, in general, and tempestuous financial market periods. Buyuksahin et al. (2010) document a similar result over the early 1990's to mid-2000's while investigating structural shift in correlation dynamics over both calm and tumultuous financial market periods. In particular, they find a lack of "greater return co-movement across equities and commodities [which] suggests that commodities should retain their role as a portfolio diversification tool."

However, much more recent research finds contradictory conclusions regarding the movement of these correlations. For instance, Silvennoinen and Thorp (2013) find that conditional volatility and correlation dynamics for returns to commodity futures, stocks, and bonds have become increasingly integrated over the period 1990-2009. Furthermore, they note a structural break in conditional correlations occurring in the late 1990's. Buyuksahin and Robe (2011) document significant changes in the make-up of the open interest between 2000 and 2010 and show that these changes impact asset pricing for the energy futures market. Specifically, they find that the dynamic conditional correlations between the rates of return on energy and stock market indices increase significantly from greater activity by speculators and hedge funds.

Interestingly, there is a growing strand of literary evidence that links the growth of index funds and other investment vehicles in the commodity futures market as the means of increased integration between the commodity market and the stock and bond markets; this integration has effectively reduced or diminished the sought after benefits of commodities. Recent work by Tang and Xiong (2012) finds that since the early 2000's the futures prices of non-energy commodities in the US have become significantly more correlated with oil futures prices. They argue that this increased integration, or comovement, is largely a reflection of the financialization of commodity markets. Furthermore, they show that this trend is more pronounced for commodities in the popular SP-GSCI and DJ-UBS commodity indices, which they attribute to the growing prominence of index trading. Buyuksahin and Harris (2011) look at the impact of financialization and speculation in the crude oil futures market; however, their analysis finds little evidence that hedge funds or other non-commercial speculators position changes cause price changes. They conclude that fundamentals and not speculation were most likely behind the 2004-2008, boom-bust commodity price cycle. Yet, numerous other studies exist which investigate, and attribute, financial speculation as a primary determinant of commodity spot price correlation, but these papers are largely confined to the investigation of the crude oil markets (see Hamilton, 2009; Fattouh et al., 2013; Kilian and Murphy, 2014) or industrial metals markets (see Korniotis, 2009). More recently, Buyuksahin and Robe (2013) implement a unique non-public dataset of trader positions in US commodity futures which focuses on the trading activity of speculators. They document that "excess speculation" by investor participants, especially by hedge funds, is positively related to the commodity returns' (index) increased correlation with equity markets. Furthermore, they find that the strength of the commodity-equity linkages has fluctuated substantially over the last 20 years, but that the activities of speculators, helps to predict observed long-run fluctuations in the dynamic commodity-equity correlation.

Given that both theory and empirical work predict no common risk factor structure in the cross-section of commodity futures risk premiums (see Daskalaki et al., 2014) we demonstrate that an analysis of commodity futures within their respective sub-sectors provides a much more meaningful analysis of the futures markets and its determinants. Furthermore, given that there is some dispersion in the literature about the comovement of the equity-commodity correlations, we employ several different measures of dependence. This methodological choice is based on the discussion of Delatte and Lopez (2013) who posit that a lack of consensus regarding the correlation structure between commodity futures and traditional asset market returns is due to the different dependence measures considered. Overall, we contribute to dual strands of literature. First, we explore several potential broad macroeconomic, financial, and speculation variables as determinants of the time-varying commodity-equity correlations. However, in contrast to prior work, we analyze each of the commodity futures sub-sectors individually. Second, given the pronounced increase in participation of financial traders in the commodity futures market in the early 2000's, we analyze the evolution of these factors across different sub-sample periods, for each sub-sector, to see how they have changed. Third, and finally, we investigate each of the dependence measures of the commodity-equity relationship in a regression setting, which includes both the DCC as a baseline approach, and three popular copulas in finance. This approach allows us the advantage of viewing our determinants against different forms of dependence structures and viewing the sensitivity of the determinants against different dependent variables. Overall, we aim to bring to light new facts regarding the commodity futures market.

3. METHODOLOGY AND DATASET

3.1 Measures of Dependence

The DCC framework, of Engle (2002), has become a largely popular approach to measuring the dependence structure between different financial assets. Notably, this dependence measure relies on the marginal distribution of returns. Hence, some empirical studies have taken a different approach to estimating the dependence structure using a copula methodology which, in contrast to the DCC approach, separates the dependence structure from the choice of marginal distribution creating a more robust approach to measuring dependence. Though the copula methodology is widely known and has been around for quite some time, its application to financial markets has become increasingly momentous in finance and risk management valuation within the last decade (see Patton, 2006; Kole et al., 2007; Chollete et al., 2010; Aloui et al., 2011) as copulas provide an important way to appropriately define a correlation structure, which may be non-linear, between different variables. We employ the commonly implemented DCC dependence measure as a baseline approach to our investigation, as well as three time-varying copulas popular in the field of finance—the normal copula, the student's t copula, and the rotated-gumbel copula.[3]

3.1.1 DCC Model

The multivariate GARCH model with DCC, a process whereby correlations are driven by the cross product of the lagged standardized residuals and an autoregressive term, was initially proposed by Engle (2002) and has since become a mainstream econometric methodology in finance and related applications. Briefly, the model is specified as:

[3] We disregard the constant dependence structure and focus solely on time-varying relationships as copious amounts of prior work have found the dependence relation among financial assets to indicate that it is anything but constant (see Erb et al., 1994; Longin and Solnik, 1995; Engle, 2002).

$$H_t = D_t R_t D_t \quad (1)$$

where, $D_t = \text{diag}\{\sqrt{h_{i,t}}\}$, R_t is a time-varying correlation matrix containing conditional correlations, and the expressions for h, the conditional standard deviations, are generally thought of as univariate GARCH models, but can include functions of other variables in the system as either pre-determined or exogenous.[4] We outline the details of our estimation procedure for the DCC model in Section 3.2.1.

3.1.2 Copula Functions

Copulas provide a convenient way to join or "couple" the marginal distributions of random variables into a joint distribution. Conversely, they can also allow one to separate a joint distribution into two contributions: the marginal distribution of each variable and the copula which combines these into a joint distribution (see Sklar, 1959). Copulas generally have a convenient parametric form and provide a large degree of flexibility in the specification of the marginal distributions and their dependence structure. Further, the choice of copula provides a great deal of control over what parts of the distribution the variables are most strongly associated; this convenience is particularly intriguing to market practitioners who are concerned with strong left tail dependence (i.e. the comovement of asset returns during market crises).

The theorem of Sklar (1959) illuminates the role copulas play in the relationship between multivariate distribution functions and their univariate marginals. Formally, in the bivariate case, if $F(X_{1t}, X_{2t})$ is a joint distribution function with marginal distribution functions $F_1(X_{1t})$ and $F_2(X_{2t})$, for random variables X_{1t} and X_{2t}, then there exists a copula, $C(u, v)$, mapping the marginal distributions of X_{1t} and X_{2t} to their joint distribution:

[4] Given the popularity of the DCC model we refer the interested reader to Engle (2002) for additional details regarding the technical notes of the model and estimation procedure.

$$F(X_{1t}, X_{2t}) = C\big(F_1(X_{1t}), F_2(X_{2t})\big) \quad (2)$$

If $F_1(X_{1t})$ and $F_2(X_{2t})$ are continuous, then the copula is unique, otherwise, the copula will not necessarily be unique. Thus, in the bivariate case, that means:

$$C(u, v) = \Pr[U \leq u, V \leq v] \quad (3)$$

where, U and V are uniformly distributed on [0,1].[5] Equation (2) explicitly highlights the practicality of copulas, in that one can simplify the analysis of dependence for a particular joint (return) distribution, $F(X_{1t}, X_{2t})$, by merely studying the copula. Conversely, if $C(u, v)$ is a copula, and F_1 and F_2 are univariate distribution functions, then $F(X_{1t}, X_{2t})$ is a joint distribution function with $F_1(X_{1t})$ and $F_2(X_{2t})$. Assuming that each marginal distribution is continuous and strictly increasing, we can write the copula as:

$$C(u, v) = F(F_{1^{-1}}(u), F_{2^{-1}}(v)) \quad (4)$$

where, $u = F_1(X_{1t}) \Leftrightarrow X_{1t} = F_{1^{-1}}(u)$ and $v = F_2(X_{2t}) \Leftrightarrow X_{2t} = F_{2^{-1}}(v)$ holds. Furthermore, assuming the marginals can be modeled parametrically, the probability integral transformation of equation (2) is given as:

$$U_{it} = F_i(X_{it}; \phi_i) \quad (5)$$

where, ϕ_i is the vector of parameters. The function $F_i(X_{it}; \phi_i)$ can be a conditional distribution (as it is in our analysis), where X_{it} is modeled by an ARMA-GARCH model, whose residuals are treated as independent and identically distributed (i.i.d.) random variables.[6] Following Manner and Reznikova (2012), it is also assumed that each variable depends only on its own past, but not

[5] While copulas also work in the multivariate context, we give our primary attention to the bivariate case.
[6] It is also assumed that the copula belongs to a parametric family $C_\theta, \theta \in \Theta \subset \mathbb{R}^K$.

on the past of the other variable, and that only instantaneous causality between the two variables exists. This assumption implies that the parameters of the copula are separate from the parameters of the marginal distributions.

Given, once again, that the copula function and the marginals are continuous, the following equation for the join probability density function (PDF) holds:

$$f(X_{1t}, X_{2t}) = c(U_{1t}, U_{2t}; \theta) \prod_{i=1}^{2} f_i(X_{it}; \phi_i) \tag{6}$$

where, $c(\cdot,\ \cdot)$ is the copula density. Further, assuming a sample for X_{1t} and X_{2t} where, $t = 1, \dots, T$, then the log-likelihood function is given as:

$$L(\theta, \phi) = \sum_{t=1}^{T} \{\log c(U_{1t}, U_{2t}; \theta) + \log f_1(X_{1t}; \phi_1) + \log f_2(X_{2t}; \phi_2)\} \tag{7}$$

This statement is equivalent to:

$$L(\theta, \phi) = L_C(\theta, \phi) + L_{X_1}(\phi_1) + L_{X_2}(\phi_2) \tag{8}$$

where, $\phi = (\phi_1', \phi_2')'$. Hence, the full log-likelihood function $L(\theta, \phi)$ can be split into two parts, the copula likelihood $L_C(\theta, \phi)$ and the likelihood of the marginals $L_{X_1}(\phi_1)$ and $L_{X_2}(\phi_2)$. The parameters θ and ϕ are estimated via a two-step process proposed by Genest et al. (1995). First, since the marginal models are unknown, the marginal distributions are estimated with the empirical CDF, based on the i.i.d. of the residuals, via the following form:

$$u = \widehat{F_1}(X_{1t}) = \frac{1}{n+1} \sum_{j=1}^{n} 1_{\{X_{1,t-j} \leq x_{1t}\}} \text{ and } v = \widehat{F_2}(X_{2t}) = \frac{1}{n+1} \sum_{j=1}^{n} 1_{\{X_{2,t-j} \leq x_{2t}\}} \tag{9}$$

Second, the copula parameters are estimated based on the rank of the data by maximizing the corresponding copula likelihood function given the results from the first step. This method proves useful in that it is robust to misspecification of the marginals which can cause biased estimates of the copula parameter.[7]

Patton (2006) proposes an extension of the copula model where the time-varying dependence parameter of a copula is a function of an autoregressive term, which captures persistence in the dependence term, and a forcing variable, which captures any variation in dependence. We follow Patton's extension to facilitate our analysis. For the normal and student's t copula, the evolution equation for the dependence parameter, ρ_t, is given as:

$$\rho_t = \bar{\Lambda}\left(\omega_\rho + \beta_\rho(\rho_{t-1}) + \alpha\frac{1}{n}\sum_{j=1}^{n}\phi^{-1}\big(F_1(X_{1,t-j})\big)\phi^{-1}\big(F_2(X_{2,t-j})\big)\right) \tag{10}$$

where, $\bar{\Lambda}(x) = \frac{(1-e^{-x})}{(1+e^{-x})}$ is a modified logistic transformation, designed to keep the correlation parameter ρ_t between (-1,1) at all times, and n is an arbitrary window length.[8] The average of the product of the last n observations of the transformed variables is the forcing variable. For the rotated-gumbel copula, the evolution equation for the dependence parameter, φ_t, is given as:

$$\varphi_t = \bar{\Lambda}\left(\omega_\rho + \beta_\rho(\rho_{t-1}) + \alpha\frac{1}{n}\sum_{j=1}^{n}\left|\big(F_1(X_{1,t-j})\big) - \big(F_2(X_{2,t-j})\big)\right|\right) \tag{11}$$

where, $\bar{\Lambda}(x)$ is a modified logistic transformation to ensure the parameter always remains in its domain, for the rotated-gumbel copula ($\varphi_t = \delta_t$) it is $1 + e^{-x}$. In this instance, the mean

[7] The theoretical properties of this estimator in a time series are derived by Chen and Fan (2006).

[8] We follow the dynamic framework methodology proposed by Creal et al. (2013) for the programming procedure of the student's t copula. The authors derive a Generalized Autoregressive Score (GAS) specification for the time-varying correlation parameter, ρ_t, using the density of the Gaussian (normal) copula, following Patton (2006).

absolute difference of the transformed variables over the previous n periods is the forcing variable.[9]

The merits and importance of the copula methodology in our analysis comes from the convenience (and ability) to impose a particular distributional dependence structure between our two variables of interest (i.e. commodity futures returns and equity returns). For instance, we can observe both symmetrical dynamics and asymmetrical dynamics. Consequently, we can also measure the strength of the relation with the appropriate density function; however, our primary objective is to measure the time-varying relation between our two choice variables and determine the factors which influence the particular dependency relationship, as well as examine the evolution of these factors over time.[10] While prior research has focused on a number of different parametric copula specifications, we focus on three types in our investigation of the determinants between commodity futures sub-sectors and equity returns: the normal, the student's t, and the rotated-gumbel. The normal copula specification, with zero tail dependence, is a common distributional assumption in finance and provides a reasonable benchmark for our analysis. Further, it provides a practical basis in which to compare the results from the baseline approach using the DCC dependence measure. The student's t copula is a useful measure as it has symmetric but non-zero tail dependence and consequently nests the normal copula. Finally, the rotated-gumbel copula is appealing because it provides the ability to measure the potential for the joint occurrence of left-tail extreme events, or lower tail dependence; that is, it captures the comovement of the return series jointly taking extremely low values. The rotated-gumbel has non-linear dependence as well as asymmetric tail dependence present during extreme negative events (i.e. the mass in the left tail is far larger than the mass in the right tail) and is a member of

[9] See Manner and Reznikova (2012) for additional details.
[10] Table A1 in Appendix A provides the best fit measures based on the log-likelihood criteria for all three copula functions, for each sub-sector, over the full sample period.

the extreme value copula family. Longin and Solnik (2001) find evidence of both extreme and asymmetrical forces at work in asset markets, thus, it seems reasonable to investigate the presence and determinants of these effects in equity-commodity correlations. In particular, investors are very much interested in how different markets commove together during severe downturns or crisis situations, as strong market comovements indicate a lack of diversification benefits, and weak comovements indicate the contrary. As noted in Chollete et al. (2011), practically speaking, these copulas are the most important shapes for finance as they represent a large subset of those implemented in empirical work. The corresponding copula functions and their dependence parameters are outlined in Table 1.

3.2 Estimation Procedure and Dataset

Our primary interest is the determinants of the dynamic equity-commodity return correlations for the five commodity futures sub-sectors, both over the entire sample period and two sub-sample periods so as to gauge the evolution of the determinants. Given this, we employ two conditional-based methodologies in order to obtain dynamically efficient estimates of the intensity of the equity-commodity return comovements. First, we implement the well-known DCC methodology, which Buyuksahin and Robe (2013) use in a similar vein of research. Second, we implement the copula methodology which yields a more robust approach to measuring dependence over different portions of the return distribution.

In order to facilitate both types of analyses we have to construct return series for both the futures and equity indices. We extract daily price data from the Commodity Research Bureau (CRB) database, over the period October 1992 to October 2013, for each of the individual

Table 1
Copula Distributions

	Copula	Parameter Range
Normal		
	$C_N(u,v;\rho)=\Phi_\rho\left(\Phi^{-1}\left(F_1(X_{1t})\right),\Phi^{-1}\left(F_2(X_{2t})\right)\right)$	$\rho\,\epsilon(-1,1)$
Student's t		
	$C_t(u,v;\rho,d)=t_{d,\rho}\left(t_d^{-1}\left(F_1(X_{1t})\right),t_d^{-1}\left(F_2(X_{2t})\right)\right)$	$\rho\,\epsilon(-1,1)$
Rotated-Gumbel		
	$C_{RG}(u,v;\delta)=F_1(X_{1t})+F_2(X_{2t})-1+\ e^{\left\{-\left[\left(-\ln\left(F_1(X_{1t})\right)\right)^{\delta}+\left(-\ln\left(F_2(X_{2t})\right)\right)^{\delta}\right]^{\frac{1}{\delta}}\right\}}$	$\delta\,\epsilon[1,\infty)$

Note. This table provides the various distributions for the copulas examined. For the normal copula, Φ^{-1} is the inverse of the cumulative distribution function (CDF) of a standard normal distribution, and the dependence parameter ρ is the Pearson's correlation coefficient, where the value 1 or -1 indicates complete dependence and 0 indicates complete independence. For the student's t copula, if the dependence parameter, ρ, takes the value 1 or -1 it indicates complete dependence, and 0 indicates complete independence. Both the left (lower) tail and right (upper) tail dependence measures take the form $2t_{d+1}\left(-\sqrt{\frac{(d+1)(1-\rho)}{1+\rho}}\right)$. For the rotated-gumbel copula, the dependence parameter, δ, takes the value of 1 for the case of independence and does not allow for negative dependence. The left (lower) tail dependence measure takes the form $2-2^{\frac{1}{\delta}}$.

commodity futures we consider in our analysis.[11] The individual commodity futures (along with their respective CRB symbol) are listed in Panel A of Table 2, along with the sub-sector to which it belongs. The inclusion of the specific commodity futures for this study is based on two criteria. First, each commodity future must have a continuous price series over the entire sample period considered. Second, the commodity futures must also have corresponding speculation data which can be extracted from the US Commodity Futures Trading Commission (CFTC)—described in section 3.3. The daily price series are averaged on a weekly (Tuesday-Tuesday) basis to obtain individual weekly price series. The same process is repeated for the equity index, the S&P 500, to obtain weekly equity price series; we extract daily equity price data from Bloomberg. Panel B of Table 2 lists the equity index as well as its respective Bloomberg identification symbol.

Table 2

Commodity Futures Groupings and Equity Index

Panel A: Commodity Sub-sectors

Commodity Futures	CRB Symbol
Energy	
Crude Oil, Brent	CB
Heating Oil #2	HO
Unleaded Gasoline	HU/RB
Natural Gas	NG
Foods & Fibers	
Cocoa	CC
Coffee	KC
Orange Juice	OJ
Sugar	SB
Cotton	CT
Lumber	LB

[11] In constructing the individual commodity futures price series we follow the typical methodology of rolling over the futures prices to the next-nearby contract when the nearest futures contract is one month from expiration (see Asness et al., 2013).

Table 2 (continued)

Commodity Futures Groupings and Equity Index

Grains & Oilseeds	
Corn	C_
Oats	O_
Soybeans	S_
Soybean Meal	SM
Soybean Oil	BO
Wheat	W_
Livestock	
Feeder Cattle	FC
Live Cattle	LC
Lean Hogs	LH
Precious Metals	
Gold	GC
Palladium	PA
Platinum	PL
Silver	SI

Panel B: Equity Index

Financial Index	BLM Symbol
S&P 500	SPX

Note. This table provides an overview of the commodity futures and equity index examined. Panel A displays the composition of the five commodity futures sub-sectors and their respective Commodity Research Bureau (CRB) symbols. Panel B displays the equity index and its respective Bloomberg (BLM) symbol.

We calculate the return series for all financial assets using the common log transformation on two consecutive weeks, formally, this sequence is given as:

$$X_{it} = \log P_{it} - \log P_{it-1} \tag{12}$$

where, X_{it} represents the log return series for each individual commodity future, or equity index, based on the price series, P_{it}. The weekly return series for each of the five commodity futures sub-sectors (energy, foods and fibers, grains and oilseeds, livestock, and precious metals) are calculated by taking an equally-weighted average of all futures returns, X_{it}, which comprise that particular sub-sector. For instance, the energy sub-sector is composed of an equally-weighted index of returns from Brent crude oil, heating oil #2, unleaded gasoline, and natural gas. Table 3 provides the summary statistics for the weekly rates of return. Specifically, Panel A summarizes the statistics for the weekly returns of the value-weighted equity index, while Panel B encapsulates the weekly return statistics for the five commodity futures sub-sectors. The excess skewness and kurtosis that the equity and commodity futures returns exhibit confirm the non-normality assumption; hence, reaffirming the need to use alternative measures of correlation structure to those based on simple linear assumptions. In general, we see that the returns of the energy and precious metals sub-sectors seem to most closely mimic those of the equity indices in terms of average return, skewness, and kurtosis. However, the standard deviation of returns for the equity index is markedly lower than that of all commodity sub-sectors, except livestock (0.007169). Interestingly, livestock is the only sub-sector (or composite index if we include equity as well) which exhibits positive skewness (0.030680) over the sample period. Overall, one can observe that the return properties of the five sub-sectors appear to decidedly differ, giving rise to the notion that their determinants may likely be heterogeneous.

3.2.1 DCC Estimation

The DCC model is in effect a two-step process to estimate the time-varying correlations between two different financial series. First, an ARMA model is fit to the specified return series and used to estimate the time-varying GARCH parameters. Second, the parameters driving the correlation dynamics are estimated using the standardized residuals from the first step estimation. This estimation procedure bears some similarities to that of the copula models.

Table 3

Summary Statistics for Weekly Rates of Return for the Equity Index and Commodity Futures Sub-sectors (October 1992 - October 2013)

Panel A: Equity Index

	S&P 500 Index
Mean	0.000567
Median	0.001232
Maximum	0.034318
Minimum	-0.053787
Std. Dev.	0.008138
Skewness	-0.665001
Kurtosis	6.772715
Obs.	1,111

Panel B: Commodity Futures Sub-sectors

	Energy	Foods & Fibers	Grains & Oilseeds	Livestock	Precious Metals
Mean	0.000619	0.000328	0.000317	0.000232	0.000622
Median	0.001756	0.000181	-0.000033	0.000168	0.000965
Maximum	0.097980	0.043599	0.036699	0.029603	0.040602
Minimum	-0.067262	-0.061074	-0.058338	-0.034269	-0.054108
Std. Dev.	0.014677	0.012034	0.010841	0.007169	0.010641
Skewness	-0.225218	-0.022605	-0.175441	0.030680	-0.898585
Kurtosis	5.537885	4.354079	4.525559	4.710997	6.409541
Obs.	1,111	1,111	1,111	1,111	1,111

Note. This table provides the summary statistics for the weekly rates of return for both the equity index and commodity futures sub-sectors over the period October 1992 to October 2013. Panel A displays the summary statistics for the unlevered rates of return for the S&P 500 index. All equity data is retrieved from Bloomberg. Equity index returns are calculated by taking the average value of daily index returns each week (Tuesday-Tuesday) and then taking the log difference on two consecutive weeks. Panel B displays the summary statistics for the rates of return for the various commodity futures sub-sectors. All commodity futures data is taken from the Commodity Research Bureau (CRB). Returns are calculated by taking the average value of daily individual commodity futures returns each week (Tuesday-Tuesday) and then taking the log difference on two consecutive weeks; the commodity futures sub-sector returns are then calculated by taking an equally-weighted average of all weekly futures returns which comprise that particular sub-sector. One month prior to the expiration of each individual commodity futures contract we roll the futures price series over to the next-nearby futures contract.

Since high frequency asset returns have a tendency to display fat-tails along with conditional heteroskedasticity and autoregressive characteristics, we select a mean equation for each return series via an AR(k) model based on the Bayesian Information Criterion (BIC)—which in our estimation provides the most parsimonious model. We then implement the Glosten-Jagannathan-Runkle-GARCH model, or more aptly the GJR-GARCH(p, q) model, which includes a leverage term for modeling asymmetric volatility clustering. In the GJR design, large negative changes are much more likely to be clustered with positive changes.[12] We apply $p = q = 1$ to our data sample given that this option usually best fits financial time series information. Thus, the model for each sub-sector and equity index log return series, X_t, is described via the following set of equations:

$$\left.\begin{aligned} X_t &= \mu + \sum_{n=1}^{k} \theta_j X_{t-j} + \varepsilon_t \\ \varepsilon_t &= \sigma_t \xi_t, \quad \text{where } \xi_t \sim \text{i.i.d.} \\ \sigma_t^2 &= \omega + \sum_{i=1}^{p} \beta \sigma_{t-1}^2 + \sum_{j=1}^{q} \alpha \varepsilon_{t-1}^2 + \sum_{j=1}^{q} \xi\, I[\varepsilon_{t-1} < 0] \varepsilon_{t-1}^2 \end{aligned}\right\} \tag{13}$$

where, the indicator function $I[\varepsilon_{t-1} < 0]$ equals 1 if $\varepsilon_{t-1} < 0$, and 0 otherwise. Hence, the leverage coefficient is applied to the negative innovations giving them more weight.[13] Table 4 provides the parameter estimates of the AR GJR-GARCH models for the S&P 500 equity index and the five commodity futures sub-sectors. In all cases, the BIC criteria chooses an AR(1)

[12] There are close similarities between the threshold GARCH (or T-GARCH) model and the GJR-GARCH model—the T-GARCH model is a recursive equation for the standard deviation process, while the GJR-GARCH model is a recursive equation for the variance process. If the leverage coefficient is zero, then the GJR-GARCH model reduces to a GARCH model.

[13] See Glosten et al. (1993) for further details on stationarity and positivity constraints.

specification for each financial time series. Moreover, we find significant leverage coefficients for the S&P 500 index, and the energy, livestock, and precious metals commodity sub-sectors.

Table 4
AR GJR-GARCH Model Parameters

	Parameter	t-statistic
Panel A: S&P 500 Index		
ARMA(1,0)		
Constant	0.000615	3.13
AR(1)	0.142240	4.43
GJR-GARCH(1,1)		
Constant	0.000002	2.41
GARCH(1)	0.824478	33.06
ARCH(1)	0.017029	0.82
LEVERAGE(1)	0.230510	7.30
Panel B: Energy		
ARMA(1,0)		
Constant	0.000228	0.56
AR(1)	0.218450	7.03
GJR-GARCH(1,1)		
Constant	0.000005	1.99
GARCH(1)	0.907876	46.31
ARCH(1)	0.041361	2.14
LEVERAGE(1)	0.045607	2.12
Panel C: Foods & Fibers		
ARMA(1,0)		
Constant	0.000285	0.84
AR(1)	0.211172	6.97
GJR-GARCH(1,1)		
Constant	0.000002	1.60
GARCH(1)	0.933194	83.55
ARCH(1)	0.065129	5.19
LEVERAGE(1)	-0.021456	-1.60
Panel D: Grains & Oilseeds		
ARMA(1,0)		
Constant	0.000414	1.46
AR(1)	0.245357	7.91
GJR-GARCH(1,1)		
Constant	0.000004	2.57
GARCH(1)	0.834690	32.75
ARCH(1)	0.145544	5.07
LEVERAGE(1)	-0.022028	-0.66

Table 4 (continued)
AR GJR-GARCH Model Parameters

Panel E: Livestock		
ARMA(1,0)		
Constant	0.000122	0.61
AR(1)	0.173848	5.73
GJR-GARCH(1,1)		
Constant	0.000001	1.80
GARCH(1)	0.912555	69.31
ARCH(1)	0.017020	1.37
LEVERAGE(1)	0.086612	4.53
Panel F: Precious Metals		
ARMA(1,0)		
Constant	0.000486	1.85
AR(1)	0.204885	6.85
GJR-GARCH(1,1)		
Constant	0.000002	1.74
GARCH(1)	0.870060	44.23
ARCH(1)	0.169278	6.40
LEVERAGE(1)	-0.096992	-3.97

Note. This table provides the autoregressive (AR) Glosten-Jagannathan-Runkle GARCH (GJR-GARCH) parameter estimates for modeling asymmetric volatility of the equity and commodity financial return series over the period October 1992 to October 2013.

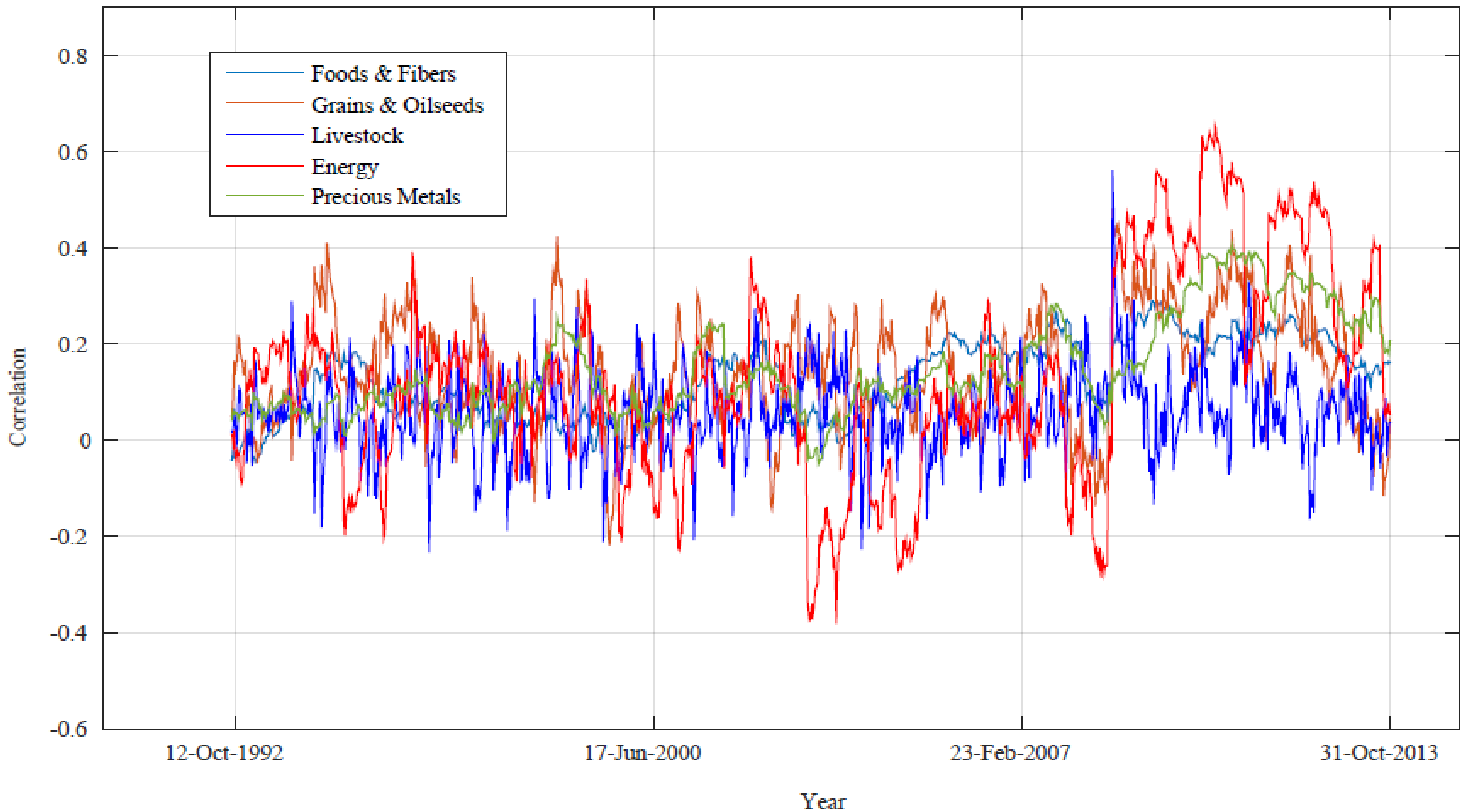
The residual series from (13) are standardized and used to estimate the time-varying correlation matrix between each of our five commodity futures sub-sectors and the equity indices, respectively, via maximum likelihood. Panel A of Table 5 presents the summary statistics for the DCC correlations between the S&P 500 and each of the commodity futures sub-sectors. The highest mean correlations occur for grains and oilseeds (0.149219) and precious metals (0.148714), while the greatest variation belongs to the energy sub-sector which has a standard deviation of 0.207809, far surpassing that of any other sub-sector. Most importantly, all sub-sector correlations, which are bounded between above (+1) and below (-1), are stationary according to the Augmented Dickey-Fuller (ADF) test results presented at the bottom of the panel, thus permitting our use of the dependence measure as a reliable dependent variable in our regression analysis in Section 4.

The most interesting aspect of the DCC model comes from an inspection of Figure 2 which plots the time-varying correlations between the S&P 500 and the five sub-sectors. Two things become readily apparent from the figure. First, the correlations between the equity markets and the commodity sub-sectors differ tremendously over the sample period, for example, the foods and fibers sub-sector shows much less variation in its correlation with the equity markets than does the energy sub-sector. Second, starting around mid-2003 the correlations between the commodity sub-sectors and the equity market seems to experience a slight upward trend, which becomes readily apparent in the post-2007 period. This rise in correlations corresponds to Buyuksahin et al. (2010) who note that this particular period (i.e. post-May 2003) is characterized by increasing participation of financial traders in the commodity futures market.

3.2.2 Copula Estimation

The initial steps of the copula estimation procedure are similar to those of the DCC described previously, whereby we utilize equation set (13) which contains the mean equation to accommodate for autocorrelation of each return series, via an AR(k) model, and the GJR-GARCH(p, q) model, using $p = q = 1$, to accommodate for heteroskedasticity. The parameters of this estimation procedure were described and reported in Table 4. The residuals from each series of this procedure are then standardized and used to estimate the empirical CDF of the filtered return series. Following the work of Patton (2006), we use these values to estimate, via maximum likelihood, the parameters of the normal, student's t, and rotated-gumbel copulas, outlined in Table 1. Panels B, C, and D of Table 5 present the summary statistics for the normal,

Figure 2
Dynamic Conditional Correlation between S&P 500 and Commodity Sub-sectors



student's t, and rotated-gumbel copula correlations between the S&P 500 and each of the commodity futures sub-sectors, respectively. The correlations for the normal and student's t copulas closely resemble those of the DCC model in terms of mean, maximum, and minimum values. However, there are a few distributional changes regarding the correlation measures; in particular, the normal copula tends to exhibit greater kurtosis over that of the DCC specification. In addition, the skewness measures for both copulas differ (both positively and negatively) from the DCC case. These distributional differences in univariate statistics underlie the fundamental differences in how copulas disentangle the unique characteristics of each return series from the dependence structure which links them together in order to estimate its dependence series. As in the DCC model, the equity-commodity correlations are bounded above (+1) and below (-1) for the normal and student's t copulas; moreover, rejection of the null hypothesis of the ADF tests permits the use of the correlation measures as dependent variables in our regression analysis due to their stationarity. Alternatively, in Panel D of Table 5, the rotated-gumbel offers another view of the dynamic equity-commodity correlations. The measure itself captures the lower left tail dependence and is unbounded above (∞), but bounded below (+1). The mean correlation measure across all sub-sectors ranges from 1.03 to 1.13 and have, in general, skewness and kurtosis distributional measures which are much greater than those in the (baseline) DCC case or the other two copulas.[14] Since this particular copula measures the comovement of the different financial assets in an extreme sense, these distributional differences are not too surprising. Further examination of the properties of the rotated-gumbel correlations also reveal that they are stationary, per the ADF results, and also permissible as a dependent variable in a regression setting.

[14] Even though the rotated-gumbel copula is unbounded above, we perform Monte Carlo simulations of the copula and find that, in the bivariate case, financial asset return series would have to comove very, very strongly to achieve a dependence parameter greater than two.

Figure 3A
Normal Copula Correlations between S&P 500 and Commodity Sub-sectors

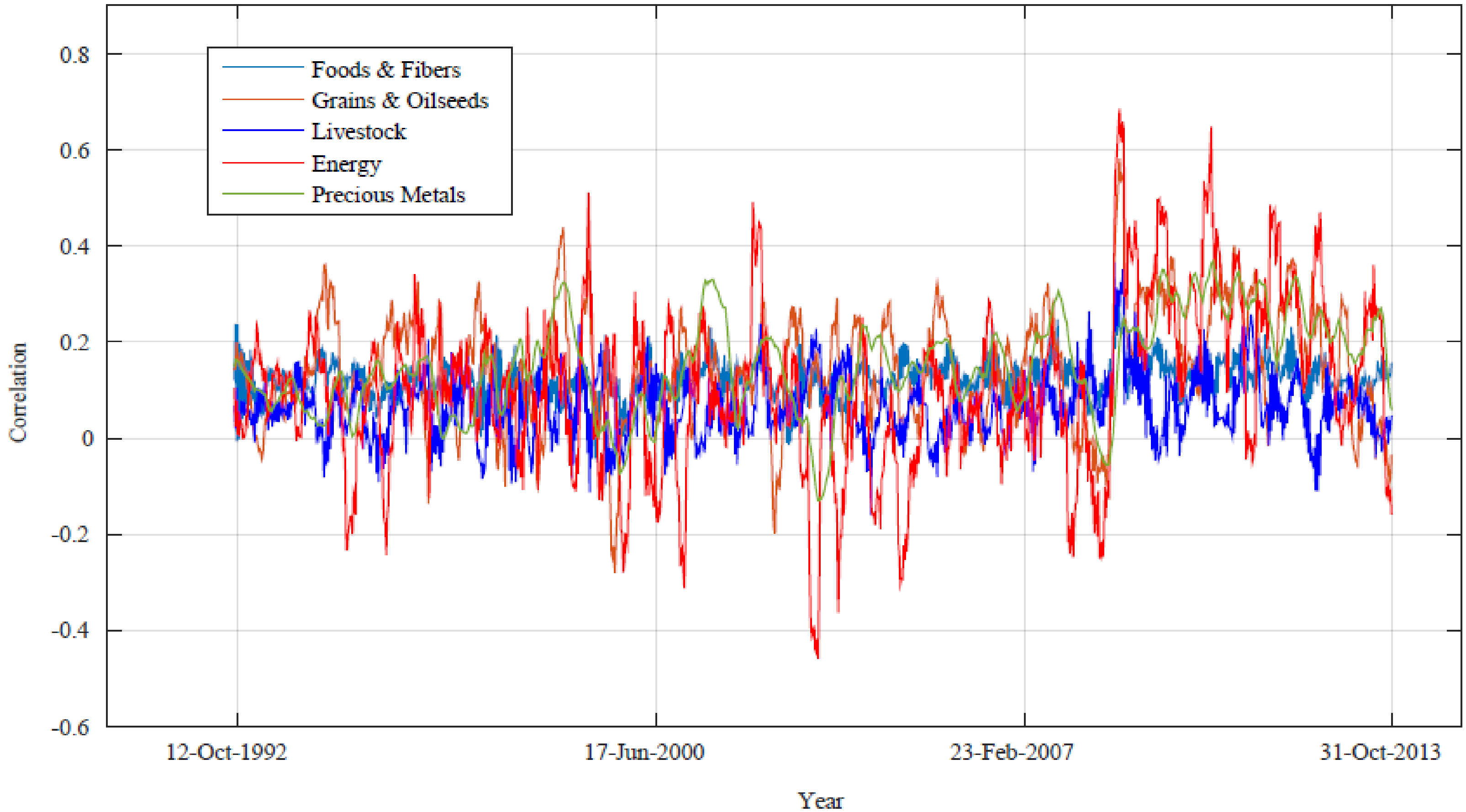


Figure 3B

Student's t Copula Correlation between S&P 500 and Commodity Sub-sectors

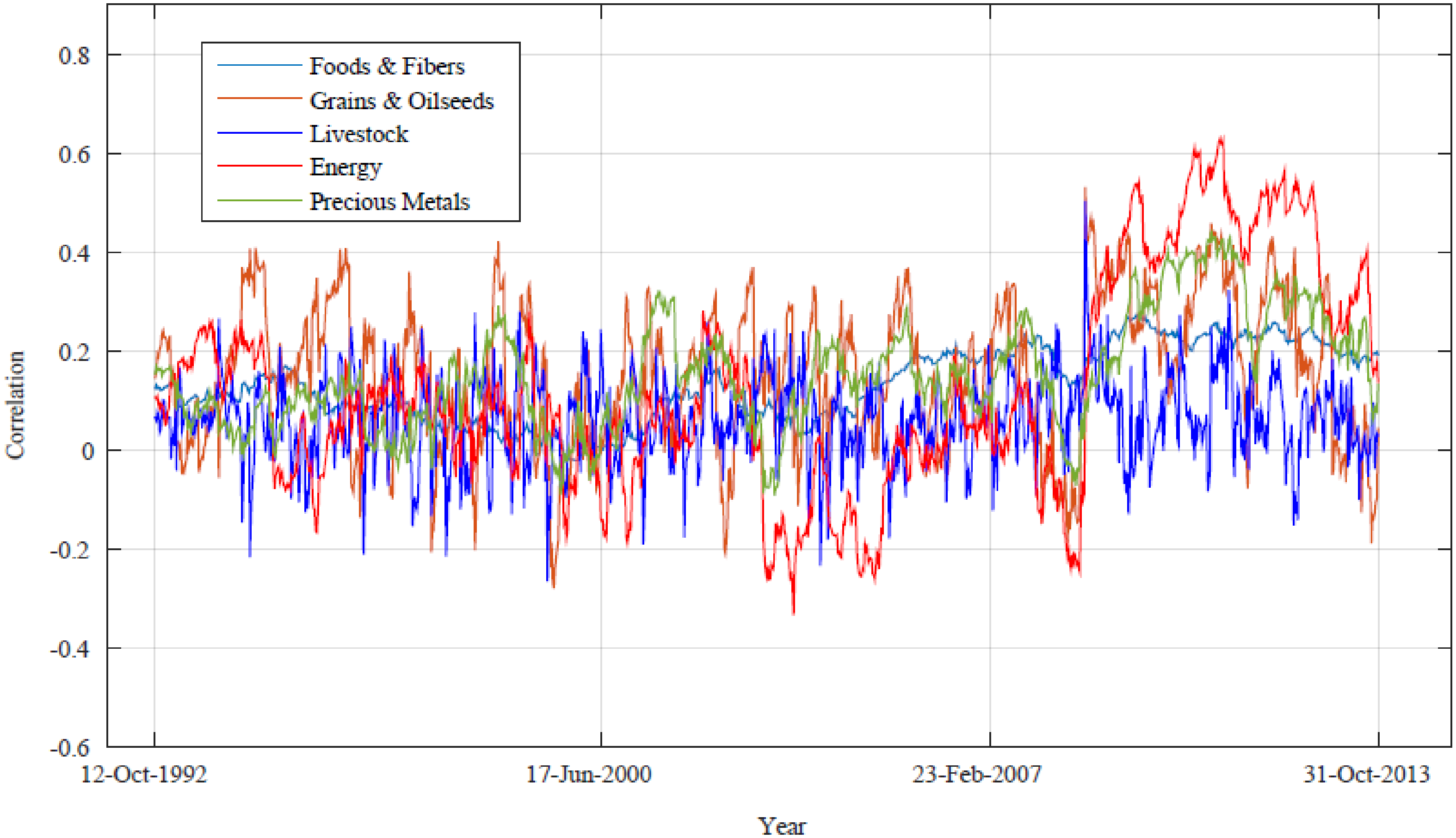


Figure 3C

Rotated-Gumbel Copula Correlation between S&P 500 and Commodity Sub-sectors

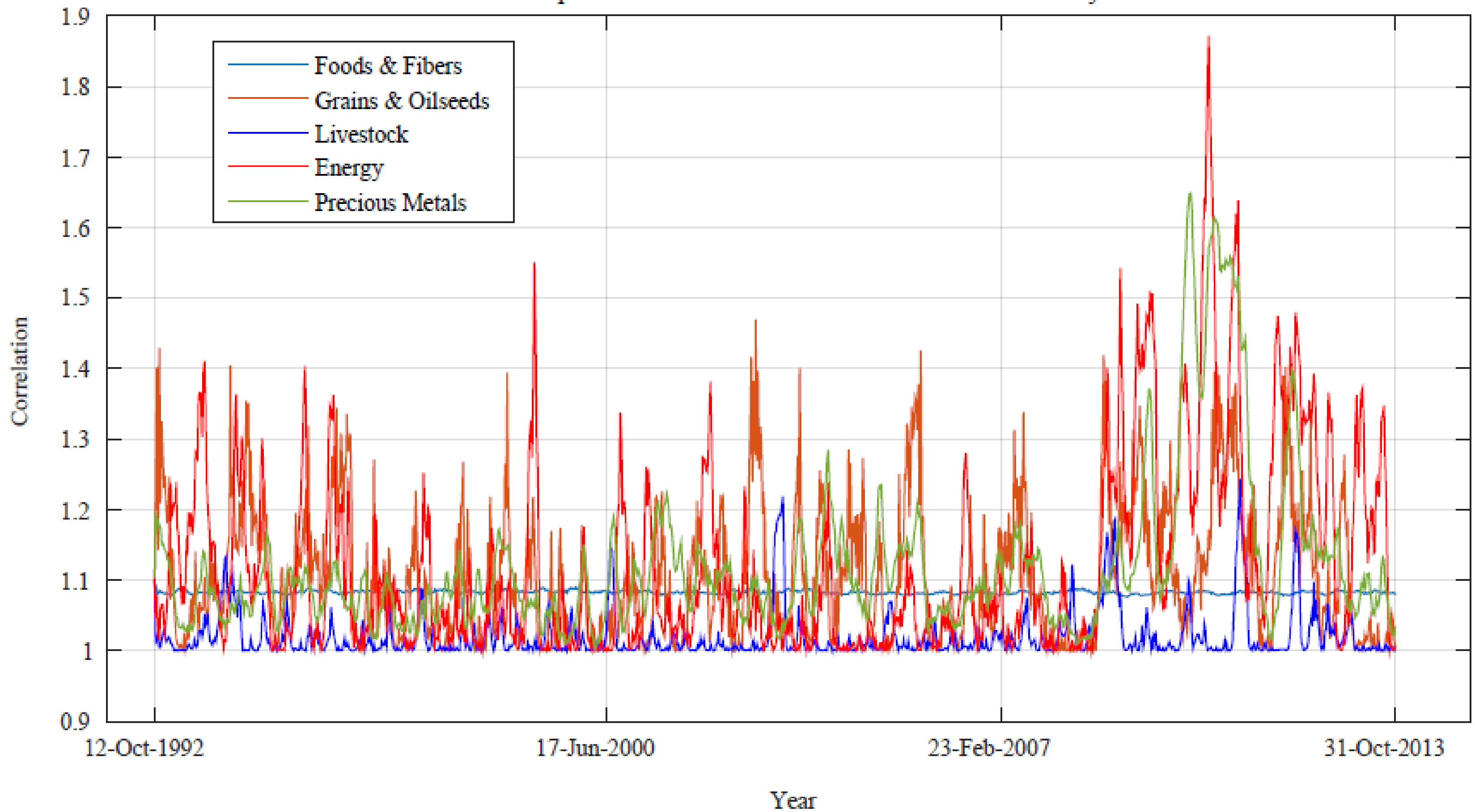


Table 5

Summary Statistics for Time-Varying Correlation Measures between S&P 500 and Commodity Futures Sub-sectors (October 1992 - October 2013)

Panel A: Dynamic Conditional Correlations

	Energy	Foods & Fibers	Grains & Oilseeds	Livestock	Precious Metals
Mean	0.126069	0.125489	0.149219	0.058425	0.148714
Median	0.107409	0.130840	0.140200	0.055995	0.115488
Maximum	0.654980	0.289116	0.482622	0.562071	0.407364
Minimum	-0.381763	-0.044773	-0.219353	-0.233247	-0.048541
Std. Dev.	0.207809	0.080540	0.117315	0.085951	0.100704
Skewness	0.308420	-0.013501	0.004694	0.237842	0.800524
Kurtosis	2.796507	1.814036	2.698508	4.874993	2.760354
Obs.	1,111	1,111	1,111	1,111	1,111
ADF level	-3.6603	-2.8313	-6.7141	-13.8197	-3.2697
ADF first diff.	-32.5152	-34.0796	-35.9554	-15.5627	-35.5540

Panel B: Normal Copula

	Energy	Foods & Fibers	Grains & Oilseeds	Livestock	Precious Metals
Mean	0.106621	0.121179	0.143618	0.064962	0.151442
Median	0.103570	0.120225	0.132827	0.064368	0.148810
Maximum	0.685544	0.307746	0.581658	0.365862	0.369051
Minimum	-0.459251	-0.050378	-0.280172	-0.159836	-0.130011
Std. Dev.	0.171787	0.041576	0.117906	0.067869	0.098891
Skewness	0.074327	-0.016404	0.121557	0.441763	-0.094123
Kurtosis	3.701664	4.304427	3.388926	4.034291	2.647368
Obs.	1,111	1,111	1,111	1,111	1,111
ADF level	-3.8620	-4.5766	-4.2010	-5.0842	-4.1359
ADF first diff.	-12.2665	-11.6330	-12.7967	-12.3530	-6.5630

Table 5 (continued)

Summary Statistics for Time-Varying Correlation Measures between S&P 500 and Commodity Futures Sub-sectors (October 1992 - October 2013)

Panel C: Student's t-Copula

	Energy	Foods & Fibers	Grains & Oilseeds	Livestock	Precious Metals
Mean	0.121276	0.130730	0.153388	0.060871	0.153499
Median	0.082237	0.123567	0.154018	0.062701	0.140089
Maximum	0.630170	0.274209	0.531798	0.503927	0.444629
Minimum	-0.333668	-0.021026	-0.279869	-0.265786	-0.092299
Std. Dev.	0.209177	0.074892	0.145963	0.087834	0.113016
Skewness	0.513095	0.009141	-0.129689	-0.087889	0.394606
Kurtosis	2.621861	1.905078	2.461630	3.800448	2.660061
Obs.	1,111	1,111	1,111	1,111	1,111
ADF level	-2.5227	-2.5681	-7.0675	-14.5816	-3.3463
ADF first diff.	-31.9363	-33.7978	-35.7528	-15.9471	-35.1611

Panel D: Rotated-Gumbel Copula

	Energy	Foods & Fibers	Grains & Oilseeds	Livestock	Precious Metals
Mean	1.122521	1.083149	1.107318	1.020608	1.119811
Median	1.066057	1.083026	1.085596	1.006884	1.089794
Maximum	1.871284	1.100000	1.469600	1.244000	1.648219
Minimum	1.000100	1.069708	1.000100	1.000100	1.003236
Std. Dev.	0.139193	0.002383	0.095962	0.034708	0.112460
Skewness	1.610933	0.344062	1.013706	3.008329	2.655377
Kurtosis	5.765241	5.683194	3.518673	13.184990	10.494070
Obs.	1,111	1,111	1,111	1,111	1,111
ADF level	-4.7136	-4.7546	-6.2368	-9.4629	-3.3308
ADF first diff.	-13.7744	-12.2148	-13.7591	-26.4828	-7.8700

Note. This table provides the summary statistics for the time-varying correlation measures between the S&P 500 and the five commodity futures sub-sectors over the full sample period (October 1992 to October 2013). Panel A provides the dynamic conditional correlations (DCC), while panels B, C, and D provide the correlations from normal, student’s t, and rotated-gumbel copulas, respectively. ADF represents the Augment Dickey-Fuller test for a unit root.

Figures 3A, 3B, and 3C highlight the various correlation time paths between the S&P 500 and the five different commodity futures sub-sectors for the normal, student's t, and rotated-gumbel copulas, respectively. While all three copulas highlight the heterogeneity between the dynamic equity-commodity correlation measures, the student's t copula, and to a lesser extent the normal copula, visibly illustrate a rise in the correlations that is documented as beginning in mid-2003. Interestingly, over the latter part of the sample period, we also witness a spike in the lower tail dependence for several equity-commodity pairings, in particular, energy, precious metals, and grains and oilseeds.

3.3 Explanatory Variables

We employ a series of macroeconomic and financial market variables along with a measure which captures aggregate market speculation for each commodity sub-sector, in order to determine what factors determine the dynamic correlations for each equity-commodity sub-sector pairing. We follow the suggestions of prior literature in our choice and implementation of these variables.

3.3.1 Macroeconomic Fundamentals

It is well-known that business cycle factors have an impact on commodity returns (see Erb and Harvey, 2006; Gorton and Rouwenhorst, 2006). Given this observation, we use an aggregate measure of US macroeconomic conditions called the Aruoba-Diebold-Scotti Index (ADSI), which tracks real business conditions at a high frequency (see Aruoba et al., 2009; Buyuksahin and Robe, 2013). The ADSI variable is a composite of several underlying seasonally adjusted (high- and low-frequency) economic indicators which include: weekly initial jobless claims, monthly payroll employment, industrial production, personal income less transfer payments, manufacturing and trade sales, and quarterly real GDP. The index is normalized such that the

ADSI variable is zero. Progressively larger positive values indicate better-than-average business conditions, whereas progressively more negative values indicate worse-than-average business conditions. Using historical statistics dating back to 1960, Bhardwaj and Dunsby (2013) find that the equity-commodity correlation business cycle component increases during period of economic weakness, and that the link between the equity-commodity correlation and business cycle is stronger for industrial commodities than agricultural commodities. However, Buyuksahin and Robe (2013) find that business conditions impart a positive, though not consistently significant, impact on dynamic equity-commodity correlations.

While US macroeconomic conditions are of substantial importance to the prices of financial assets, worldwide economic activity also plays a central role, particularly for commodities. Thus, we implement a measure of real global economic activity called the Baltic Dry Shipping Index (BDSI). The BDSI is an indicator of transportation costs for raw materials shipped by sea. It's based on a daily quote, published by the Baltic Exchange in London, for booking vessels of various sizes and across multiple maritime routes. Specifically, the BDSI is calculated as a weighted-average of the Baltic Exchange's indices for the shipping costs of the four largest dry-vessel classes. Our interest in this measure is based on the idea that the supply structure of the shipping industry is generally predictable and that changes in shipping costs are largely due to changes in the worldwide demand for raw materials. Kilian (2009) uses a similar type of freight measure and finds that increases in the shipping rates of freight can be used as indicators of both demand and supply shifts in global commodity markets. This link to global demand has prompted some interest in the BDSI as a leading indicator of global economic activity. Withstanding recent work by Bakshi et al. (2011), who investigate the BDSI as a predictor for global stock and commodity returns, not many studies have used the variable for analysis beyond that of economic growth.

Panel A of Table 6 provides the summary statistics of these macroeconomic variables. Most importantly diagnostic tests reveal that the variables are both stationary in the level form, thus permitting them as usable variables in our regression. Additionally, it is apparent that the magnitude of the mean of the BDSI variable (2,356.17) is much greater than the mean of the ADSI (-0.1499), or any of our dependent variables, so we use the natural logarithm of the BDSI to remedy the issue in our regression analysis in Section 4.

3.3.2 Financial Market Indicators

Recent work by Silvennoinen and Thorp (2013) documents that for certain commodity futures higher than expected US stock volatility can help to predict higher volatility in those markets. Alternatively, for a small sample of other commodity futures they note the opposite effect. Overall, they conclude that an increase in stock market volatility, as proxied by the VIX index, can be linked to an increase in correlations across markets. Based on this finding we include the VIX index, a measure of implied volatility of S&P 500 index options, or better regarded as a gauge of investor sentiment (or "fear" index), as a regressor in our analysis. A general interpretation of the index is as follows, higher values of the VIX correspond to greater investor uncertainty about the equity markets.

While equity market volatility is well captured using the VIX index, broad market financial stress may not be so easily encapsulated. The finance literature has acknowledged that an increase in cross-market correlations in crisis periods occurs due to arguments such as spill-over effects and flight-to-quality (see Danielsson et al., 2011; Pavlova and Rigobon, 2008; Kyle and Xiong, 2001). Therefore, following the work of Hong and Yogo (2012), who investigate the

Table 6

Summary Statistics for Macroeconomic, Financial Market, and Speculation Variables (October 1992 - October 2013)

Panel A: Macroeconomic Variables

	Aruoba-Diebold-Scotti Index (ADSI)	Baltic Dry Shipping Index (BDSI)
Mean	-0.1499	2,356.17
Median	-0.0452	1,562.00
Maximum	1.8019	11,573.40
Minimum	-3.9308	653.60
Std. Dev.	0.7528	1,964.65
Skewness	-2.0591	2.2540
Kurtosis	9.8130	8.3934
Obs.	1,111	1,111
ADF level	-3.1758	-2.8736
ADF first diff.	-13.3703	-13.5897

Panel B: Financial Market Variables

	Market Volatility Index (VIX)	Yield Spread (YS)
Mean	20.3922	0.9684
Median	18.9220	0.8640
Maximum	72.7200	3.4480
Minimum	9.5775	0.5260
Std. Dev.	8.3245	0.4458
Skewness	1.9062	3.0442
Kurtosis	9.0562	14.6389
Obs.	1,111	1,111
ADF level	-4.2232	-3.3236
ADF first diff.	-31.8152	-12.3040

Table 6 (continued)

Summary Statistics for Macroeconomic, Financial Market, and Speculation Variables (October 1992 - October 2013)

Panel C: Sub-sector Excess Speculation Measures

	Energy (SPE_ENERGY)	Foods & Fibers (SPE_FOODFIB)	Grains & Oilseeds (SPE_GRAINS)	Livestock (SPE_LIVESTK)	Precious Metals (SPE_PMETALS)
Mean	0.333579	1.049820	1.052079	1.289389	0.489811
Median	0.226707	0.958957	0.778949	1.091181	0.450626
Maximum	0.949566	4.567558	13.753700	6.753103	1.450270
Minimum	0.000000	0.109335	0.234880	0.320550	0.071367
Std. Dev.	0.246275	0.477917	1.250135	0.729180	0.225018
Skewness	0.572599	1.947755	5.326990	2.175143	1.086559
Kurtosis	1.916844	10.410460	36.967110	10.532740	4.322239
Obs.	1,111	1,111	1,111	1,111	1,111
ADF level	-4.0186	-5.8459	-9.4453	-7.0662	-7.4378
ADF first diff.	-26.2772	-25.9532	-14.2246	-31.9454	-31.7010

Note. This table provides the summary statistics for the macroeconomic, financial, and speculation variables over the period October 1992 to October 2013. Panel A displays the summary statistics of the weekly macroeconomic variables, the Aruoba-Diebold-Scotti Index (ADSI) tracks real business conditions at a high frequency and the Baltic Dry Shipping Index (BDSI) provides an assessment of the price of moving major raw commodity materials by sea. Panel B displays the summary statistics for the weekly financial market variables, the market volatility index (VIX) represents the market's expectation of stock market volatility and the Yield Spread (YS) is the difference between Moody's Aaa and Baa corporate bond yields, which represents a reflection of the overall broad corporate economy (and therefore credit quality and financial stress). Panel C displays the summary statistics regarding the calculation of the excess speculation index for each commodity futures sub-sector. Excess speculation for each individual commodity futures series is calculated via Working's "T" method based on weekly (Tuesday-Tuesday) speculation data provided by the US Commodity Futures Trading Commission (CFTC), and then aggregated to its respective sub-sector. The variables SPE_ENERGY, SPE_FOODFIB, SPE_GRAINS, SPE_LIVESTK, and SPE_PMETALS represent speculation in the energy, foods & fibers, grains & oilseeds, livestock, and precious metals commodity futures sub-sectors, respectively. ADF represents the Augment Dickey-Fuller test for a unit root.

predictability of commodity futures as well as other asset returns, we proxy for aggregate financial market stress using a slight variation of the yield spread (YS). Here, YS is defined as the difference between Moody's Aaa corporate bond yield and Baa corporate bond yield.

Panel B of Table 6 provides the summary statistics of the financial market variables. As in the case of BDSI, we use the log of VIX to facilitate our regression analysis given that its mean value (20.3922) is substantially larger than the mean value of all other independent and dependent variables. Additionally, both financial indicators are stationary in their level permitting reasonable inferences from the regression in Section 4.

3.3.3 Excess Speculation

Recent literature on the financialization of commodity markets recognizes the idea that "who trades matters," and that the presence of increased market participation may in fact propagate the linkage between cross-market (price) correlation dynamics (see Etula, 2009; Tang and Xiong, 2012; Buyuksahin and Robe, 2013). We address this issue by acknowledging that speculators and index investors perform very different economic roles in the commodity futures market and that these differences should have dissimilar influences on commodity prices. A survey by Greely and Currie (2008) highlights that speculators bring information to the commodity futures markets on future supply and demand fundamentals, while index investors merely earn a passive return as payment for bearing the risk of price fluctuations. We postulate that the role of speculators may be unique among the different sub-sectors of commodity futures given the distinctiveness of the commodities themselves, as well as their individual trading volume.[15]

In order to create an index which accounts for market speculation we appeal to the Commitment of Traders (COT) reports which aggregate the positions of "major players" in the

[15] See volume statistics at www.futuresindustry.org.

US commodity futures markets each week. It is exclusively devoted to the domain of open interest with no price or volume data. Traders are divided into commercials traders, non-commercial traders, and small traders. Commercial traders (or hedgers) participate in order to hedge their inherent commodity price risk exposure, whereas non-commercial traders (or speculators) participate in order to profit from the anticipation of future price movements. We utilize the US Commodity Futures Trading Commission's (CFTCs) sub-classification of open interest data to measure speculation in the market. Prior studies analyzing the role of speculation have utilized Working's "T" index, defined as the ratio of positions held by speculators to that of hedgers (see Buyuksahin and Robe, 2013). Working's "T" measures the extent to which speculation is in "excess" of the level required to satisfy hedgers' net demand for hedging at the market clearing price. It is common to interpret a high index or high volatility of the index as indicative of excess speculation. For each of our 23 commodity futures $(\mathrm{i} = 1,2,\dots,23)$ we calculate Working's "T" on a weekly basis (Tuesday-Tuesday), as that is when COT publically publishes their trading data.

Formally, for the $\mathrm{i^{th}}$ commodity market in week t we calculate the speculation index as follows:

$$\mathrm{T_{it}} = \begin{cases} \dfrac{\mathrm{SS_i}}{\mathrm{HL_{it}} + \mathrm{HS_{it}}} \text{ if } \mathrm{HS_{it}} \geq \mathrm{HL_{it}} \\ \dfrac{\mathrm{SL_i}}{\mathrm{HL_{it}} + \mathrm{HS_{it}}} \text{ if } \mathrm{HS_{it}} < \mathrm{HL_{it}} \end{cases} \quad (\text{for i} = 1, \dots, 23) \tag{14}$$

where, $\mathrm{SS_i} \geq 0$ and represents the "Speculator Short" positions held in aggregate by all non-commercial traders, $\mathrm{SL_i} \geq 0$ and represents the "Speculator Long" positions held in aggregate by all non-commercial traders, $\mathrm{HS_{it}} \geq 0$ and represents all commercial "Hedge Short" positions, and $HL_{it} \geq 0$ and represents all commercial "Hedge Long" positions. After calculating excess speculation in each individual market, we aggregate the measure for each sub-sector as follows:

$$SI_{nt} = \sum_{i=1}^{k} T_{it,n} \quad (\text{for } n = 1, \dots, 5) \tag{15}$$

where, SI_{nt} is the "Speculation Index" for each of the five sub-sectors, n, composed of the individual commodity futures which belong to it.

Panel C of Table 6 summarizes the speculation measures for each of the five sub-sectors investigated in this paper. The univariate statistics differ quiet drastically between markets, with livestock reporting the highest mean measure of excess speculation (1.289389) and energy reporting the lowest (0.333579). Furthermore, tests of non-stationarity reveal that all speculation indices are, in fact, stationary in their level form and hence usable variables in our regression analysis.

4. REGRESSION RESULTS

Given the unique characteristics among the different types of commodities, the factors which determine the time-varying correlations among the various sub-sectors with other asset markets, in particular equities, should also be unique. In order to explore this hypothesis we utilize the following regression model:

$$y_{nt} = \alpha w_t + \beta x_t + \gamma z_{nt} + \varepsilon_t \tag{16}$$

where, y_{nt} is a $t \times 1$ vector of dependent variable dynamic correlations from either the DCC model or one of the copula specifications for a given sub-sector (n), x is a $t \times k$ vector of regressors consisting of macroeconomic and financial market variables, z_{nt} is $t \times 1$ vector of

Working's "T" excess speculation measure for a given sub-sector (n), and w is a $t \times 1$ vector of one's. We estimate the model parameters using ordinary least squares (OLS) and report Newey-West t-statistics which correct for both autocorrelation and heteroskedasticity.[16] The results from the regression analysis provide a more detailed level of understanding of the commodity futures market and its dependence with the equity markets. Furthermore, accurate and current knowledge on the determinants of the dynamic equity-commodity correlations at the sub-sector level has implications for non-index commodity futures investors and portfolio managers alike.

4.1 Equity-Commodity Correlation Determinants

In order to evaluate the evolution of the determinants of the dynamic correlations we decompose our overall sample (October 1992 to October 2013) into two sub-samples: sub-sample A (October 1992 to May 2003) and sub-sample B (May 2003 to October 2013). The justification for splitting the sample around mid-2003 stems from Buyuksahin et al. (2010) who note that the latter sub-period is characterized by increasing participation of financial traders in the commodity futures market. This observation and the emerging literature which argues that the increased financialization of commodities has contributed to the increase in dynamic correlations with more traditional assets makes the break in the overall sample period a natural choice.[17]

4.1.1 The DCC Model

Table 7 presents our baseline approach using the DCC as the dependent variable. Panel A presents the results for the energy sub-sector. For the full sample period, we find that the variables ADSI, YS, and SPE_ENERGY are all positive and highly statistically significant.

---

[16] We also estimate model parameters using time series regression techniques and find similar results.

[17] The literature on commodities lacks a complete consensus on when investing in commodity futures became highly popularized; however, the general consensus is that it occurred in the early 2000's. Given this, we analyze individual trading volumes of commodity futures in conjunction with Buyuksahin et al. (2010) and find mid-2003 to be a suitable place to divide the sample.

Hence, a 1% increase in ADSI, YS, and SPE_ENERGY, ceteris paribus, results in a 0.0786%, 0.1396%, and 0.4285% increase in the dynamic equity-commodity correlations, respectively. These findings are largely consistent with those of Buyuksahin and Robe (2013), which is not too surprising since they examine the DCC between the SP-GSCI (which is heavily weighted in energy futures) and the S&P 500. Conversely, we find that coefficient on BDSI is negative, but highly statistically significant. Its economic influence seems to be substantially less than the other factors however, as only a 1% decrease in BDSI results in a 0.003396% increase in the equity-commodity correlations. Since aggregate worldwide demand is approximated using BDSI, this result confirms the intuition that cross-market correlations increase in poor global economic conditions. Interestingly, when we decompose these results into the two sub-periods, we find that in the latter sub-period, B, all of the explanatory variables become highly significant and are of the expected sign, whereas in the former sub-period, A, just three of the variables (BDSI, VIX, and SPE_ENERGY) are statistically significant, but VIX and SPE_ENERGY are not of the expected sign. Thus, in the period characterized by increased market participation, we find that the equity-commodity correlations for the energy sub-sector are strongly determined by our macroeconomic, financial, and speculation indicators. Prior to this period, however, the determinants of the dynamic correlations are less linked to overall macroeconomic and global market conditions. Moreover, two of the variables which are significant, VIX and SPE_ENERGY, take the opposite (negative) sign indicating that prior to 2003 the return properties between the two asset classes drastically differed.

Panel B presents the results for the foods and fibers sub-sector. Over the full sample period, all of the explanatory variables are significant at conventional levels. Interestingly, the coefficient on VIX is significantly negative; however, given the findings of Silvennoinen and

Thorp (2013) we would expect the sign to be positive. Yet, the overall economic significance of this change seems rather small, as a 1% increase in the VIX results in a decrease of 0.00716% in the equity-commodity correlations. The coefficient on BDSI tells a similar story. Interestingly, the results across the two sub-periods are strikingly similar to what we observed in Panel A. All the coefficients on the variables in the latter sub-period take the expected sign and are statistically significant, whereas in the first sub-sample period just three of the variables are significant, and the sign of VIX is the opposite of expectations. Disparate from Panel A, we see the economic impact of all variables is markedly smaller in Panel B. For instance, the speculation measure though significant in the latter sub-period, its economic impact is decidedly smaller; a 1% increase in speculation leads to a 0.0314% increase in the equity-commodity correlations. Nonetheless, this result lends some credence to the argument that increased market participation, by market speculators, is in fact a prime contributor to the increased comovement between commodity futures and equities, hence deteriorating the long-run benefits of commodity futures.

Panel C, which shows the grains and oilseeds sub-sector, displays some interesting results regarding the evolution of the determinants of the equity-commodity correlations. In the full sample period, the only coefficients which are not rendered insignificant are ADSI and YS. However, an examination of sub-period A shows that none of the factors surveyed help to explain the dynamic correlations. Yet, sub-period B reveals that all of the factors are now highly significant at conventional levels. Overall, we see a drastic shift in the determinants of the dynamic correlations of the sub-sector in the period characterized by an increase in speculative activity. This characteristic is particularly interesting given that the latter sub-period results are similar to those found in Panel B, except that the sign of the coefficient on the speculative

Table 7

Determinants of Dynamic Conditional Correlations between the S&P 500 and Commodity Futures Sub-sectors

	Oct. 1992 - Oct. 2013 (Full Sample Period)	Oct. 1992 - May 2003 (Sub-period A)	June 2003 - Oct. 2013 (Sub-period B)
Panel A: Energy			
Constant	0.8424	2.2233	0.0816
	(4.01)	(5.39)	(0.34)
ADSI	0.0786	-0.0047	0.1704
	(3.26)	(-0.26)	(6.23)
Log(BDSI)	-0.3396	-0.5712	-0.2938
	(-7.22)	(-5.09)	(-5.06)
Log(VIX)	0.0999	-0.2769	0.5297
	(1.31)	(-3.74)	(5.05)
YS	0.1396	0.0675	0.1815
	(3.68)	(1.14)	(4.05)
SPE_ENERGY	0.4285	-0.4720	0.5055
	(6.56)	(-2.35)	(4.07)
R^2	0.4427	0.2717	0.6055
Panel B: Foods & Fibers			
Constant	-0.1047	0.2528	0.0901
	(-1.12)	(1.41)	(1.20)
ADSI	0.0294	-0.0140	0.0335
	(2.33)	(-1.12)	(3.93)
Log(BDSI)	0.0523	-0.0599	-0.0322
	(2.57)	(-1.22)	(-1.82)
Log(VIX)	-0.0716	-0.0684	0.0964
	(-1.90)	(-1.70)	(2.86)
YS	0.1176	0.0808	0.0439
	(6.53)	(2.95)	(2.42)
SPE_FOODFIB	0.0386	0.0358	0.0314
	(3.87)	(4.31)	(2.67)
R^2	0.3117	0.1953	0.1974

Table 7 (continued)

Determinants of Dynamic Conditional Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel C: Grains & Oilseeds			
Constant	0.1279	-0.1703	0.2728
	(0.92)	(-0.65)	(1.60)
ADSI	0.0361	0.0072	0.0717
	(2.54)	(0.43)	(3.71)
Log(BDSI)	-0.0168	0.1103	-0.0983
	(-0.53)	(1.48)	(-2.43)
Log(VIX)	-0.0379	-0.0655	0.1579
	(-0.60)	(-0.76)	(1.77)
YS	0.1324	0.0447	0.1295
	(6.08)	(1.01)	(3.37)
SPE_GRAINS	0.0017	0.0030	-0.1060
	(0.35)	(0.72)	(-3.12)
R^2	0.1160	0.0197	0.3094
Panel D: Livestock			
Constant	-0.0478	-0.2534	-0.0286
	(-0.65)	(-1.05)	(-0.31)
ADSI	0.0031	0.0149	-0.0105
	(0.41)	(1.32)	(-0.96)
Log(BDSI)	0.0034	0.0736	-0.0133
	(0.22)	(1.11)	(-0.61)
Log(VIX)	0.0329	0.0190	0.0912
	(1.01)	(0.38)	(1.74)
YS	0.0487	0.0624	0.0151
	(2.92)	(1.90)	(0.61)
SPE_LIVESTK	0.0048	0.0005	0.0027
	(0.95)	(0.07)	(0.35)
R^2	0.0637	0.0165	0.1092

Table 7 (continued)

Determinants of Dynamic Conditional Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel E: Precious Metals			
Constant	-0.0023	0.8589	0.3982
	(-0.02)	(5.09)	(3.61)
ADSI	0.0293	-0.0384	0.0297
	(1.98)	(-3.69)	(2.32)
Log(BDSI)	0.0054	-0.2562	-0.1702
	(0.20)	(-5.23)	(-7.07)
Log(VIX)	0.0394	0.0561	0.3972
	(0.78)	(1.64)	(6.78)
YS	0.0776	-0.0377	-0.0893
	(2.98)	(-1.86)	(-3.22)
SPE_PMETALS	0.0245	-0.0107	-0.0067
	(0.78)	(-0.73)	(-0.18)
R^2	0.0672	0.3885	0.3938

Note. This table provides the ordinary least squares (OLS) regression results for each of the commodity futures sub-sectors over the full sample period (October 1992 to October 2013) and two sub-periods (October 1992 to May 2003 and May 2003 to October 2013). In panels A, B, C, D, and E the dependent variable is the time-varying dynamic conditional correlation (DCC) between the weekly rates of return on the S&P 500 equity index and the equally-weighted weekly futures returns on the energy, foods & fibers, grains & oilseeds, livestock, and precious metals sub-sectors, respectively. The variables ADSI, BDSI, VIX, and YS represent the Aruoba-Diebold-Scotti Index, the Baltic Dry Shipping Index, the market volatility index, and yield spread, respectively. The variables SPE_ENERGY, SPE_FOODFIB, SPE_GRAINS, SPE_LIVESTK, and SPE_PMETALS represent speculation in the energy, foods & fibers, grains & oilseeds, livestock, and precious metals commodity futures sub-sectors, respectively. In all sample periods, Newey-West t-statistics are reported in parentheses below the corresponding coefficients, along with the R^2 of the regression.

variable (SPE_GRAINS) is negative. According to prior work using commodity indices the variable's sign should be positive (as increased speculator participation has been shown to increase correlations). As such, our results curiously indicate that for the grains and oilseeds market a 1% increase in speculative activity results in a 0.1060% decrease in dynamic correlations between the equity and commodity futures market. This result highlights the point that outside of a commodity index setting the factors which affect the dynamic correlations between commodity futures and traditional assets are not homogenous across all futures markets.

Panel D highlights the regression results for the livestock sub-sector. A quick inspection of the results reveals that over all sample periods only a few of the explanatory variables generally aid in explaining the dynamic equity-commodity correlations. Over the full sample period, yield spread (YS) is highly significant and positive, which is observed in all the other commodity sub-sectors examined up to this point as well. Similar to Panel C, the regression model does a poor job of explaining the equity-commodity correlations in sub-period A. However, unlike Panel C, the explanatory variables in sub-period B also seem to do a rather inadequate job of explaining the dynamic correlations as well, as only VIX is significant at standard confidence levels. Furthermore, the regression in Panel D also shows an R-squared for sub-period B that is approximately 11% (comparably, sub-period A only has an R-squared of 1.7%). Thus, the livestock sub-sector correlations with the equity market are far removed from the broad market explanatory variables we saw in the other groups. This result provides further credence for the heterogeneity among the different sub-sectors and that the differences in financial asset return determinants translate into real opportunities for price formation and risk management strategies.

Finally, the results panel E, which contains the precious metals sub-sector, tells a strikingly similar story of a drastic shift in the sub-sector determinants as seen in Panels A, B, and C. Sub-period A presents the ADSI, BDSI, and YS as significantly negative determinants of the dynamic correlations. However, the shift to sub-period B shows not only increased significance of the VIX variable, but a change in the sign of ADSI. Unlike prior analysis of the sub-sectors the YS is unexpectedly negative, and the speculation variable is statistically insignificant altogether. These findings once again highlight the heterogeneous effects of not only speculation, but also of all the determinants considered across the various sub-sectors.

Implementing the DCC model as the dependent variable reveals some very interesting results across the different commodity futures sub-sectors. In general, we see that in moving from sub-period A to sub-period B the dynamic equity-commodity return correlations are increasingly explainable by the series of macroeconomic and financial market indicators. Furthermore, the effect of increased investor speculation in the market is heterogeneous across the various commodity futures sub-sectors. For both the energy and foods and fibers sub-sectors the speculation variable is positive and significant for their respective equity-commodity return correlations, in both the full and sub-sample B periods. However, the economic magnitude of the coefficients for the foods and fibers group is noticeably less than that of the energy sector. Contrastingly, for the grains and oilseeds sub-sector we find the speculation variable is insignificant for the full sample period but negative and significant sign in sub-period B. This result suggests that increased speculation actually decreases dynamic equity-commodity correlations within the group. Finally, in the livestock and precious metals sub-sectors the speculation variable is insignificant in all regression periods. One additional thought-provoking point is that for each sub-sectors' full sample period regressions the proxy for financial stress

(YS) is positive and highly significant. Overall, a 1% increase in the yield spread results in roughly a 0.05%-0.14% increase in dynamic equity-commodity return correlations.

4.1.2 The Normal Copula

Table 8 presents our regression results using the normal copula as the dependent variable. The normal copula is a symmetrical dependence structure which allows for no tail dependence and provides in many senses a more robust measure of the return comovement. Panel A summarizes our findings for the energy sub-sector. In general, we find a similar pattern between the normal copula and the DCC results for the sub-sector. The significance of the results do not vary much, although the magnitude of the relevant coefficients seems to decrease for both the full sample and sub-sample B in comparison to the results in Panel A of Table 6. Overall results again illuminate the unique fact that the returns of the energy sub-sector and equity market both seem to be highly intertwined and determined via broad market macroeconomic and financial market indicators as well as speculative activity since May 2003.

Interestingly, the regression results of Panel B differ quite a bit from those of the (baseline) DCC model. In the normal copula case, only YS is significant (and positive) for the full sample period. However, the most prominent changes come in evaluating the results in sub-period A where only the VIX is significant. In the latter period, the variables ADSI, VIX, and YS are all highly significantly positive for the normal copula, and the magnitudes of the coefficients are economically similar to those in the DCC model. Moreover, as in Panel A, sub-period B elicits a higher R-squared over the earlier period. The normal copula based findings paint a markedly similar picture of increased integration among the equity-commodity return determinants in the last 10 years.

The results of Panels C and D, which summarize the grains and oilseeds and livestock sub-sectors, respectively, do not dramatically differ from those found in Table 6, both the magnitude and sign of the coefficients are relatively unchanged. The only major difference is that the ADSI coefficient becomes significantly positive in sub-period A of the livestock sub-sector. Regarding grains and oilseeds, the speculation variable, interestingly, remains significantly negative, indicating that a 1% increase in speculative activity results in a 0.0982% decrease in the equity-commodity return correlations. Results based on the dynamic copula correlations for the two sub-sectors also highlight the heterogeneous effects of the broad market indicators. Furthermore, they also detail a story of increased integration among the deterministic return factors between the equity and commodity futures markets; though the copula based results tend to be less acute.

The results for Panel E, the precious metals sub-sector, are somewhat different from the DCC (baseline) case. The entire set of coefficients are insignificant for the full sample period. However, in sub-sample A we find the variables ADSI, BDSI, and YS all become statistically significant, just as in the DCC model. We again see that ADSI and YS are of the opposite than expected sign. Yet, in sub-sample B, ADSI changes its sign to positive, YS becomes statistically insignificant, and VIX becomes positive and significant. The fact that ADSI and YS are of the negative sign in sub-period A, but of the positive sign in sub-period B, highlights an evolution of the determinants in the particular sub-sector. Interestingly, the speculation variable is insignificant in all cases, similar to the findings in Panel D. The latter sub-period results are similar to the DCC model with the exception of the insignificant YS variable. Panel E brings to light some interesting points to debate. For example, the macroeconomic conditions variable (ADSI) shows that during the period October 1992 to May 2003 a 1% increase in the

Table 8

Determinants of Normal Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

	Oct. 1992 - Oct. 2013 (Full Sample Period)	Oct. 1992 - May 2003 (Sub-period A)	June 2003 - Oct. 2013 (Sub-period B)
Panel A: Energy			
Constant	0.4898	2.2189	-0.0492
	(2.71)	(4.54)	(-0.25)
ADSI	0.0336	-0.0311	0.0891
	(1.70)	(-1.44)	(4.15)
Log(BDSI)	-0.2036	-0.5578	-0.1942
	(-4.85)	(-4.09)	(-4.09)
Log(VIX)	0.0854	-0.2595	0.5254
	(1.19)	(-3.20)	(5.42)
YS	0.1165	-0.0045	0.1100
	(3.12)	(-0.06)	(2.34)
SPE_ENERGY	0.1967	-0.4918	0.1594
	(3.56)	(-2.30)	(1.69)
R^2	0.2594	0.1924	0.4858
Panel B: Foods & Fibers			
Constant	0.0744	0.2820	0.0318
	(2.03)	(3.04)	(0.82)
ADSI	0.0031	-0.0094	0.0100
	(0.71)	(-1.58)	(2.48)
Log(BDSI)	0.0077	-0.0367	-0.0018
	(0.86)	(-1.48)	(-0.19)
Log(VIX)	-0.0090	-0.0470	0.0581
	(-0.61)	(-2.48)	(2.93)
YS	0.0294	0.0086	0.0188
	(3.70)	(0.56)	(2.04)
SPE_FOODFIB	0.0048	-0.0019	0.0124
	(0.95)	(-0.38)	(2.39)
R^2	0.0728	0.0197	0.1243

Table 8 (continued)

Determinants of Normal Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel C: Grains & Oilseeds			
Constant	0.0400	-0.1847	0.1534
	(0.29)	(-0.69)	(0.85)
ADSI	0.0429	0.0168	0.0756
	(3.17)	(1.06)	(3.97)
Log(BDSI)	-0.0077	0.0984	-0.0792
	(-0.24)	(1.25)	(-1.88)
Log(VIX)	0.0072	-0.0262	0.1903
	(0.11)	(-0.30)	(1.88)
YS	0.1308	0.0438	0.1299
	(5.29)	(0.93)	(3.18)
SPE_GRAINS	-0.0005	0.0005	-0.0982
	(-0.10)	(0.10)	(-3.03)
R^2	0.1194	0.0082	0.3015
Panel D: Livestock			
Constant	-0.0159	-0.1686	-0.0136
	(-0.26)	(-0.86)	(-0.18)
ADSI	0.0017	0.0143	-0.0134
	(0.27)	(1.72)	(-1.34)
Log(BDSI)	0.0009	0.0599	-0.0149
	(0.07)	(1.09)	(-0.77)
Log(VIX)	0.0241	-0.0015	0.1015
	(0.91)	(-0.04)	(2.56)
YS	0.0474	0.0558	0.0084
	(3.30)	(2.01)	(0.37)
SPE_LIVESTK	0.0012	-0.0022	-0.0063
	(0.29)	(-0.42)	(-0.97)
R^2	0.1032	0.0258	0.1814

Table 8 (continued)

Determinants of Normal Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel E: Precious Metals			
Constant	0.0412	1.0248	0.2669
	(0.34)	(3.71)	(2.23)
ADSI	0.0094	-0.0645	0.0399
	(0.53)	(-4.16)	(2.39)
Log(BDSI)	0.0001	-0.2842	-0.1048
	(0.00)	(-3.61)	(-4.31)
Log(VIX)	0.0504	0.0491	0.2608
	(0.99)	(0.81)	(3.60)
YS	0.0387	-0.1090	-0.0223
	(1.55)	(-2.79)	(-0.73)
SPE_PMETALS	0.0193	0.0020	-0.0180
	(0.78)	(0.08)	(-0.52)
R^2	0.0329	0.2522	0.2514

Note. This table provides the ordinary least squares (OLS) regression results for each of the commodity futures sub-sectors over the full sample period (October 1992 to October 2013) and two sub-periods (October 1992 to May 2003 and May 2003 to October 2013). In panels A, B, C, D, and E the dependent variable is the time-varying normal copula correlation between the weekly rates of return on the S&P 500 equity index and the equally-weighted weekly futures returns on the energy, foods & fibers, grains & oilseeds, livestock, and precious metals sub-sectors, respectively. The variables ADSI, BDSI, VIX, and YS represent the Aruoba-Diebold-Scotti Index, the Baltic Dry Shipping Index, the market volatility index, and yield spread, respectively. The variables SPE_ENERGY, SPE_FOODFIB, SPE_GRAINS, SPE_LIVESTK, and SPE_PMETALS represent speculation in the energy, foods & fibers, grains & oilseeds, livestock, and precious metals commodity futures sub-sectors, respectively. In all sample periods, Newey-West t-statistics are reported in parentheses below the corresponding coefficients, along with the R^2 of the regression.

macroeconomic conditions results in an decrease of 0.0645% in commodity-equity return correlations, and vice versa. However, over the period June 2003 to October 2013, the effect is the opposite, a 1% increase in macroeconomic conditions increases the commodity-equity return correlations by 0.0399%, and vice versa. Given that equity markets tend to prosper during economic booms, sub-period A underlies the fact that when equity returns were increasing, returns in the precious metals sub-sector were not commoving with them. Alternatively, when economic times were poor and equity markets were retrogressing, the commodity futures returns were moving in the opposite direction. Hence, the sub-sector returns were acting as a type of diversifying tool. A similar story is evident in the sub-periods of Panels A and B; these observations are very intriguing as they coincide with the literature which posits that the recent financialization and increased participation has otherwise diminished the benefits of commodity futures.

Overall, the results using the normal copula correlations as the dependent variable reveal many similarities and a few interesting differences from the DCC (baseline) case. It is clear that the broad market macroeconomic and financial indicator variables, which are generally associated with traditional asset (equity) market return fluctuations, now similarly influence commodity futures returns; thus, providing a measure of increasing return comovement between the two markets. These findings provide general support for the financialization of commodity futures argument. Additionally, the results also point out that the equity-commodity return correlations tend to increase in times of market distress (as proxied by YS), a downfall in global economic conditions (as proxied by BDSI), during periods of domestic market uncertainty (proxied by VIX), and with an improvement in domestic business conditions (proxied by ADSI) particularly over the full and sub-sample B periods. These findings seem to imply that commodity futures returns act less like a hedge or diversifying tool than the used too. However, our analysis also reveals that the magnitude and significance of these effects is, again, heterogeneous across sub-sectors.

4.1.3 The Student's t Copula

Table 9 presents our regression results using the student's t copula as the dependent variable. The student's t copula is a symmetrical but non-zero tail dependence structure which nests the normal copula. Panel A summarizes our findings for the energy sub-sector. Overall, regression results closely mirror that of the DCC and Normal copula approaches. Noticeably, however, the coefficient for the speculation factor (SPE_ENERGY) is quite larger under the student's t copula correlations, implying a greater role for the variable regarding the correlations movement. The results also reveal a larger R-squared for both the full sample and sub-period B. The overall consistency of the findings for the energy sub-sector across all correlation measures further solidify the results that its returns with the equity market are strongly determined by all of our broad macroeconomic and financial indicators as well as excess speculative activity. Furthermore, the evolution of the deterministic factors as seen by the change in significance and sign of the coefficients across the two sub-periods supports the postulations of the financialization argument, in which large capital inflows to the energy sub-sector have integrated its return structure with the traditional financial markets, hence altering the behavior of the assets and theoretically the potential benefits.

Panel B results, for the foods and fibers sub-sector, more so resemble the DCC findings than the normal copula. The allowance of symmetrical tail dependence gives the highest R-squared (approximately 40%) for the full sample period out of all dependence measures considered. We also find that all explanatory variables are highly significant in explaining the dynamic equity-commodity return correlations, whereas in the case of the normal copula only the variable YS is significant at conventional levels. In sub-period B we observe that YS has lost its explanatory power (when compared to the DCC model), yet the remaining significant coefficients are largely the same as seen in Table 7. The regression results using the student's t copula for the foods and fibers sub-sector show that the macroeconomic and financial indicator variables are relatively more important in the most recent sub-sample period compared to the initial sub-period. Moreover, the effects of the determinants have changed over the different sub-sample periods. Even though the results of the sub-sector are sensitive to the measure of dependence, the aggregate copula and DCC findings tell a story of increasing return correlation between the two asset markets via the deterministic factors.

Panels C and D, of the grains and oilseeds and livestock sub-sectors, strongly resemble those found in both Tables 7 and 8. Over the full sample period the results are unchanged regardless of the dependence measure used. Across the two sub-samples there are only minor changes in the deterministic factors of the equity-commodity return correlations. Interestingly, the findings of Panel C again document a negative speculation coefficient (SPE_GRAINS), such that a 1% increase in speculation activity results in a decrease of 0.1390% in the sub-sectors' equity-commodity correlations. Additionally, the sub-period B results of Panel D show that none of the explanatory variables are significant in explaining the dynamic equity-commodity correlations, whereas the prior tables found just the VIX variable to be of any importance. Overall, findings provide considerable credence to the observation that the determinants of the grains and oilseeds sub-sector have become more integrated with those of the equity markets, as seen by the considerable change in the explanatory variables; however, the livestock sub-sector

Table 9

Determinants of Student's t Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

	Oct. 1992 - Oct. 2013 (Full Sample Period)	Oct. 1992 - May 2003 (Sub-period A)	June 2003 - Oct. 2013 (Sub-period B)
Panel A: Energy			
Constant	1.0439	1.8784	0.3198
	(5.36)	(5.50)	(1.38)
ADSI	0.0839	0.0133	0.1594
	(3.88)	(0.99)	(5.64)
Log(BDSI)	-0.3933	-0.4636	-0.3734
	(-8.89)	(-5.00)	(-6.49)
Log(VIX)	0.0707	-0.3255	0.5853
	(0.94)	(-5.08)	(5.12)
YS	0.1064	0.1066	0.0983
	(2.75)	(2.10)	(1.69)
SPE_ENERGY	0.5307	-0.2247	0.6061
	(8.35)	(-1.28)	(5.66)
R^2	0.5265	0.2918	0.6661
Panel B: Foods & Fibers			
Constant	0.0236	0.3620	0.1582
	(0.30)	(3.08)	(2.66)
ADSI	0.0282	-0.0019	0.0241
	(3.56)	(-0.25)	(3.40)
Log(BDSI)	0.0478	-0.0317	-0.0499
	(2.71)	(-0.93)	(-3.49)
Log(VIX)	-0.1624	-0.2206	0.1158
	(-5.15)	(-9.46)	(4.11)
YS	0.1346	0.1048	0.0195
	(8.94)	(6.53)	(1.24)
SPE_FOODFIB	0.0311	0.0188	0.0305
	(3.95)	(4.40)	(2.83)
R^2	0.3956	0.4571	0.2330

Table 9 (continued)

Determinants of Student's t Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel C: Grains & Oilseeds			
Constant	0.2388	-0.2281	0.2665
	(1.38)	(-0.67)	(1.74)
ADSI	0.0403	0.0044	0.0869
	(2.33)	(0.22)	(3.54)
Log(BDSI)	-0.0260	0.1703	-0.1214
	(-0.67)	(1.78)	(-2.39)
Log(VIX)	-0.1191	-0.1884	0.1502
	(-1.55)	(-1.77)	(1.42)
YS	0.1644	0.0879	0.1604
	(6.36)	(1.55)	(3.37)
SPE_GRAINS	-0.0011	0.0003	-0.1390
	(-0.18)	(0.05)	(-3.41)
R^2	0.1057	0.0581	0.2895
Panel D: Livestock			
Constant	-0.0209	-0.1914	-0.0092
	(-0.28)	(-0.75)	(-0.10)
ADSI	0.0030	0.0118	-0.0076
	(0.39)	(1.04)	(-0.69)
Log(BDSI)	0.0024	0.0611	-0.0120
	(0.15)	(0.87)	(-0.55)
Log(VIX)	0.0180	0.0047	0.0736
	(0.55)	(0.09)	(1.44)
YS	0.0462	0.0564	0.0173
	(2.98)	(1.74)	(0.73)
SPE_LIVESTK	0.0051	0.0024	0.0016
	(1.03)	(0.32)	(0.20)
R^2	0.0463	0.0080	0.0816

Table 9 (continued)

Determinants of Student's t Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel E: Precious Metals			
Constant	0.0901	1.2855	0.4492
	(0.64)	(5.33)	(3.38)
ADSI	0.0219	-0.0670	0.0430
	(1.20)	(-4.82)	(2.67)
Log(BDSI)	-0.0024	-0.3569	-0.1594
	(-0.08)	(-5.27)	(-5.49)
Log(VIX)	-0.0081	-0.0388	0.3167
	(-0.15)	(-0.65)	(3.82)
YS	0.0732	-0.0270	-0.0604
	(2.56)	(-0.80)	(-1.69)
SPE_PMETALS	0.0290	0.0069	-0.0260
	(0.95)	(0.27)	(-0.60)
R^2	0.0390	0.3338	0.3299

Note. This table provides the ordinary least squares (OLS) regression results for each of the commodity futures sub-sectors over the full sample period (October 1992 to October 2013) and two sub-periods (October 1992 to May 2003 and May 2003 to October 2013). In panels A, B, C, D, and E the dependent variable is the time-varying student's t copula correlation between the weekly rates of return on the S&P 500 equity index and the equally-weighted weekly futures returns on the energy, foods & fibers, grains & oilseeds, livestock, and precious metals sub-sectors, respectively. The variables ADSI, BDSI, VIX, and YS represent the Aruoba-Diebold-Scotti Index, the Baltic Dry Shipping Index, the market volatility index, and yield spread, respectively. The variables SPE_ENERGY, SPE_FOODFIB, SPE_GRAINS, SPE_LIVESTK, and SPE_PMETALS represent speculation in the energy, foods & fibers, grains & oilseeds, livestock, and precious metals commodity futures sub-sectors, respectively. In all sample periods, Newey-West t-statistics are reported in parentheses below the corresponding coefficients, along with the R^2 of the regression.

seems to remain largely segmented from the equity markets via the insignificant coefficients on the explanatory variables.

The precious metals sub-sector results of Panel E are a hybrid of the findings of Tables 7 and 8. The macroeconomic and financial market determinants of the return correlations for this sub-sector are all significant in sub-period B, whereas only ADSI and BDSI are relevant in sub-period A. The interesting points of Panel E are that ADSI changes sign across the sub-periods and YS is statistically negative in the latter sample period (as in Table 7), which is opposite of expectations and means that a 1% increase in the YS (i.e. financial market stress) results in a decrease of 0.0604% in the sub-sectors equity market correlations. Provided that asset returns tend to commove more closely during times of financial distress, this result suggests that precious metals returns tend to move in the opposite direction of equities in periods of stress; this result is likely a product of the flight-to-quality, particularly for gold, during market downturns. Lastly, we again document that the speculation variables is insignificant in all regressions—highlighting the heterogeneity of the explanatory variable across the different commodity groups.

Given the findings from Table 9, as well as those in Tables 7 and 8, we conclude that the factors which explain the equity-commodity return correlations for the energy, grains and oilseeds, precious metals, and to a lesser degree the foods and fibers sub-sector, have significantly changed over the last decade. The return correlations between the two different asset classes have become increasingly explained by both macroeconomic and financial market variables. However, the inferences seem to be somewhat sensitive to the dependence measure used. The livestock sub-sector displays a different pattern with its equity market return correlations. The sub-sectors return dependence is generally not explained by the broad macroeconomic, financial, or speculation variables; the determinants of the sub-sectors equity-commodity correlations show the weakest evidence of increasing return integration.

The findings of Table 9 supplement the conclusions reached in Tables 7 and 8 in that the variables which proxy for local market distress (YS), uncertainty (VIX), the business cycle (ADSI), and global financial market destabilization (BDSI) show that commodity futures returns act less like a hedge or diversification tool as they tend to commove more strongly with equity market returns under shocks to the explanatory variables in the more recent sample period.

4.1.4 The Rotated-Gumbel Copula

Table 10 presents our regression results using the rotated-gumbel copula as the dependent variable. The rotated-gumbel copula is a left tail, non-linear, asymmetrical dependence structure, which is mostly present during extreme negative events, and is suitable for analyzing dependence over the left portion of the distribution. Panel A summarizes our findings for the energy sub-sector. In general, the findings reveal a similar pattern to the DCC model, normal copula, and student's t copula in the prior tables. This is interpreted to mean that the factors which drive the equity-commodity return dependence relationship under the previous dependence structures examined are very similar to those which drive the time-varying relationship in left-tail crises situations. However, the overall explanatory power of the model does decrease for both the full sample period and sub-period B upon examination of the R-squared which are 23.8% and 38.9%, respectively. The foods and fibers sub-sector, in Panel B, consistently shows that the VIX and yield spread (YS) are important determinants of the left-tail dynamic dependence structure for both the full and sub-sample periods. Notably, the coefficients across all periods are decidedly smaller than in previous tables. Additionally, the relatively smaller R-squared for the sample periods suggests that these explanatory variables matter less when left tail dependence is the dependent variable.

The results of Panel C, the grains and oilseeds sub-sector, are very similar to the findings of Table 9; the characteristics of the coefficients are not drastically different then what has already been documented. What is intriguing is that the variable SPE_GRAINS is negative and highly significant, which is what the other tables report as well. Hence, the effect of increased speculation on the dynamic correlations of the grains and oilseeds group is strikingly different than any other sub-sector examined. In addition, the R-squared of the regressions tend to be considerably lower than what was found in the prior tables. Similar to the interpretation of Panel B, this means that while these factors do help to explain the dynamic equity-commodity correlations of lower tail dependence, they do not have the substantial impact found under the more general or symmetric dependence models.

The results of Panel D, the livestock sub-sector, show a moderate increase in the significance of the variables which explain the correlations structure. While the R-squared remains low overall, it is on par with the results found in Table 8. Overall, we see that ADSI and BDSI are highly negatively significant, suggesting that macroeconomic conditions, and not financial market, are the primary drivers of the sub-sectors' dynamic correlations. The precious metals sub-sector, in Panel E, seems to moderately resemble the findings from the DCC model, except that in this case we find the coefficient on SPE_PMETALS is statistically significant for the first time. In sub-period B we document a negative relationship between speculation activity and equity-commodity correlations; this result has otherwise only been seen in the grains and oilseeds sub-sector.

Table 10

Determinants of Rotated-Gumbel Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

	Oct. 1992 - Oct. 2013 (Full Sample Period)	Oct. 1992 - May 2003 (Sub-period A)	June 2003 - Oct. 2013 (Sub-period B)
Panel A: Energy			
Constant	1.5606	1.9080	1.0817
	(10.81)	(7.31)	(6.01)
ADSI	0.0182	-0.0372	0.0768
	(1.12)	(-2.75)	(3.54)
Log(BDSI)	-0.1567	-0.1384	-0.0156
	(-5.15)	(-1.84)	(-3.75)
Log(VIX)	-0.0542	-0.3196	0.3291
	(-0.86)	(-5.06)	(2.85)
YS	0.0768	0.0635	0.0736
	(2.84)	(1.71)	(1.50)
SPE_ENERGY	0.2158	-0.1739	0.2407
	(4.14)	(-1.32)	(3.83)
R^2	0.2378	0.1809	0.3858
Panel B: Foods & Fibers			
Constant	1.0810	1.0730	1.0826
	(389.34)	(179.11)	(353.13)
ADSI	-0.0001	0.0007	-0.0060
	(-0.22)	(1.86)	(-1.66)
Log(BDSI)	-0.0004	0.0016	-0.0003
	(-0.66)	(0.96)	(-0.35)
Log(VIX)	0.0042	0.0055	0.0026
	(3.79)	(3.48)	(1.75)
YS	-0.0020	-0.0020	-0.0021
	(-3.68)	(-2.01)	(-2.44)
SPE_FOODFIB	0.0001	0.0002	-0.0002
	(0.33)	(0.65)	(-0.52)
R^2	0.0966	0.0937	0.0383

Table 10 (continued)

Determinants of Rotated-Gumbel Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel C: Grains & Oilseeds			
Constant	1.2279	1.1398	1.2879
	(11.27)	(4.04)	(8.50)
ADSI	0.0177	0.0036	0.0354
	(1.55)	(0.31)	(1.84)
Log(BDSI)	-0.0204	0.0331	-0.0600
	(-0.92)	(0.46)	(-1.80)
Log(VIX)	-0.0102	-0.1722	0.0256
	(-2.10)	(-2.54)	(0.32)
YS	0.0857	0.1009	0.0758
	(4.64)	(2.39)	(2.10)
SPE_GRAINS	-0.0039	-0.0024	-0.0851
	(-1.30)	(-0.89)	(-3.10)
R^2	0.0690	0.0752	0.1342
Panel D: Livestock			
Constant	1.0892	0.9634	1.1008
	(31.79)	(6.19)	(25.29)
ADSI	-0.0049	0.0020	-0.0123
	(-1.39)	(0.60)	(-1.82)
Log(BDSI)	-0.0185	0.0276	-0.0307
	(-2.58)	(0.66)	(-2.34)
Log(VIX)	-0.0257	-0.0305	0.0160
	(-2.07)	(-1.31)	(0.58)
YS	0.0220	0.0124	0.0028
	(2.41)	(0.69)	(0.16)
SPE_LIVESTK	0.0020	-0.0013	-0.0005
	(0.82)	(-0.42)	(-0.14)
R^2	0.1028	0.0246	0.1933

Table 10 (continued)

Determinants of Rotated-Gumbel Copula Correlations between the S&P 500 and Commodity Futures Sub-sectors

Panel E: Precious Metals			
Constant	0.9466	1.3979	1.2469
	(6.32)	(11.45)	(5.48)
ADSI	0.0335	-0.0268	0.0553
	(1.78)	(-3.03)	(1.88)
Log(BDSI)	0.0318	-0.0842	-0.0887
	(1.07)	(-2.50)	(-2.00)
Log(VIX)	0.0181	-0.0419	0.2478
	(0.35)	(-1.22)	(1.83)
YS	0.0673	-0.0120	-0.0076
	(2.55)	(-0.60)	(-0.14)
SPE_PMETALS	-0.0290	0.0241	-0.1401
	(-1.05)	(1.40)	(-2.28)
R^2	0.0396	0.1527	0.1813

Note. This table provides the ordinary least squares (OLS) regression results for each of the commodity futures sub-sectors over the full sample period (October 1992 to October 2013) and two sub-periods (October 1992 to May 2003 and May 2003 to October 2013). In panels A, B, C, D, and E the dependent variable is the time-varying rotated-gumbel copula correlation between the weekly rates of return on the S&P 500 equity index and the equally-weighted weekly futures returns on the energy, foods & fibers, grains & oilseeds, livestock, and precious metals sub-sectors, respectively. The variables ADSI, BDSI, VIX, and YS represent the Aruoba-Diebold-Scotti Index, the Baltic Dry Shipping Index, the market volatility index, and yield spread, respectively. The variables SPE_ENERGY, SPE_FOODFIB, SPE_GRAINS, SPE_LIVESTK, and SPE_PMETALS represent speculation in the energy, foods & fibers, grains & oilseeds, livestock, and precious metals commodity futures sub-sectors, respectively. In all sample periods, Newey-West t-statistics are reported below the coefficients in parentheses, along with the corresponding R^2 of the regression.

Overall, regression results pertaining to the rotated-gumbel dependence structure seem to indicate that across the full sample and two sub-periods macroeconomic and financial market variables are important in determining the equity-commodity dynamic correlations in crisis situations. However, the overall explanatory power of the relevant variables seems to be substantially reduced. Most importantly, these results, similar to our prior findings, reflect the fact that the impact of the explanatory variables across the different sub-sectors is heterogeneous both in terms of magnitude and sign.

4.2 Speculation and Dummy Variable Analysis

Given the importance which prior work has placed on speculation as a first-order determinant of increasing equity-commodity correlations, we conduct a dummy variable analysis to gain a better understanding of such effects. Instead of utilizing sub-sample analysis we create a dummy variable (DUM) which takes on a value of zero between the period October 1992 and May 2003 and is one otherwise. Subsequently, we create an interaction term (DUM*SPE_) between each sub-sectors' speculation measure and the dummy variable. In this setting, if we consider our macroeconomic and financial market variables as control variables and merely analyze the role of speculation on dynamic correlations we gain some additional insights. Table 11 provides the regression results using the dummy variable analysis. Panel A, B, C, and D report the results using the DCC, normal copula, student's t copula, and rotated-gumbel copula respectively.

The DUM variable is interpreted as the "premium" that post-period (i.e. sub-period B) correlation measures experience over the pre-period (i.e. sub-period A) correlations, provided that the sub-sectors respective speculation measure is zero. In general, we see that the DUM variable is positive for most sub-sectors and across different dependence measures; this means that dynamic correlations across the different groups have been on the rise since mid-2003

Table 11

Determinants of Correlations between S&P 500 and Commodity Futures Sub-sectors with Time Period Dummy Variable

	Constant	ADSI	Log(BDSI)	Log(VIX)	YS	SPE_	DUM	DUM*SPE_	R^2
Panel A: DCC									
Energy	0.8842	0.1522	0.0860	-0.3157	0.1105	-0.7091	-0.2041	1.2381	0.4791
	(3.81)	(4.22)	(3.77)	(-5.05)	(1.48)	(-2.87)	(-2.28)	(4.34)	
Foods & Fibers	0.1024	0.0416	0.0107	-0.0362	0.0131	0.0370	0.1120	-0.0144	0.4806
	(1.33)	(2.21)	(1.10)	(-1.87)	(0.40)	(4.58)	(5.00)	(-0.93)	
Grains & Oilseeds	0.2746	0.1102	0.0340	-0.0718	-0.0123	0.0048	0.1130	-0.0879	0.1429
	(1.80)	(4.28)	(2.39)	(-1.78)	(-0.19)	(1.04)	(2.59)	(-2.28)	
Livestock	-0.0521	0.0491	0.0031	0.0044	0.0331	0.0053	0.0016	-0.0021	0.0620
	(-0.64)	(2.50)	(0.40)	(0.21)	(0.93)	(0.75)	(0.08)	(-0.20)	
Precious Metals	0.4563	-0.0807	-0.0037	-0.1804	0.2080	-0.0057	0.2212	-0.0517	0.5092
	(4.65)	(-4.06)	(-0.34)	(-7.41)	(5.95)	(-0.30)	(7.03)	(-1.06)	
Panel B: Normal Copula									
Energy	0.5967	0.1094	0.0346	-0.2138	0.1218	-0.7881	-0.1040	0.9829	0.2973
	(3.06)	(2.80)	(1.79)	(-4.13)	(1.70)	(-2.72)	(-1.44)	(3.18)	
Foods & Fibers	0.1088	0.0170	0.0011	-0.0053	0.0037	0.0008	0.0063	0.0083	0.0892
	(2.81)	(1.89)	(0.25)	(-0.51)	(0.25)	(0.14)	(0.64)	(1.08)	
Grains & Oilseeds	0.1698	0.1121	0.0412	-0.0562	0.0287	0.0023	0.1010	-0.0799	0.1404
	(1.11)	(4.03)	(3.08)	(-1.38)	(0.42)	(0.45)	(2.41)	(-2.22)	
Livestock	-0.0315	0.0471	0.0016	0.0032	0.0275	0.0041	0.0100	-0.0097	0.1035
	(-0.45)	(2.69)	(0.25)	(0.18)	(0.95)	(0.73)	(0.55)	(-1.10)	
Precious Metals	0.3984	-0.0764	-0.0156	-0.1416	0.1623	0.0175	0.1893	-0.0885	0.2961
	(3.44)	(-3.11)	(-0.96)	(-4.95)	(3.40)	(0.60)	(5.43)	(-1.57)	
Panel C: Student's t Copula									
Energy	1.0478	0.1254	0.0925	-0.3588	0.0657	-0.3826	-0.2000	1.0382	0.5524
	(4.76)	(3.20)	(4.40)	(-6.01)	(0.89)	(-1.85)	(-2.46)	(4.29)	
Foods & Fibers	0.2236	0.0614	0.0149	-0.0359	-0.0821	0.0256	0.0950	-0.0024	0.5741
	(3.56)	(4.57)	(2.33)	(-2.29)	(-3.11)	(5.30)	(4.44)	(-0.17)	
Grains & Oilseeds	0.3908	0.1492	0.0398	-0.0805	-0.1021	0.0024	0.1253	-0.1097	0.1285
	(2.03)	(4.83)	(2.27)	(-1.58)	(-1.31)	(0.41)	(2.32)	(-2.36)	
Livestock	-0.0273	0.0456	0.0029	0.0030	0.0202	0.0066	0.0054	-0.0046	0.0448
	(-0.34)	(2.53)	(0.36)	(0.14)	(0.57)	(0.94)	(0.25)	(-0.41)	
Precious Metals	0.5641	-0.0791	-0.0111	-0.1903	0.1393	0.0280	0.2523	-0.1204	0.3939
	(4.57)	(-2.94)	(-0.73)	(-6.21)	(2.96)	(0.92)	(6.47)	(-1.92)	

Table 11 (continued)

Determinants of Correlations between S&P 500 and Commodity Futures Sub-sectors with Time Period Dummy Variable

Panel D: Rotated-Gumbel Copula									
Energy	1.5618	0.0866	0.0227	-0.1389	-0.0562	-0.2776	-0.1060	0.5583	0.2538
	(10.00)	(2.96)	(1.39)	(-3.45)	(-0.86)	(-1.78)	(-2.21)	(3.23)	
Foods & Fibers	1.0797	-0.0016	0.0000	0.0001	0.0038	0.0002	-0.0003	-0.0002	0.1023
	(368.89)	(-2.50)	(0.04)	(0.08)	(3.20)	(0.64)	(-0.45)	(-0.40)	
Grains & Oilseeds	1.2952	0.0885	0.0195	-0.0414	-0.1055	-0.0020	0.0651	-0.0708	0.0897
	(10.32)	(4.01)	(1.67)	(-1.30)	(-2.18)	(-0.71)	(1.69)	(-2.41)	
Livestock	1.0872	0.0195	-0.0054	-0.0203	-0.0208	0.0039	0.0079	-0.0044	0.1032
	(27.60)	(1.60)	(-1.43)	(-1.73)	(-1.28)	(1.39)	(0.63)	(-0.98)	
Precious Metals	1.2916	-0.0204	0.0117	-0.0961	0.0704	0.0321	0.2320	-0.2247	0.2246
	(7.96)	(-0.76)	(0.70)	(-2.46)	(1.24)	(1.51)	(4.34)	(-3.24)	

Note. This table provides the ordinary least squares (OLS) regression results for each of the commodity futures sub-sectors over the full sample period (October 1992 to October 2013) and two sub-periods (October 1992 to May 2003 and May 2003 to October 2013). In panels A, B, C, and D the dependent variables are the time-varying dynamic conditional correlation (DCC), time-varying normal copula correlations, time-varying student's t copula correlations, and time-varying rotated-gumbel copula correlations, respectively, between the weekly rates of return on the S&P 500 equity index and the equally-weighted weekly futures returns on the energy, foods & fibers, grains & oilseeds, livestock, and precious metals sub-sectors. The variables ADSI, BDSI, VIX, and YS represent the Aruoba-Diebold-Scotti Index, the Baltic Dry Shipping Index, the market volatility index, and yield spread, respectively. The variable SPE_ represent speculation in the energy, foods & fibers, grains & oilseeds, livestock, and precious metals commodity futures sub-sectors, respectively. The variable DUM is a dummy variable which takes on the value of 0 between the period October 1992 and May 2003 and 1 otherwise. The variable DUM*SPE_ is an interaction term between the period dummy (DUM) and the respective sub-sectors' speculation (SPE_). In all sample periods, Newey-West t-statistics are reported in parentheses below the corresponding coefficients, along with the R^2 of the regression.

irrespective of speculation. However, not all of the DUM coefficients are significantly different from zero. The unique sub-sector to this generalization is the energy sub-sector which has a negative and statistically significant DUM coefficient in three out of four panels. This actually means that with speculation is zero, the energy group has experienced an average decrease in equity-commodity correlations over the post-period. The interaction term, DUM*SPE_, affords us the ability to interpret the effect of increased speculation on the correlation measures over the two different sub-periods. If the interaction term is positive then the effect of speculation for increasing correlations is greater on post-period correlations than pre-period correlations. That is, speculation causes larger increases in the sub-period B correlations than the sub-period A correlations. If the interaction term is negative then the effect of speculation for increasing correlations is smaller on post-period correlations then pre-period correlations. That is, speculation causes smaller increases in the sub-period A correlations than the sub-period B correlations. Results across all panels indicate a positive and significant interaction term for the energy sub-sector, which aligns with our prior findings regarding the evolution of the energy sub-sector determinants. Furthermore, consistent with our prior analysis, we find that the interaction term for the grains and oilseeds sub-sector is negative and significant across all panels. The interaction terms for the foods and fibers and livestock sub-sectors are both highly insignificant across all panels. Lastly, in two of the four panels (DCC and normal copula) we find that the interaction term for the precious metals sub-sector is insignificant, while in the other two (student's t and rotated-gumbel copulas) the interaction term is negative and significant.

The dummy variable analysis allows us to view the effects of speculation on dynamic equity-commodity correlations in isolation. The results are consistent with the sub-sample examination but provide another view of the evolution of the determinant via the dummy variable and the interaction term. The findings once again confirm the intuition that the effects of speculation (and other determinants as well) are unique among the sub-sectors.

5. CONCLUDING REMARKS

Recent empirical research has documented an increase in equity-commodity return correlations over the last decade. Given this observation we bring to light new facts regarding the commodity futures market by investigating the determinants of the correlations and the evolution of these variables through time. Moreover, we undertake an analysis of the various sub-sectors of the futures market, as opposed to a commodity index, to highlight the heterogeneity of the different commodity groups.

In this paper we calculate the dynamic dependence structure between the returns of five commodity futures sub-sectors—energy, foods and fibers, grains and oilseeds, livestock, and precious metals—and a well-known value-weighted equity index—S&P 500. We then investigate the determinants of the equity-commodity return dependence structures, via sub-sample and dummy variable regression analysis, using several comprehensive macroeconomic, financial market, and speculation explanatory variables. We utilize the well-known DCC model as a baseline approach to our investigation as well as three time-varying copulas. We analyze (i) the normal copula—a symmetrical and frequent dependence structure which has no tail dependence, (ii) the student's t copula—a symmetrical but non-zero tail dependence structure which nests the normal copula, and (iii) the rotated-gumbel copula—a left tail, non-linear, asymmetrical dependence structure. Practically speaking, these copulas represent the most relevant shapes for finance and are frequently used in empirical papers.

We find that while copulas offer a more robust measure of time-varying dependence, there are many similarities between the DCC and copula dependence measures. We document that the equity-commodity correlations for the energy, grains and oilseeds, precious metals, and to a lesser extent the foods and fibers sub-sectors have become increasingly explainable by macroeconomic and financial market indicators, particularly after the period May 2003. The livestock sub-sector exhibits the smallest increase in integration with the equity market returns as the majority of explanatory variables are, generally, statistically insignificant across all sample periods examined. In particular, we note that the variables which relate to market distress (YS), uncertainty (VIX), the business cycle (ADSI), and global financial market destabilization (BDSI) show that commodity futures returns act less like a hedge or diversification tool in more recent years as they tend to commove more strongly with equities. However, the macroeconomic, financial, and speculation variables do exhibit heterogeneous effects in terms of significance, magnitude, and sign. Moreover, we document that increased participation by financial market speculators is not a primary determinant for all sub-sectors' dynamic equity-commodity return correlations. This suggests that other forces are at play regarding the previously documented market-wide increase in commodity correlations. Alternatively, our results underlie the importance of examining commodity sub-sectors as opposed to value-weighted commodity indices due to the heterogeneity among the different groups.

Our results have interesting implications for academicians and practitioners alike. Knowledge of the factors which drive the return dependence between different commodity futures and the equity market, how they have evolved over time, and the sensitivity of these factors to different forms of dependence will provide investors, particularly those involved in commodity futures a more detailed level of understanding of the overall market as we bring to light the empirical facts of the market. These details help economists grasp the links between the real economy and finance and inform policymakers how exogenous shocks affect the markets and how to design reforms of the financial system. Moreover, the analysis also helps to highlight potential investment benefits for non-index futures investors regarding asset allocation and risk management. For instance, given that all commodity sub-sectors are not equally affected by the broad macroeconomic and financial market variables which traditionally play a strong role in equity market returns, non-index commodity investors could utilize certain commodity sub-sectors, such as livestock for example, where the potential for diversification benefits are likely more intact.

REFERNCES

Aloui, R., Ben Aissa, M., and Nguyen, D., 2011. Global financial crisis, extreme interdependences, and contagion effects: The role of economic structure? *Journal of Banking and Finance* 35(1), 130-141.

Aruoba, S.B., Diebold, F.X., and Scotti, C., 2009. Real-Time Measurement of Business Conditions. *Journal of Business and Economic Statistics* 27(4), 417-427.

Asness, C.S., Moskowitz, T.J., and Pedersen, L.H., 2013. Value and Momentum Everywhere. *Journal of Finance* 68(3), 929-985.

Bakshi, G., Panayotov, G., and Skoulakis, G., 2011. The Baltic Dry Index as a Predictor of Global Stock Returns, Commodity Returns, and Global Economic Activity. *Working Paper*.

Bhardwaj, G. and Dunsby, A., 2013. The Business Cycle and the Correlation between Stocks and Commodities. *Journal of Investment Consulting* 14(2), 14-25.

Billio, M., Caporin, M., and Gobbo, M., 2006. Flexible Dynamic Conditional Correlation multivariate GARCH models for asset allocation. *Applied Financial Economics Letters* 2(2), 123-130.

Buyuksahin, B., Haigh, M.S., Harris, J.H., Overdahl, J.A., and Robe, M.A., 2009. Fundamentals, trading activity and derivative pricing. In: *Paper Presented at the 2009 Meeting of the European Finance Association*.

Buyuksahin, B., Haigh, M.S., and Robe, M.A., 2010. Commodities and equities: ever a 'market of one'? *Journal of Alternative Investments* 12(3), 75-95.

Buyuksahin, B. and Harris, J., 2011. Do Speculators Drive Crude Oil Futures Prices. *The Energy Journal* 32(2), 167-202.

Buyuksahin, B. and Robe, M.A., 2011. Does “Paper Oil” Matter? *Energy Market Financialization and Equity-Energy Linkages*, Working Paper.

Buyuksahin, B. and Robe, M.A., 2013. Speculators, commodities and cross-market linkages. *Journal of International Money and Finance* 42, 38-70.

Chen, X. and Fan, Y., 2006. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of Econometrics* 135(1-2), 125-154.

Chollete, L., de la Pena, V., and Lu, C.C., 2011. International diversification: A copula approach. *Journal of Banking and Finance* 35(2), 403-417.

Chong, J. and Miffre, J., 2010. Conditional correlation and volatility in commodity futures and traditional asset markets. *Journal of Alternative Investments* 12, 61-75.

Conover, C.M., Jensen, G.R., Johnson, R.R., and Mercer, J.M., 2010. Is now the time to add commodities to your portfolio? *Journal of Investing* 19, 10-19.

Creal, D., Koopman, S.J., and Lucas, A., 2013. Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics* 28(5), 777-795.

Danielsson, J., Shin, H.S., and Zigrand, J.P., 2011. Balance Sheet Capacity and Endogenous Risk. *Working Paper*. London School of Economics and Princeton University.

Daskalaki, C., Kostakis, A., and Skiadopoulos, G., 2014. Are there common factors in individual commodity futures returns? *Journal of Banking and Finance* 40, 346-363.

Daskalaki, C. and Skiadopoulos, G., 2011. Should investors include commodities in their portfolios after all? New evidence. *Journal of Banking and Finance* 35, 2606-2626.

Deheuvels, G., 1979. La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. *Academie Royale de Belgique, Bulletin de la Classe des Sciences* 65(5), 274-292.

Delatte, A. and Lopez, C., 2013. Commodity and equity markets: Some stylized facts from a copula approach. *Journal of Banking and Finance* 37(12), 5346-5356.

Embrechts, P., McNeil, A., and Straumann, D., 2002. Correlation and dependence in risk management: properties and pitfalls. In: Dempster, M. (Ed.), *Risk Management: Value at Risk and Beyond*. Cambridge University Press, Cambridge, UK, 176-223.

Engle, R., 2002. Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics* 20(3), 339-350.

Erb, C. and Harvey, C., 2006. The strategic and tactical value of commodity futures. *Financial Analysts Journal* 62(2), 69-97.

Erb, C., Harvey, C., and Viskante, T., 1994. Forecasting international equity correlations. *Financial Analysts Journal* 50(6), 32-45.

Fattouh, B., Kilian, L., and Mahadeva, L., 2013. The Role of Speculation in Oil Markets: What Have We Learned So Far? *The Energy Journal* 34(3), 7-33.

Genest, C., Ghoudi, K., and Rivest, L.P., 1995. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika* 82(3), 543-552.

Glosten, L.R., Jagannathan, R., and Runkle, D.E., 1993. On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance* 48(5), 1779–1801.

Gorton, G. and Rouwenhorst, K., 2006. Facts and fantasies about commodity futures. *Financial Analysts Journal* 62, 47–68.

Greely, D. and Currie, J., 2008. Speculators, Index Investors, and Commodity Prices. *The Goldman Sachs Group: Global Investment Research*, 1-22.

Hamilton, J., 2009. Causes and Consequences of the Oil Shock of 2007-08. *NBER Working Paper* No. 15002.

Hong, H. and Yogo, M., 2012. What does futures market interest tell us about the macroeconomy and asset prices? *Journal of Financial Economics* 105(3), 473-490.

Jensen, G.R., Johnson, R.R., and Mercer, J.M., 2000. Efficient use of commodity futures in diversified portfolios. *Journal of Futures Markets* 20, 489-506.

Kilian, L., 2009. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. *American Economic Review* 99(3), 1053-1069

Kilian, L. and Murphy, D.P., 2014. The role of inventories and speculative trading in the global market of crude oil. *Journal of Applied Econometrics* 29(3), 454-478.

Kole, E., Koedijk, K., and Verbeek, M., 2007. Selecting copulas for risk management. *Journal of Banking and Finance* 31(8), 2405-2423.

Korniotis, G.M., 2009. Does Speculation Affect Spot Price Levels? The Case of Metals with and without Futures Markets. In: *Division of Research and Statistics and Monetary Affairs*, Federal Reserve Board.

Kyle, A. and Xiong, W., 2001. Contagion as a Wealth Effect. *Journal of Finance* 56(4), 1401-1440.

Longin, F. and Solnik, B., 1995. Is the correlation in international equity returns constant: 1960-1990. *Journal of International Money and Finance* 14(1), 3-26.

Longin, F. and Solnik, B., 2001. Extreme correlation of international equity markets. *Journal of Finance* 56(2), 649-676.

Manner, H. and Reznikova, O., 2012. A Survey on Time-Varying Copulas: Specification, Simulations, and Application. *Econometric Reviews* 31(6), 654-687.

Patton, A., 2004. On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics* 2(1), 130-168.

Patton, A., 2006. Modelling asymmetric exchange rate dependence. *International Economic Review* 47(2), 527-556.

Patton, A., 2009. Copula-based models for financial time series. In: Andersen, T., Davies, R., Kreiss, J., Mikosch, T. (Ed.), *Handbook of Financial Time Series*. Springer, New York, 767-785.

Pavlova, A. and Rigobon, R., 2008. The role of portfolio constraints in the international propagation of shocks. *Review of Economic Studies* 75(4), 1215-1256.

Rosenberg, J. and Schuermann, T., 2006. A general approach to integrated risk management with skewed, fat-tailed risks. *Journal of Financial Economics* 79, 569-614.

Silvennoinen, A. and Thorp, S., 2013. Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money* 24, 42-65.

Sklar, A., 1959. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistiqué de l'Université de Paris* 8, 229-231.

Symeonidis, L., Prokopczuk, M., Brooks, C., and Emese, L., 2012. Futures basis, inventory and commodity price volatility: An empirical analysis. *Economic Modelling* 29, 2651-2663.

Tang, K. and Xiong, W., 2012. Index investing and the financialization of commodities. *Financial Analysts Journal* 68(5), 54-74.

Chapter 2

US Community Bank Failure: An Empirical Investigation

1. INTRODUCTION

The community banking industry has long been a pivotal cornerstone in US financial sector intermediation. The importance of the characteristically American enterprise stems from the provision of a unique combination of relationship lending services based on the knowledge and history of their generally smaller, rural clientele. This inherent flexibility and willingness of community banks to work with customers comes in stark contrast to big bank processes which are best described as transactional, quantitative, and standardized. However, the last 40 years have seen the overall number of small, locally owned community lenders markedly shrink. The savings and loan crisis of the 1980's saw the first culling of the industry. This was followed by broad sweeping changes in regulation and industry practices in the early 1990's which effectively removed the barriers to bank consolidation. Moreover, economic conditions, such as low interest rates and financial product innovations created a more favorable banking environment for larger financial institutions. Finally, the far-reaching financial shocks of the 2008 global financial crisis propelled the most recent attenuation of the industry.

According to a recent 2012 FDIC community banking study, the share of US credit market debt held by the domestic financial intermediaries declined by almost 50 percent between 1984 and 2011. Furthermore, the share of US banking assets held by community banks declined from 38 percent to 14 percent over the same period.[18] Lux and Greene (2015) report a similar decline in the community bank lending-market over the last two decades, but document a much larger decline, of roughly 50 percent, in total banking assets. Yet, in spite of the changes the US

[18] Statistics come from Federal Reserve, *Flow of Funds*.

financial industry has undergone community banks continue to play a critical role in key lending segments of the US economy. As of 2012, the FDIC notes that approximately 92 percent of FDIC insured banks and 94 percent of US banking organization are made up of community banks. Additionally, Lux and Greene (2015) report that the unique financial sector provides nearly 77 percent of agricultural loans and approximately 50 percent of small business loans. Furthermore, community banks play a pivotal role in real estate lending, particularly for housing, where knowledge of local market conditions and borrowers is paramount. In 2013, it was reported that the default rates for loans secured by family residential properties was approximately 3.47 percent for community banks with $1 billion or less in assets, while banks with over $1 billion in assets reported default rates three times that, at 10.42 percent.[19] The geographical and economic importance of these institutions cannot be overstated as the ability of small businesses and consumers to find and obtain adequate credit is vital to the overall health and growth of the US economy. Hence, by this standard, the failure of community banks is not merely trivial. We feel that for these reasons an analysis of why certain community banks fail, while others do not, is of particular merit and has important regulatory and policy implications.

While prior research has largely focused on US commercial banks as a whole, we feel that the indicators and implementation of early warning systems of community bank failure should recognize the distinct differences and risk profiles of community and non-community banks. As argued in Lux and Greene (2015), the financials of community banks are different than their larger (non-community) bank counterparts, with less leverage, less-robust returns, and less of an emphasis on technology. Additionally, community banks are not as intensely involved in capital or securitization markets, and their earnings streams tend to be less diverse which make

[19] Hester Peirce, Senior Research Fellow, The Mercatus Center at George Mason University, testimony on July 18, 2013, before the House Committee on Oversight and Government Reform, 113th Congress, 1st session.

them more vulnerable to economic and financial disruptions. Hence, recognizing the early symptoms of community bank failure is an initial step towards failure prevention and ensuring overall economic stability. A more accurate warning model can provide regulators with additional lead time to avert an institutions failure, or mitigate the impact of failure on the banking system and economy, if failure does in fact occur.

Prior studies confirm that accounting data can be exploited to differentiate between sound and unsound banking (see Pettway and Sinkey, 1980; Whalen, 1991; Demirguc-Kunt, 1989; Thomson, 1992) and that is the avenue we pursue. Using a broad set of bank-specific accounting data we incorporate information from balance sheet, income statement, and the CAMELS ratings to select the appropriate covariates for our various model specifications. In the spirit of Francis and Schipper (1999), who examine the claim(s) that financial statement information has lost its explanatory power for exchange-listed firms' returns, and Beaver et al. (2005), who investigate a secular change in the ability of financial ratios to predict firm bankruptcy, we similarly explore the relevance of such accounting information and financial ratios (i.e. CAMELS ratings) in their ability to predict community bank failure. Pappas et al. (2013) pursue a similar undertaking in their study of Islamic and conventional banks from Middle and Far East countries. Additionally, we incorporate a well-known US market liquidity component, the TED spread, into our models to link how macroeconomic liquidity shocks contribute to community bank failure. The rise of the TED spread precipitated the 2008 financial crisis foretelling trouble for not only the US economy but the banking sector as well.[20] Cole and Wu (2009) similarly extend their analysis of US commercial bank failure to incorporate macroeconomic components. Though they find

[20] Through 2006 and into the second quarter of 2007 the TED spread remained steadfast around 50 basis points; however, in August of 2007 the spread began to widen, reaching over 200 basis points in November 2007 (a year before the collapse of Lehman Brothers) and peaking at 315 basis points in September 2008. The TED spread remained elevated through year-end 2008, averaging nearly 150 basis points, before returning to pre-crisis levels in mid-2009.

evidence that declining economic growth contributes to the failure of banks with high non-performing loans, and shocks to interest rates make banks heavily relying on long-term borrowing more susceptible to failure, both GDP growth and short-term interest rates do not improve the predictive accuracy of their model.

We employ survival analysis to examine the characteristics of failed US community banks relative to a sample of non-failed US community banks in order to determine the significant indicators of community bank failure. Specifically, we adopt the semiparametric Cox Proportional Hazards model which has the advantage of requiring no distributional assumptions about the failure times of the banks. We utilize both Federal Deposit Insurance Corporation (FDIC) and Federal Reserve Bank historical data over the period 1992-2013 to examine the bank-specific and macroeconomic characteristics of community banks which failed between the years 2000-2013. Specifically, we examine 452 failed community banks consisting of 6,350 bank-year observations and 6,217 non-failed community banks consisting of 124,167 bank-year observations, for a total of 6,669 community banks and 130,517 bank-year observations. Our empirical results indicate that ordinary balance sheet and income statement information are a relatively ineffective way to predict community bank failure, notwithstanding balance sheet information offers an informational edge over income statement information and highlights that smaller banks (based on total assets) are actually less likely to fail than their larger community bank counterparts. Financial ratio information, in particular ratios which encompass capital adequacy, asset quality and liquidity, and earnings, provide a dramatic increase in the ability to predict community bank failure risk. We find that community banks which have declining amounts of equity and loan loss provisions as a percentage of total assets are much more likely to fail. We also find that banks with more commercial and industrial loans as well as real estate loans relative to total assets have increased failure risk. Interestingly, community banks which reduce their proportion of consumer lending as a percentage of total assets are more likely to fail—emphasizing the importance of community banks and their smaller, rural lending practices. Overall, income-based ratios such as net operating income to total assets and return on equity seem to have less relative importance in predicting community bank failure. Finally, given the strong relational nature of community banking, we find that a decrease in salary and wage expenses as a proportion of total assets results in a dramatic increase in community bank failure rates. We posit that this result is a byproduct of two qualitative factors specific to the community banking industry—excellent management and quality employee retention. The relevant covariate (i.e. salary and wage expenses to total assets), we argue, is an indirect proxy of managerial effectiveness and efficiency. Provided that "better" more valuable managers require increased compensation for their efforts, a reduction in salary and wages as a proportion of total assets results in a strong increase in failure rates. Thus, managerial effectiveness and efficiency seems to be a vital component of community banking industry survival. Moreover, quality employee retention for the small institutions is paramount. Community banks strongly rely on the flexible relationship banking paradigm much more so than non-community banks, as such the interpersonal business model is only as good as its employees and their connections with the local community. In order to retain quality, knowledgeable employees and reduce turnover community banks must pay reasonable wages and benefits to their valued employees. Additionally, we document that the use of the macroeconomic indicator of liquidity conditions, the TED spread, provides a substantial improvement in modeling predictive community bank failure. Specifically, we document that as macroeconomic liquidity conditions deteriorate (i.e. as the TED spread widens) there is a considerable rise in the risk of failure for community banks, especially for those banks which are already suffering from financial duress.

The remainder of this paper is organized as follows. Section 2 provides a brief review of the pertinent literature on failure prediction. Section 3 describes our methodology, detailing our community bank definition, data sources, variables selection, and research design. Section 4 presents our empirical results. Finally, section 5 provides concluding remarks.

2. LITERATURE REVIEW

Trying to identify why some banks fail while others continue to thrive has been an ongoing issue in financial literature since the late 1960's. Ravi Kumar and Ravi (2007) provide a comprehensive review of the research methodologies employed to try and solve bankruptcy prediction in both banks and firms. Their survey spans the period 1968-2005 and is largely organized by the type of techniques applied to the bankruptcy problem, such as: statistical models, neural networks, case-based reasoning, decision trees, operational research, evolutionary approaches, and soft computing techniques, among others. However, they emphasize that the most precise way of monitoring the financial condition of banks is via on-site examinations. These examinations are conducted on a bank's properties by official regulators every 12-18 months and are required by the FDIC Improvement Act of 1991. To this end, regulators utilize a six part ratings system, known as CAMELS, which evaluates the safety and soundness of the institution under review. The ratings system appraises banks according to the following basic functional areas: Capital adequacy, Asset quality, Management expertise, Earnings strength, Liquidity, and Sensitivity to market risk. These CAMELS ratings provide regulators with imperative information about the financial condition of the banks; however, Cole and Gunther (1995) report that these ratings decay quite rapidly.

The majority of prior research relating to bank failure and prediction has largely relied on methodologies such as discriminant analysis, logit regressions, and neural networks (see Bell, 1997; Kolari et al., 2002; Swicegood and Clark, 2001). Shumway (2001) was the earliest study to show how the dynamic hazard model outperforms these more traditional bankruptcy models, and that a new hazard model which combines both accounting and market information substantially improves predicting bankruptcy.[21] While the work of Shumway (2001) was focused on US corporate bankruptcies, over the period 1962-1992, he comprehensively demonstrates that a (dynamic) hazard model provides more consistent in-sample estimations and more accurate out-of-sample predictions. Beaver et al. (2005) extend the work of Shumway (2001) by evaluating corporate bankruptcy data over a more recent period, 1962-2002, and find there is a slight decline in predictive ability of financial ratios, but that it can be compensated for by adding market variables into the hazard estimation. While the hazard model has been widely applied to corporate bankruptcy prediction (see Chava and Jarrow, 2004; Campbell et al., 2008; Bonfim, 2009) it has received relatively less attention in the realm of predicting bank failure until recently. The notable exception to this is Wheelock and Wilson (2000) who implement the Cox Proportional Hazards model with time-varying covariates to analyze commercial banks over the period 1984-1993.[22] Using traditional CAMELS variables as covariates they find that poorly capitalized banks, less liquid banks, less (managerially) efficient banks, and less profitable banks are much more likely to fail. They also document that banks holding more risky asset portfolios are more likely to fail. Their results provide strong evidence in favor of the CAMELS ratings

[21] Dynamic hazard models are preferable to static models in predicting bankruptcy for three reasons: (1) hazard models control for how long a firm is at risk of failure, (2) hazard models incorporate information from panel data, and (3) hazard models incorporate information from many observations, producing more accurate out-of-sample forecasts.

[22] A few early studies implement the hazard framework pertaining to predictive bank failure; however, they use static, single period, hazard models (see Lane et al., 1986; Whalen, 1991).

system. Cole and Wu (2009) take a different approach and analyze the forecasting accuracy of the dynamic hazard model using both bank-specific and macroeconomic variables on a very large sample of US commercial banks. They find that the hazard model significantly improves the accuracy of in-sample and out-of-sample failure forecasting relative to a static probit model.

More recently, Alali and Romero (2013) use survival analysis, in particular the Cox Proportional Hazards model, to determine how early the indicators of bank failure can be observed. Utilizing a large sample of US commercial banks which failed between 2000 and 2012, they note that banks with high loan-to-asset and high personal-loan-to-assets are more likely to survive. Moreover, they document that older banks and banks with high real estate and agricultural loans, loan charge-offs, loan loss allowance, and non-performing loan-to-asset ratios are more likely to fail. Pappas et al. (2013) undertake a similar survival analysis in the banking sector except at the international level.[23] They employ the Cox model to estimate the conditional hazard rates for both Islamic and conventional banks from 20 Middle and Far East countries. They find that Islamic banks have significantly lower risk of failure both unconditionally and conditionally on time-varying bank-specific and macroeconomic (i.e. GDP growth and inflation) covariates. Hence, the implementation of early warning systems of bank failure should recognize the distinct risk profiles of the two banks.

Our research complements the more recent line of literature on bank failure prediction. We employ survival analysis, and in particular the (semiparamteric) Cox Proportional Hazards model, to investigate the early warning signs of community bank failure. We feel that the obvious differences between community bank and non-community bank profiles (discussed

[23] Other international survival analysis studies include: Molina (2002), Gomez-Gonzalez and Kiefer (2009), Sales and Tannuri-Pianto (2007), and Mannasoo and Mayes (2009); these studies utilize both semiparametric and parametric survival models.

earlier) may lend themselves towards less obvious differences in terms of bank-specific predictors of failure. As in the case of Islamic and conventional banks, the implementation of early warning systems and indicators of bank failure should recognize the distinct risk profiles of community and non-community banks. To this end we utilize a much broader set of bank-specific data than previous studies which incorporates information from balance sheets, income statements, and the CAMELS ratings to provide an extremely thorough overview of the community banking industry. This line of research, while extremely important to a well-functioning economy, has been somewhat overlooked. Additionally, prior work has also demonstrated that macroeconomic conditions can play a role in bank and firm failure. As such, we specifically incorporate a well-known US market liquidity component—the TED spread—into our survivor analysis models to ascertain the magnitude of the link between macroeconomic liquidity shocks and their contribution to community bank failure. The TED spread is a great indicator of interbank credit risk, liquidity, and the perceived health of the banking system as a whole. A rising TED spread indicates a decrease in market liquidity and a rise in the risk of default rates, particularly for non-performing loans. Moreover, during liquidity crises the interbank lending market does not function smoothly and any shock in interest rates can make banks which heavily rely on long-term borrowing much more vulnerable to failure.

3. METHODOLOGY

3.1 Community Banking Definition and Data Sample

In order to analyze the community banking industry in the US it is first necessary to define what it means to be a community bank. In general, a rough agreement exists on the characteristics which define a community bank as most of the attributes encompass how and where a community bank conducts its business practices. For instance, community banks primarily focus on traditional banking services in their local communities where they obtain the majority of their core deposits and provide the lion's share of their loans to small businesses and local consumers. This form of banking practice is often referred to as "relationship" lending and borrowing (as opposed to "transactional") since the small institutions have a specialized knowledge of their local community and customers.[24] This expertise allows community banks to base their lending decisions on unique local knowledge and (non-standard) long-term relationship data rather than customary underwriting models alone, which are typically implemented by larger banks.

While there is a general consensus about the characteristics of community banks, clearly defining them has been a more difficult venture; what constitutes small, basic banking activities is very subjective and hard to measure. Most regulators and practitioners cannot even fully agree as to what constitutes a community bank. In fact, the three largest banking regulators: the Office of the Comptroller of Currency (OCC), the FDIC, and the Federal Reserve Board all use different definitions for community banks. In fact, the FDIC recently changed their definition of community banks (in 2012) to establish standard requirements for lending and deposit gathering as well as limits on the geographic scope of operations that an institution must meet to be designated as a community bank. The new definition still remains loosely based on the $1 billion total asset (i.e. size) threshold, but goes beyond the typical size criteria alone in separating community from non-community banks.[25] The Federal Reserve Board defines community banks as having $10 billion or less in total assets, while the OCC uses a $1 billion total asset threshold. In general, the standard method used by prior research has been to define community banks according to a size threshold, per total assets, which has ranged anywhere from $750 million to

[24] Hein et al. (2005), Critchfield et al. (2004), and Berger and Udell (2001) provide more information on the practice of "relationship" banking.
[25] See the FDIC Community Banking Study (December 2012) for specific criteria; under the new thresholds 94% of US banking organizations are defined as community banks.

$10 billion. Though the threshold alone may be an imperfect criterion and may seem rather arbitrary, many studies use $1 billion in total assets as an approximate limit, which is typically applied at the charter level rather than the banking organization level.[26] In this study, we similarly define a community bank as an FDIC-chartered institution which has $1 billion of total assets, in inflation-adjusted dollars. We adjust the nominal total asset values for the entire sample into 2013 constant-dollars using the CPI-U measure of inflation as reported by the Bureau of Labor and Statistics (BLS). It is essential that any dollar-based yardstick be adjusted over time to account for inflation, economic growth, and the size of the banking industry. More aptly, $1 billion is not what it used to be.

We collect FDIC bank data, on an annual (year-end) basis, for all US banking institutions that are FDIC insured in the Statistics and Depository Institutions (SDI) database. The SDI database collects financial data for nearly 1,000 different variables for FDIC-chartered institutions; it includes information from income statements and balance sheets as well as other sources pertaining to derivatives, risky assets, and much more. Additionally, we obtain an indicator of US macroeconomic conditions (i.e. perceived credit risk in the general economy), the TED spread, from the Federal Reserve Bank of St. Louis. The TED spread is the difference between interest rates on interbank loans and short-term US government debt. More specifically, the TED spread is the (percentage) difference between the three-month LIBOR and the three-month T-bill interest rate. A rising TED spread is a sign that lenders believe the risk of default on interbank loans is increasing and it often precipitates a downturn in the US financial markets as liquidity is “drying up.” When the risk of bank default is considered to be decreasing, the TED spread decreases, and market liquidity accordingly increases.

[26] Critchfield et al. (2004) and the Federal Reserve Bank of Kansas City (June 2003) apply the $1 billion limit at the banking organization level; DeYoung et al. (2004), Hassan and Hippler (2014), the FDIC, and OCC apply the $1 billion limit at the charter level.

Initially, we apply our definition of a community bank to a sample of 516 US banking institutions from the "FDIC Failed Banks List" which failed between the period January 2000 and December 2013 and have at least three years of data. This list includes banks that failed and were subsequently acquired by another institution and banks which failed and were not acquired at all.[27] Given that the asset size of some community banks may have grown beyond the $1 billion inflation-adjusted threshold during some periods of the study we analyze the initial sample to ensure that all banks are at or below the size threshold during at least one of the last five years of available institutional data, and those banks which do not meet this criteria are removed. Additionally, to ensure that institutions which initially qualified as community banks (based on size) but grew far beyond that threshold in the very latter years of the sample are removed from the failed community banks group, we winzorize the right-tail of the failed banks group at the 0.5% level when sorted on total assets.[28] The results of this procedure yield a total of 452 failed community banks and 6,350 bank-year observations.

Next, we apply our definition of a community bank to a sample of 9,350 non-failed US banking institutions from the FDIC database which have financial data reported over the period January 1992 (as that is when the FDIC institutional data becomes available), or since the institutions inception, to December 2013. Similar to the procedure used for the failed banks sample, we remove community banks that do not have at least three years of data, whose asset size has grown beyond the $1 billion inflation-adjusted threshold during the last five years of the sample period, and those which grew far beyond the asset threshold in the very latter years of the

[27] Of the 516 failed US banks, 485 were acquired by other institutions and 31 were not acquired at all. There were 255 unique acquirers, and of those 167 made single bank acquisitions, while the rest (88) made multiple bank acquisitions (which accounted for approx. 3.6 acquisitions on average).

[28] We subsequently apply another filter where community banks missing return on equity (ROE) and return on assets (ROA) data are removed from the sample of failed community banks. However, none of the 452 community banks failed this filtering process.

sample by winzorizing the right-tail of the non-failed banks group at the 1% level when sorted on total assets. Furthermore, we delete banks which do not have FDIC reported ROE and ROA data.[29] The results of this procedure yield a total of 6,217 non-failed community banks and 124,167 bank-year observations. The combined sample of failed and non-failed community banks results in a total of 6,669 subjects and 130,517 bank-year observations. This comprehensive dataset allows for a thorough, retrospective examination of failed community banks relative to non-failed community banks to determine the bank-specific and macroeconomic characteristics of why US community banks fail.

3.2 Variable Selection

This study is exploratory by design, and as such we examine prominent balance sheet and income statement variables as bank failure predictors; we also identify and utilize a comprehensive set of financial ratios, using prior literature (from section 2) as a guide, which fall under the umbrella of the accounting-based CAMELS ratings. Market-based models have found past stock return and volatility data to also be useful in failure studies (see Pettaway and Sinkey, 1980; Curry et al. 2007; Bharath and Shumway, 2008; Campbell et al. 2008). Given that stock return data is only applicable to publicly listed banks and the universe of community banks are, generally, privately held we cannot use market data for the banks in our sample, and as such we only use accounting information to explain community bank failures.[30] However, as discussed in Pettway and Sinkey (1980) accounting information generally leads market price information so

[29] The ROE and ROA criterion results in the removal of only four community banks from the sample.
[30] Community bank accounting data is obtained via FDIC certificate numbers for each bank.

that the sole use of accounting information is justifiable.[31] We further extend the analysis by incorporating the macroeconomic TED spread covariate (Ted_Spread) into our failure analysis.[32]

Continued regulatory efforts, such as the Basel Accords, aimed at safeguarding the financial stability of banks continually rely on the assumption that capital (and liquidity) regulation make banks more resilient to shocks from the real economy. Kaplan and Minoiu (2013) highlight the role of bank balance sheet strength in the transmission of financial sector shocks to the real economy, noting that banks with strong balance sheets were better able to maintain lending during the 2008 financial crisis. Cole and Gunther (1995) find capital and troubled assets to be among the important variables in explaining the timing of bank failure. Given the empirical importance of balance sheet information in bank stability we analyze the following variables: total assets (Asset), total liabilities (Liab), total equity capital (Eqtot), Tier-one core capital (Riskcapt1), commercial and industrial loans (Loanci), loans to individuals (Loancon), all real estate loans (Loanre), farm loans (Loanag), and total loans and leases (Loanlease).[33] We similarly explore income statement information as conditioning variables to community bank failure. A careful review of a bank's financial statement can reveal key factors of the institution's financial condition; in fact, Cole and Gunther (1995) find net income to be an extremely important element of bank failure. The income statement variables analyzed include: total interest income (Intinc), income before extraordinary items (Incext), total interest expense (Intexp), total non-interest expense (Nonintinc), net income (Netinc), net operating income

[31] Management variables (ratios) are also excluded from this study due to data unavailability.
[32] Some more recent empirical research suggests that taking macroeconomic conditions into account may improve the prediction accuracy of default in corporate firms—although firm-specific characteristics are the major determinant of corporate failure (see Carling et al., 2007; Bonfim, 2009). Pertaining to bank failure, Arena (2008) studies the 1990's Latin America and East Asia banking crises and notes that individual bank conditions explain bank failures, while macroeconomic shocks (which triggered the crises) primarily destabilized the weaker banks.
[33] While we consider farm loans and its accompanying ratios in our preliminary analysis, it is largely excluded from our final analysis because data is missing for many institutions; consequently, it dramatically reduces the overall number of failed and non-failed community banks available for analysis.

(Netopinc), and provisions for loan and lease losses (LLLP). Lastly, we capture three components of the CAMELS ratings using financial ratios based on balance sheet and income information—Capital adequacy, Asset quality and liquidity, and Earnings. Prior work by Wheelock and Wilson (2000), Cole and Wu (2009), Kolari et. al (2002), Bell (1997), and Alali and Romero (2013) find success analyzing CAMELS variables to identify the characteristics that cause banks to fail using neural network, logit, and hazard models. As such, we include the following capital adequacy ratios for analysis: total equity capital to total assets (Eq_Asset), total equity capital to total loans and leases (Eq_Loanlease), and total equity capital to risk-weighted adjusted assets (Eq_RWA). The asset quality and liquidity ratios include: commercial and industrial loans to total assets (Loanci_Asset), loans to individuals to total assets (Loancon_Asset), real estate loans to total assets (Loanre_Asset), farm loans to total assets (Loanag_Asset), total loans and leases to total assets (Loanlease_Asset), loss allowance to total assets (Lossallow_Asset), net charge-offs to total assets (Chargeoff_Asset), total loan and lease loss provision to total assets (LLLP_Asset), loss allowance to total loans and leases (LLLP_Loanlease), net charge-offs to total loans and leases (Chargeoff_Loanlease), total loans and leases to total deposits (Loanlease_Dep), and Tier-one capital to risk-weighted assets (Riskcapt1_RWA). Finally, the earnings ratios include: income before extraordinary items to total assets (Incext_Asset), net operating income to total assets (Netopinc_Asset), net interest margin (Netint_Asset), salary and wage expenses to total assets (Wage_Asset), return on assets (ROA), and return on equity (ROE). Table 1 provides a summary of the independent (conditioning) variables.

Table 1

Dependent and Conditioning Variables

Variables	Symbol	Type	Definition
Dependent			
Bank Failure	Bank_Fail	Qualitative	Binary indicator variable which is equal to 1 for failed banks in the year in which they fail and 0 in all preceding years. The variable is equal to 0 in all sample years for surviving banks.
Independent			
Total Assets	Asset	Balance Sheet	The sum of all assets owned by the institution including cash, loans, securities, bank premises and other assets. This total does not include off-balance-sheet accounts.
Total Liabilities	Liab	Balance Sheet	Deposits and other borrowings, subordinated notes and debentures, limited-life preferred stock and related surplus, trading account liabilities and mortgage indebtedness.
Total Equity Capital	Eqtot	Balance Sheet	Total equity capital on a consolidated basis (note: beginning march 2009, includes the non-controlling (minority) interests in consolidated subsidiaries for CALL report and TFR filers).
Tier-one (core) Capital	Riskcapt1	Balance Sheet	Tier-one (core) capital includes: common equity plus noncumulative perpetual preferred stock plus minority interests in consolidated subsidiaries less goodwill and other ineligible intangible assets. The amount of eligible intangibles (including mortgage servicing rights) included in core capital is limited in accordance with supervisory capital regulations. As of March 2014, Advanced Approaches Institutions began reporting regulatory capital according to the amended Market Risk Capital Rule.
Commercial and Industrial Loans	Loanci	Balance Sheet	Commercial and industrial loans. Excludes all loans secured by real estate, loans to individuals, loans to depository institutions and foreign governments, loans to states and political subdivisions and lease financing receivables.
Loans to Individuals	Loancon	Balance Sheet	Loans to individuals for household, family, and other personal expenditures including outstanding credit card balances and other secured and unsecured consumer loans.
All Real Estate Loans	Loanre	Balance Sheet	Loans secured primarily by real estate, whether originated by the bank or purchased.

Table 1 (continued)
Dependent and Conditioning Variables

Farm Loans	Loanag	Balance Sheet	Loans to finance agricultural production and other loans to farmers. Excludes savings institutions filing a Thrift Financial Report.
Total Loans and Leases	Loanlease	Balance Sheet	Total loans and lease financing receivables, net of unearned income.
Total Interest Income	Intinc	Income Statement	Sum of income on loans and leases, plus investment income, interest on interest bearing bank balances, interest on federal funds sold and interest on trading account assets earned by the institution.
Income before Extraordinary Items	Incext	Income Statement	Income (loss) before security transactions, extraordinary items and other adjustments.
Total Interest Expense	Intexp	Income Statement	Total interest expenses.
Total Non-interest Income	Nonintinc	Income Statement	Income from fiduciary activities, plus service charges on deposit accounts in domestic offices, plus trading gains (losses) and fees from foreign exchange transactions, plus other foreign transaction gains (losses), plus other gains (losses) and fees from trading assets and liabilities.
Net Income	Netinc	Income Statement	Net interest income plus total noninterest income plus realized gains (losses) on securities and extraordinary items, less total noninterest expense, loan loss provisions and income taxes.
Net Operating Income	Netopinc	Income Statement	Net income excluding discretionary transactions such as gains (losses) on the sale of investment securities and extraordinary items. Income taxes subtracted from operating income have been adjusted to exclude the portion applicable to securities gains (losses).
Provisions for Loan and Lease Losses	LLLP	Income Statement	The amount needed to make the allowance for loan and lease losses adequate to absorb expected loan and lease losses (based upon management's evaluation of the bank’s current loan and lease portfolio). Prior to 2001 and after 2002, an allowance for transfer risk is also included to cover losses on international assets. Additionally, from 1997 to 2000, includes provision for credit losses on off-balance sheet credit exposures. Reflects net provision for losses on interest-bearing assets.

Table 1 (continued)
Dependent and Conditioning Variables

Total Equity Capital to Total Assets	Eq_Asset	Financial Ratio: Capital Adequacy	Total equity capital as a percent of total assets.
Total Equity Capital to Total Loans and Leases	Eq_Loanlease	Financial Ratio: Capital Adequacy	Total equity capital as a percent of total loans and lease financing receivables, net of unearned income.
Total Equity to Risk-Weighted Adjusted Assets	Eq_RWA	Financial Ratio: Capital Adequacy	Total equity capital to total risk-weighted adjusted assets. Risk-weighted assets are adjusted for risk-based capital definitions which include on-balance-sheet as well as off-balance-sheet items multiplied by risk-weights that range from zero to 200 percent. A conversion factor is used to assign a balance sheet equivalent amount for selected off-balance-sheet accounts.
Commercial and Industrial Loans to Total Assets	Loanci_Asset	Financial Ratio: Asset Quality and Liquidity	Commercial and industrial loans as a percent of total assets.
Loans to Individuals to Total Assets	Loancon_Asset	Financial Ratio: Asset Quality and Liquidity	Loans to individuals as a percent of total assets.
Real Estate Loans to Total Assets	Loanre_Asset	Financial Ratio: Asset Quality and Liquidity	All real estate loans as a percent of total assets.
Farm Loans to Total Assets	Loanag_Asset	Financial Ratio: Asset Quality and Liquidity	Farm loans as a percent of total assets.
Total Loans and Leases to Total Assets	Loanlease_Asset	Financial Ratio: Asset Quality and Liquidity	Total loans and lease financing receivables, net of unearned income, as a percent of total assets.
Loss Allowance to Total Assets	Lossallow_Asset	Financial Ratio: Asset Quality and Liquidity	Allowance reserve for loan and lease losses that is adequate to absorb estimated credit losses associated with its loan and lease portfolio (which also includes off-balance-sheet credit instruments) as a percent of total assets.
Net Charge-offs to Total Assets	Chargeoff_Asset	Financial Ratio: Asset Quality and Liquidity	Gross loan and lease financing receivable charge-offs, less gross recoveries, (annualized) as a percent of total assets.
Total Loan and Lease Loss Provisions to Total Assets	LLLP_Asset	Financial Ratio: Asset Quality and Liquidity	The annualized provision for loans and lease losses as a percent of total assets on a consolidated basis.

Table 1 (continued)
Dependent and Conditioning Variables

Total Loan and Lease Loss Provisions to Total Loans and Leases	LLLP_Loanlease	Financial Ratio: Asset Quality and Liquidity	Allowance for loan and lease losses as a percent of total loan and lease financing receivables, excluding unearned income.
Net Charge-offs to Total Loans and Leases	Chargeoff_Loanlease	Financial Ratio: Asset Quality and Liquidity	Total loans and leases charged-off (removed from balance sheet because of uncollectibility), less amounts recovered on loans and leases previously charged-off as a percent of total loans and lease financing receivables.
Total Loans and Leases to Total Deposits	Loanlease_Dep	Financial Ratio: Asset Quality and Liquidity	Total loans and lease financing receivables, net of unearned income, as a percent of the sum of all deposits including demand deposits, money market deposits, other savings deposits, time deposits and deposits in foreign offices.
Tier-one Capital to Risk-Weighted Adjusted Assets	Riskcapt1_RWA	Financial Ratio: Asset Quality and Liquidity	Tier-one (core) capital as a percent of risk-weighted assets as defined by the appropriate federal regulator for prompt corrective action during that time period.
Income before Extraordinary Items to Total Assets	Incext_Asset	Financial Ratio: Earnings	Income before extraordinary items as a percent of total assets.
Net Operating Income to Total Assets	Netopinc_Asset	Financial Ratio: Earnings	Net operating income (annualized) as a percent of average total assets.
Net Interest Margin	Netint_Asset	Financial Ratio: Earnings	Total interest income less total interest expense (annualized) as a percent of average total earning assets.
Salary and Wage Expenses to Total Assets	Wage_Asset	Financial Ratio: Earnings	Salary and employee benefit expenses as a percent of total assets.
Return on Assets	ROA	Financial Ratio: Earnings	Net income after taxes and extraordinary items (annualized) as a percent of average total assets.
Return on Equity	ROE	Financial Ratio: Earnings	Annualized net income as a percent of average equity on a consolidated basis (note: negative retained earnings are shown as N/A).

Table 1 (continued)

Dependent and Conditioning Variables

TED Spread	Ted_Spread	Macroeconomic	The (percentage) difference between the three-month LIBOR and the three-month T-bill interest rate.

Note. All financial variable data and definitions are obtained from the FDIC database; macroeconomic data is obtained from the Federal Reserve Bank of St. Louis.

3.3 Research Design

We utilize survival analysis, or time to occurrence of an event, which is better suited for our purposes of analyzing community bank failure than more traditional models, such as binary logit or ordinary least squares (OLS). Survival models have the ability to easily accommodate both lifetime and censored (particularly, right-censored) data. More importantly, however, these models overcome the pitfall of assumed normality, which is the true drawback to the linear regression framework. The distributions for time to an event (in this case community bank failure) are likely disparate from the commonly assumed normal pattern and are undoubtedly non-symmetric, and may even be bimodal.[34] A linear regression framework is not robust to these types of violations. However, survival analysis models can provide proper estimates of expected time to failure and relevant parameter covariates by substituting a more reasonable distribution assumption (e.g. Weibull, exponential, etc.) in the case of parametric modeling or making no distributional assumptions at all in the case of semiparametric and nonparametric modeling, all of which can be estimated via maximum likelihood (ML). Moreover, as mentioned above, survival models of bank failure naturally and easily control for the condition that the number of observation periods of a given bank may not represent the bank's entire lifespan (i.e. censoring).[35] Left-censoring means that the event (i.e. failure) occurs prior to a subject (i.e. bank) entering the study; however, given that no banks in the sample of this study failed prior to initial analysis (beginning in Jan. 1992), left-censoring is irrelevant. Contrastingly, right-censoring, means that a subject is under study for a period of time and thereafter is no longer observed. In the context of community bank failure, it is entirely plausible a bank could remain in business

[34] See Cleves et al. (2010) for further exposition.
[35] For a more complete and detailed discussion of survival analysis and its properties see Hosmer et al. (2008), Kalbfleisch and Prentice (2002), and Nelson (1972).

beyond the conclusion of the sample period (ending in Dec. 2013) and fail at some point in time afterwards. In fact, our entire sample of non-failed community banks (i.e. those which survive beyond year-end 2013) represent right-censored observations. Fortunately, the likelihood function can be easily expressed in the presence of right-censoring to account for such data.[36]

In our analysis, we adopt the semiparametric Cox Proportional Hazards model (of Cox, 1972) in addressing the issue of community bank failure risk and its associated covariates. This methodological choice has key advantages over both the parametric and non-parametric survival analysis models. Parametric models require making distributional assumptions; this raises the concern that the assumptions, and not the data, are determining the results. Contrastingly, semiparametric and nonparametric models require no assumptions about the distribution of failure times. The key insight into removing this distributional assumption is realizing that with survival data the events occur at given times, and that these events can be ordered and the analysis can be performed using the ordered survival times. Thus, under semiparametric and nonparametric analysis time plays no significant role other than ordering the observations. Despite this (conditional) similarity a significant difference still exists between the two modelling techniques. In the case of semiparametric models, one is parameterizing the effect of the covariate(s), so that a parametric component of the analysis still exists; contrastingly, an entirely nonparametric approach is utilized when no covariates exist, or are purely qualitative in nature.[37] Semiparametric models are parametric in the sense that the effect of the covariate(s) is assumed to take a definitive form. The semiparametric analysis is a combination of separate binary-outcome analyses (one per failure time), while parametric analysis is merely a

[36] In the case of semiparametric models, if subject i is censored at time t_i, then that particular subject enters all the individual failure-time studies up to and including time t_i, and after that is merely ignored (as the subject did not fail at that time).

[37] Nonparametric methods include those of Kaplan and Meier (1958), Nelson (1972), and Aalen (1978).

combination of several analyses (at all possible failure times). If no failures occur over a certain interval, such periods are non-informative in semiparametric analysis, but very informative in parametric analysis. So it stands that while semiparametric analysis is advantageous in that it is not concerned with intervening analyses, the parametric analysis is much more efficient if the proper distributional assumptions are made around the times when failures are not observed; however, choosing the appropriate distributional assumptions is quite challenging in practice. Lastly, while the semiparametric model makes no distributional assumptions of failure times, as that is taken care of with an ordering of how the failures occurred, it does make an assumption about how each subject's observed covariate value determined the probability that a subject would fail.[38]

To provide context to the survivor analysis model, which we explore shortly, let us denote T as time to failure event, where $\mathrm{T} \epsilon [0, \infty)$, and $\mathrm{F(t)}$ as its cumulative distribution function and $\mathrm{f(t)}$ as its probability density function, where $\mathrm{F(t)} = \mathrm{Pr(T \leq t)}$ and $\mathrm{f(t)} = \frac{-\mathrm{dF(t)}}{\mathrm{dt}}$. However, in survival analysis it is far more convenient to describe the probability distribution for T in terms of $\mathrm{S(t)}$, the survivor function, and $\mathrm{h(t)}$, the hazard function, rather than $\mathrm{F(t)}$ and $\mathrm{f(t)}$, respectively. The survivor function is merely the reverse cumulative distribution function of T, and is given by:

$$\mathrm{S(t)} = 1 - \mathrm{F(t)} = \mathrm{Pr(T > t)} \tag{1}$$

The survivor function gives the probability of surviving beyond time (i.e. year) t. In other words, it is the probability that there is no failure (event) prior to time t. At $\mathrm{t} = 0$ the function is equal to one and subsequently decreases as t approaches infinity. Alternatively, the hazard function, or conditional failure rate, gives the instantaneous rate of failure. The function is given by:

[38] The nonparametric approach does away with the distributional assumption and lets the data do the speaking.

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr(t + \Delta t > T > t | T > t)}{\Delta t} = \frac{f(t)}{S(t)} \tag{2}$$

More aptly, the hazard function is the probability of failure occurring within year t, conditional upon the subject having survived to the beginning of time t, divided by the width of the interval. The hazard rate can vary from zero (i.e. no risk) to infinity (i.e. certain instantaneous failure) and provides the rate at which risk is accumulated given the one-to-one relationship between the probability of survival past a certain time and the amount of risk that has been accumulated up to that particular time. The (time-varying) hazard rate of community bank failure risk is the primary object of interest in the present context.

The Cox Proportional Hazards model is formalized as:

$$h(t|x_i) = h_0(t)\exp(x_i\beta_x) \tag{3}$$

where $h_0(t)$ is the baseline hazard function, x_i is a row vector of covariates, and β_x is a column vector of regression coefficients to be estimated by the data. The covariates can be of an indicator, categorical, or continuous nature. The beauty of the Cox (1972) model is that $h_0(t)$ is given no particular parameterization; actually, the baseline hazard is left unestimated altogether. Since semiparamteric analysis is confined to only those times for which failure occurs, the baseline hazard drops out from calculation.[39] Since the model makes no assumptions about the shape of the baseline hazard function it can take any form one can imagine. However, whatever the general shape the baseline hazard does take, it is the same for all; the hazard for one subject is merely a multiplicative replica of another subject's. More aptly, the model assumes the covariates multiplicatively shift the baseline hazard function. Comparing subject i to subject j, the model explicitly states that:

[39] For a detailed and technical treatment of how this occurs see Kalbfleisch and Prentice (2002).

$$\frac{h(t|x_i)}{h(t|x_j)} = \frac{\exp(x_i\beta_x)}{\exp(x_j\beta_x)} \tag{4}$$

assuming the covariates x_i and x_j do not change over time. Furthermore, the Cox model assumes the hazard rate increases linearly with time conditional on the covariate(s). In our context the covariates consist of bank-specific balance sheet, income statement, financial ratio, and macroeconomic variables. In order to estimate the time-dependent covariate coefficients we de-mean the lagged variables so that the baseline hazard rate, $h_0(t)$, can be interpreted as the rate of an average bank in the population sample.[40] A value of $\widehat{\beta_x}$ greater (less) than 0 indicates that a rise in the x^{th} covariate increases (decreases) failure risk and decreases (increases) survival time. The hazard rate, $\exp(\beta_x)$, can be reformulated into $100 \times \exp(\beta_x - 1)$ so that it is interpreted as the expected percentage increase in failure risk for a one unit increase in the x^{th} covariate. In our analysis the Cox model covariate selection process is based on a general-to-specific procedure as outlined in Pappas et al. (2013).[41]

We employ the exact-marginal calculation method, or continuous-time calculation method, for tied failure events in our ML calculations; in the present context, this refers to community banks which failed during the same time (i.e. the same month and day). This calculation method assumes that the institutions which failed on the same day did not all fail at the exact same time on the given day and that we are merely limited by how precisely we can measure the failure time with the dataset. The exact-marginal calculation utilizes conditional probabilities of tied failures in the likelihood calculations and assumes continuous time which

[40] See Box-Steffensmeier and Jones (2004) for more information regarding the covariate estimates.
[41] Pappas et al. (2013) build on the work of Lane et al. (1986), utilizing a forward-and-backward variable selection procedure. The name refers to the fact that the technique can both drop and add covariates sequentially. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether $\beta_x = 0$, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

makes it mathematically impossible that the community bank failures occurred at precisely the same instant.[42]

4. EMPIRICAL RESULTS

4.1 Descriptive Statistics

Table 2 provides the summary statistics for the full sample of conditioning variables. The variables are grouped by the following categories: balance sheet, income statement, financial ratio, and macroeconomic. The financial ratio section is further sub-divided into the CAMELS ratings categories of capital adequacy, asset quality and liquidity, and earnings. We report the mean, median, standard deviation, minimum, maximum, number of observations, and unit of measurement for each variable.

Table 2

Summary Statistics for Full Sample of Conditioning Variables

Variables	Units	Mean	Median	Std. Dev.	Min.	Max.	Obs.
Balance Sheet							
Asset	\$M	144,930.50	88,420.00	162,720.90	19.00	1,957,120.00	130,517
Liab	\$M	130,520.84	78,728.00	168,543.10	0.00	14,264,000.00	130,517
Eqtot	\$M	15,125.00	9,144.00	18,509.29	-47,041.00	825,213.00	130,517
Riskcapt1	\$M	14,567.81	8,910.00	21,458.34	-45,673.00	2,325,000.00	130,517
Loanci	\$M	12,274.62	5,277.00	24,219.21	0.00	1,874,000.00	130,517
Loancon	\$M	6,772.55	3,314.00	16,154.63	0.00	1,479,739.00	130,507
Loanre	\$M	69,573.39	34,376.00	111,118.12	0.00	9,783,000.00	130,517
Loanag	\$M	4,457.50	1,081.00	9,925.26	0.00	383,488.00	118,945
Loanlease	\$M	93,415.47	52,889.00	114,639.90	0.00	1,639,110.00	130,517
Income Statement							
Intinc	\$M	8,080.45	5,148.00	9,090.73	0.00	305,089.00	130,495
Incext	\$M	1,136.91	723.00	4,319.96	-351,282.00	439,941.00	130,517
Intexp	\$M	2,959.99	1,760.00	3,615.47	-2.00	63,034.00	130,495
Nonintinc	\$M	1,402.20	447.00	8,744.79	-120,461.00	909,750.00	130,517
Netinc	\$M	1,136.90	724.00	4,320.72	-351,282.00	439,941.00	130,517
Netopinc	\$M	1,102.94	704.00	4,287.38	-351,375.72	439,941.00	130,517
LLLP	\$M	595.41	112.00	3,415.49	-12,198.00	370,000.00	130,517

[42] Kalbfleisch and Prentice (2002) provide a technical treatment of the marginal calculation method.

Table 2 (continued)

Summary Statistics for Full Sample of Conditioning Variables

Financial Ratio							
Capital Adequacy							
Eq_Asset	%	12.50	9.96	99.11	-13.51	15,242.01	130,517
Eq_Loanlease	%	77.97	16.48	3,998.17	-6,763.20	758,448.29	129,904
Eq_RWA	%	21.39	15.72	111.88	-18.95	19,094.34	130,517
Asset Quality and Liquidity							
Loanci_Asset	%	19.62	6.65	959.17	0.00	111,481.26	130,517
Loancon_Asset	%	5.86	4.29	6.52	0.00	282.93	130,507
Loanre_Asset	%	95.25	40.38	4,566.13	0.00	581,975.00	130,517
Loanag_Asset	%	5.59	1.38	8.65	0.00	73.61	118,945
Loanlease_Asset	%	61.61	62.74	65.98	0.00	8,347.96	130,517
Lossallow_Asset	%	54.23	0.79	4,499.71	0.00	574,360.50	130,517
Chargeoff_Asset	%	0.26	0.07	5.43	-6.86	1,742.11	130,517
LLLP_Asset	%	2.59	0.13	190.11	-25.42	22,010.71	130,517
LLLP_Loanlease	%	1.51	1.30	11.97	-4,285.71	100.00	129,925
Chargeoff_Loanlease	%	1.20	0.36	3.34	-27.44	470.44	129,956
Loanlease_Dep	%	75.24	73.71	327.58	0.00	107,033.34	130,468
Riskcapt1_RWA	%	20.75	15.22	109.48	-19.77	18,544.72	130,517
Earnings							
Incext_Asset	%	1.02	0.94	23.90	-3,471.04	3,254.80	130,517
Netopinc_Asset	%	0.84	0.98	2.58	-138.19	272.36	130,517
Netint_Asset	%	4.13	4.08	1.39	-166.67	72.64	130,514
Wage_Asset	%	3.48	1.55	148.21	-0.30	17,489.59	130,517
ROA	%	0.87	1.00	2.57	-138.19	272.36	130,517
ROE	%	7.53	9.32	78.18	-11,095.83	14,089.74	130,517
Macroeconomic							
Ted_Spread	%	0.51	0.41	0.32	0.15	1.55	130,517

Note. Descriptive statistics are aggregated for both failed and non-failed community banks and reported over the sample period 1992-2013.

The maximum number of bank-year observations for any given variable is 130,517. The variable Loanag has the lowest recorded number of observations at 118,945, which in many cases forces us to remove it from our analysis due to the substantial loss of observations, particularly for the failed banks group. Overall, the full sample of community banks has an average size of $144.93 million; however, the standard deviation is quite large at $162.72 million indicating substantial variation among community banks even at an asset threshold of $1 billion. Given the type of loan information available for study—Loanci, Loancon, Loanre, and Loanag—Loanre makes up the largest average dollar amount of loans for community banks at $69.57 million. Provided the critical role that community banks play in consumer lending, and particularly in local real estate, this observation is not too surprising. Income statement information shows average income measures (as measured by Incext, Netinc, and Netopinc) of the order of approximately $1.10 million, but again considerable disparity exists with large standard deviation measures of approximately $4.30 million. The abundance of financial ratios tells a similar story of considerable variation amongst the sample of community banks. The bottom of Table 2 shows the average annual statistics of the macroeconomic liquidity measure TED_Spread. Over the entire sample period the average of the spread is 0.51% with a median value of 0.41%. It attains a minimum average spread value of 0.15% and a maximum of 1.55%, with the maximum values occurring over the 2007-2009 period.

The full sample summary statistics provide a useful overall picture of the community banking industry, but it is intuitively more useful to evaluate the summary information by group. Thus, we separate the sample statistics into failed and non-failed community bank categories. Table 3 reports the decomposition of the full sample summary statistics into the two groupings. Panel A presents the summary statistics for the conditioning variables for the failed community banks, while panel B provides the summary information for the non-failed community banks. Comparing the two panels we see that the average size of a failed community bank is roughly $178.13 million while that of a non-failed community bank is smaller at $143.24 million. We similarly see that the variables Liab, Loanre, and Loanlease are considerably larger for failed community banks ($163.75, $106.66, and $130.88 million, respectively) than non-failed community banks ($128.82, $67.68, and $91.50 million, respectively). Analyzing the income statement variables we find the largest discrepancies between failed and non-failed community banks for the income-based variables Incext, Netinc, and Netopinc as well as the loss provisioning variable LLLP. Respectively, for the failed banks, we observe values of -$901.32, -$895.95, and -$897.15 thousand, as well as $2.25 million; for the non-failed banks we observe values of $1.24, $1.24, and $1.21 million, as well as $510.99 thousand. The drastic difference in average income measures between groups is rather intuitive as failed community banks are expected to be noticeably less profitable than non-failed community banks.

Table 3

Summary Statistics for Conditioning Variables by Group

Variables	Units	Mean	Median	Std. Dev.	Min.	Max.	Obs.
Panel A: Failed Banks							
Balance Sheet							
Asset	$M	178,025.18	106,741.00	204,705.12	1,755.00	1,957,120.00	6,350
Liab	$M	163,746.30	98,424.00	190,032.35	96.00	1,822,604.00	6,350
Eqtot	$M	14,278.88	8,452.00	18,382.19	-47,041.00	209,684.00	6,350
Riskcapt1	$M	13,707.81	8,214.00	17,486.31	-45,673.00	209,518.00	6,350
Loanci	$M	16,230.30	8,093.50	27,579.79	0.00	637,180.00	6,350
Loancon	$M	5,354.14	2,442.50	22,511.16	0.00	856,815.00	6,350
Loanre	$M	106,675.27	55,852.00	141,868.31	0.00	1,491,289.00	6,350
Loanag	$M	1,554.36	0.00	5,686.34	0.00	86,218.00	5,606
Loanlease	$M	130,879.60	73,545.50	162,044.60	0.00	1,639,110.00	6,350
Income Statement							
Intinc	$M	10,654.46	6,155.00	14,198.21	2.00	278,482.00	6,350
Incext	$M	-901.32	377.50	8,469.10	-165,024.00	49,809.00	6,350
Intexp	$M	4,621.63	2,576.00	5,925.42	0.00	63,034.00	6,350
Nonintinc	$M	1,308.45	474.50	4,274.97	-16,357.00	104,142.00	6,350
Netinc	$M	-895.95	384.00	8,489.81	-165,024.00	51,478.00	6,350
Netopinc	$M	-897.15	355.25	8,329.74	-154,895.27	49,709.00	6,350
LLLP	$M	2,246.03	295.00	7,994.67	-5,955.00	206,150.00	6,350
Financial Ratio							
Capital Adequacy							
Eq_Asset	%	9.85	8.60	8.20	-13.51	94.69	6,350
Eq_Loanlease	%	29.10	12.25	592.76	-6,763.20	44,854.54	6,341
Eq_RWA	%	16.45	11.74	52.16	-18.95	3,424.91	6,350

Table 3 (continued)

Summary Statistics for Conditioning Variables by Group

Asset Quality and Liquidity							
Loanci_Asset	%	9.79	7.79	8.53	0.00	73.91	6,350
Loancon_Asset	%	4.28	2.38	5.85	0.00	96.75	6,350
Loanre_Asset	%	52.33	54.98	19.82	0.00	105.44	6,350
Loanag_Asset	%	1.91	0.00	5.52	0.00	65.55	5,606
Loanlease_Asset	%	68.72	71.49	15.58	0.00	97.88	6,350
Lossallow_Asset	%	1.28	0.94	1.19	0.00	20.06	6,350
Chargeoff_Asset	%	0.62	0.10	1.52	-2.01	37.84	6,350
LLLP_Asset	%	0.87	0.29	1.72	-2.09	23.95	6,350
LLLP_Loanlease	%	1.90	1.34	2.12	0.00	100.00	6,341
Chargeoff_Loanlease	%	1.72	0.50	3.61	-9.89	78.63	6,342
Loanlease_Dep	%	82.99	83.51	31.58	0.00	1,705.65	6,345
Riskcapt1_RWA	%	16.02	11.32	52.11	-19.77	3,424.91	6,350
Earnings							
Incext_Asset	%	-0.34	0.62	2.74	-27.48	7.94	6,350
Netopinc_Asset	%	-0.45	0.69	3.44	-79.48	9.94	6,350
Netint_Asset	%	4.16	4.16	1.84	-9.44	71.25	6,350
Wage_Asset	%	1.74	1.61	0.87	-0.30	14.83	6,350
ROA	%	-0.44	0.72	3.47	-79.48	9.76	6,350
ROE	%	-19.28	7.19	301.29	-11,095.83	6,375.35	6,350
Macroeconomic							
TED_Spread	%	0.56	0.47	0.34	0.15	1.55	6,350
Panel B: Non-Failed Banks							
Balance Sheet							
Asset	$M	143,238.01	87,636.00	160,095.73	19.00	1,703,388.00	124,167
Liab	$M	128,821.67	77,985.00	167,193.46	0.00	14,264,000.00	124,167
Eqtot	$M	15,168.27	9,179.00	18,514.80	-2,984.00	825,213.00	124,167
Riskcapt1	$M	14,611.79	8,945.00	21,641.03	-5,468.00	2,325,000.00	124,167
Loanci	$M	12,072.33	5,168.00	24,017.34	0.00	1,874,000.00	124,167
Loancon	$M	6,845.09	3,364.00	15,757.46	0.00	1,479,739.00	124,157
Loanre	$M	67,675.97	33,606.00	108,975.01	0.00	9,783,000.00	124,167
Loanag	$M	4,601.09	1,203.00	10,067.12	0.00	383,488.00	113,339
Loanlease	$M	91,499.52	52,059.00	111,338.70	0.00	1,515,332.00	124,167
Income Statement							
Intinc	$M	7,948.79	5,106.00	8,729.40	0.00	305,089.00	124,145
Incext	$M	1,241.14	735.00	3,965.55	-351,282.00	439,941.00	124,167
Intexp	$M	2,875.00	1,730.00	3,434.55	-2.00	61,622.00	124,145
Nonintinc	$M	1,407.00	445.00	8,913.32	-120,461.00	909,750.00	124,167
Netinc	$M	1,240.86	737.00	3,964.31	-351,282.00	439,941.00	124,167
Netopinc	$M	1,205.23	716.12	3,944.46	-351,375.72	439,941.00	124,167
LLLP	$M	510.99	106.00	2,974.48	-12,198.00	370,000.00	124,167

Table 3 (continued)

Summary Statistics for Conditioning Variables by Group

Financial Ratio							
Capital Adequacy							
Eq_Asset	%	12.64	10.02	101.59	-3.60	15,242.01	124,167
Eq_Loanlease	%	80.48	16.69	4,097.26	-4,399.33	758,448.29	123,563
Eq_RWA	%	21.64	15.91	114.09	-5.92	19,094.34	124,167
Asset Quality and Liquidity							
Loanci_Asset	%	20.12	6.60	983.39	0.00	111,481.26	124,167
Loancon_Asset	%	5.94	4.39	6.55	0.00	282.93	124,157
Loanre_Asset	%	97.44	39.74	4,681.42	0.00	581,975.00	124,167
Loanag_Asset	%	5.78	1.55	8.73	0.00	73.61	113,339
Loanlease_Asset	%	61.24	62.28	67.53	0.00	8,347.96	124,167
Lossallow_Asset	%	56.93	0.79	4,613.32	0.00	574,360.50	124,167
Chargeoff_Asset	%	0.24	0.07	5.55	-6.86	1,742.11	124,167
LLLP_Asset	%	2.68	0.13	194.91	-25.42	22,010.71	124,167
LLLP_Loanlease	%	1.49	1.30	12.26	-4,285.71	96.82	123,584
Chargeoff_Loanlease	%	1.17	0.35	3.32	-27.44	470.44	123,614
Loanlease_Dep	%	74.84	73.20	335.77	0.00	107,033.34	124,123
Riskcapt1_RWA	%	20.99	15.40	111.62	-13.52	18,544.72	124,167
Earnings							
Incext_Asset	%	1.09	0.95	24.49	-3,471.04	3,254.80	124,167
Netopinc_Asset	%	0.91	0.99	2.51	-138.19	272.36	124,167
Netint_Asset	%	4.12	4.08	1.36	-166.67	72.64	124,164
Wage_Asset	%	3.57	1.54	151.95	0.00	17,489.59	124,167
ROA	%	0.94	1.01	2.50	-138.19	272.36	124,167
ROE	%	8.90	9.38	41.76	-1,132.20	14,089.74	124,167
Macroeconomic							
Ted_Spread	%	0.51	0.41	0.32	0.19	1.55	124,167

Note. Panel A reports the descriptive statistics for failed community banks over the period 1992-2013. Panel B reports the descriptive statistics for non-failed community banks over the period 1992-2013.

Differences in financial ratios between the groups also exist. For instance, all average capital adequacy ratios tend to be larger for non-failed community banks than failed community banks. However, this type of pattern is not so evident with the asset quality and liquidity ratios; the most notable difference between the failed and non-failed banks' asset quality and liquidity ratio sample statistics is the standard deviations which are substantially larger for the non-failed sample—however, this can likely be explained by the differences in group sample size. An overview of the earnings ratios shows much poorer performance by the failed banks group. This finding is not too surprising given the subpar income statement figures we observed for the failed banks in Panel A. In fact, all earnings ratios, with the exception of Netint_asset and Wage_Asset, produce a negative mean value. We find the largest discrepancy between the groups' earning ratios is for ROE, the failed community banks produce an average return of -19.28%, while that of the non-failed banks is approximately 8.90%.

The "eye-test" for analyzing the differences between the two sample groups can only provide so much insight as to why certain community banks fail and other survive. As such, Table 4 provides the statistical difference-in-means tests for the two groups of conditioning variables observed in Table 3. Tests of the balance sheet variables show that the difference-in-means for failed and non-failed community banks are significant at better than the 1% level in all cases. Similar results are obtained for the difference-in-means tests of the income statement variables—the lone exception is the Nonintinc variable which is statistically insignificant at conventional levels with a P-value of 0.38. Under the financial ratio section, specifically the capital adequacy ratios, only two of the four means tests (Eq_Asset and Eq_RWA) are statistically significant at better than the 5% level. This result comes as a bit of a surprise for the variable Eq_Loanlease since the difference between the failed mean (29.10%) and non-failed mean (80.48%) is seemingly so large; however, the extremely large standard deviation associated with the non-failed banks provides some context to the result. Regarding the asset quality and liquidity ratios we find significant difference-in-means results, at the 5% level or better, for eight (Loancon_Asset, Loanag_Asset, Loanlease_Asset, Chargeoff_Asset, LLLP_Loanlease, Chargeoff_Loanlease, Loanlease_Dep, and Riskcapt1_RWA) of the 12 variables considered. As described earlier, the insignificant results are largely attributable due to the relatively enormous

Table 4

Univariate Difference-in-Means Tests of Conditioning Variables between Failed and Non-Failed Community Banks

	Failed Banks (452 Banks)		Non-Failed Banks (6,217 Banks)				
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean Diff.	t-statistic	P-value
Balance Sheet							
Asset	178,025.18	204,705.12	143,238.01	160,095.73	34,787.17	16.63	0.00
Liab	163,746.30	190,032.35	128,821.67	167,193.46	34,924.63	16.12	0.00
Eqtot	14,278.88	18,382.19	15,168.27	18,514.80	-889.39	-3.73	0.00
Riskcapt1	13,707.81	17,486.31	14,611.79	21,641.03	-903.98	-3.27	0.00
Loanci	16,230.30	27,579.79	12,072.33	24,017.34	4,157.97	13.35	0.00
Loancon	5,354.14	22,511.16	6,845.09	15,757.46	-1,490.95	-7.17	0.00
Loanre	106,675.27	141,868.31	67,675.97	108,975.01	38,999.29	27.357	0.00
Loanag	1,554.36	5,686.34	4,601.09	10,067.12	-3,046.74	-22.48	0.00
Loanlease	130,879.60	162,044.60	91,499.52	111,338.70	39,380.12	26.77	0.00
Income Statement							
Intinc	10,654.46	14,198.21	7,948.79	8,729.40	2,705.67	23.18	0.00
Incext	-901.32	8,469.10	1,241.14	3,965.55	-2,142.47	-38.79	0.00
Intexp	4,621.63	5,925.42	2,875.00	3,434.55	1,746.63	37.75	0.00
Nonintinc	1,308.45	4,274.97	1,407.00	8,913.32	-98.54	-0.88	0.38
Netinc	-895.95	8,489.81	1,240.86	3,964.31	-2,136.81	-38.66	0.00
Netopinc	-897.15	8,329.74	1,205.23	3,944.46	-2,102.37	-38.33	0.00
LLLP	2,246.03	7,994.67	510.99	2,974.48	1,735.04	39.72	0.00
Financial Ratio							
Capital Adequacy							
Eq_Asset	9.85	8.20	12.64	101.59	-2.79	-2.19	0.03
Eq_Loanlease	29.10	592.76	80.48	4,097.26	-51.38	-1.00	0.32
Eq_RWA	16.45	52.16	21.64	114.09	-5.19	-3.61	0.00
Asset Quality and Liquidity							
Loanci_Asset	9.79	8.53	20.12	983.39	-10.33	-0.84	0.40
Loancon_Asset	4.28	5.85	5.94	6.55	-1.66	-19.76	0.00

Table 4 (continued)

Univariate Difference-in-Means Tests of Conditioning Variables between Failed and Non-Failed Community Banks

Loanre_Asset	52.33	19.82	97.44	4,681.42	-45.11	-0.70	0.44
Loanag_Asset	1.91	5.52	5.78	8.73	-3.87	-32.85	0.00
Loanlease_Asset	68.72	15.58	61.24	67.53	7.48	8.82	0.00
Lossallow_Asset	1.28	1.19	56.93	4,613.32	-55.66	-0.96	0.34
Chargeoff_Asset	0.62	1.52	0.24	5.55	0.38	5.45	0.00
LLLP_Asset	0.87	1.72	2.68	194.91	-1.81	-0.74	0.46
LLLP_Loanlease	1.90	2.12	1.49	12.26	0.41	2.65	0.01
Chargeoff_Loanlease	1.72	3.61	1.17	3.32	0.55	12.76	0.00
Loanlease_Dep	82.99	31.58	74.84	335.77	8.15	1.93	0.05
Riskcapt1_RWA	16.02	52.11	20.99	111.62	-4.97	-3.53	0.00
Earnings							
Incext_Asset	-0.34	2.74	1.09	24.49	-1.44	-4.68	0.00
Netopinc_Asset	-0.45	3.44	0.91	2.51	-1.36	-41.08	0.00
Netint_Asset	4.16	1.84	4.12	1.36	0.04	1.97	0.05
Wage_Asset	1.74	0.87	3.57	151.95	-1.83	-0.96	0.34
ROA	-0.44	3.47	0.94	2.50	-1.38	-41.88	0.00
ROE	-19.28	301.29	8.90	41.76	-28.19	-28.11	0.00
Macroeconomic							
Ted_Spread	0.56	0.34	0.51	0.32	0.05	11.54	0.00

Note. The difference-in-means statistics are reported for all financial and macroeconomic variables; both t-statistics and P-values are reported in the far right column, respectively.

standard deviations of the non-failed banks. Lastly, we note that five (Incext_Asset, Netopin_Asset, Netintm, ROA, and ROE) of the six earnings ratios have significantly different means for the two groupings.

Overall, the sample statistics provide some general insights into the community bank failure investigation, but more so they validate the need for a more technical treatment into the reasons as to why certain community banks fail. In the following section we delve deeper into this issue by utilizing the Cox Proportional Hazards model to analyze and highlight the importance of certain covariates in predicting community bank failure.

4.2 Cox Proportional Hazards Regression Analysis

We utilize survival analysis techniques to gain a better understanding of which variables can help predict community bank failure. Given the exploratory nature of our study we estimate separate Cox Proportional Hazards models for balance sheet variables, income statement variables, capital adequacy ratios, asset quality and liquidity ratios, earnings ratios, and finally a comprehensive model which aggregates all financial ratio information. We apply the forward-and-backward variable procedure, discussed earlier, to identify the various model specifications which yield the best overall fit of the data.[43] For each regression model we report three different specifications: model I which includes only bank-specific covariates, model II which includes the same bank-specific covariates as model I but also incorporates the macroeconomic covariate, and model III which provides an alternative specification to that of model II. The results of the estimated models are provided in Tables 5 through 10.

An overall picture of the US community banking industry is given in Figure 1. It shows the Nelson-Aalen cumulative (community bank) hazard estimates as a plot of the percentage of

[43] Correlation matrices are also utilized to help identify the appropriate selection of variables in the various regression models; the correlation matrices are reported in Appendix A.

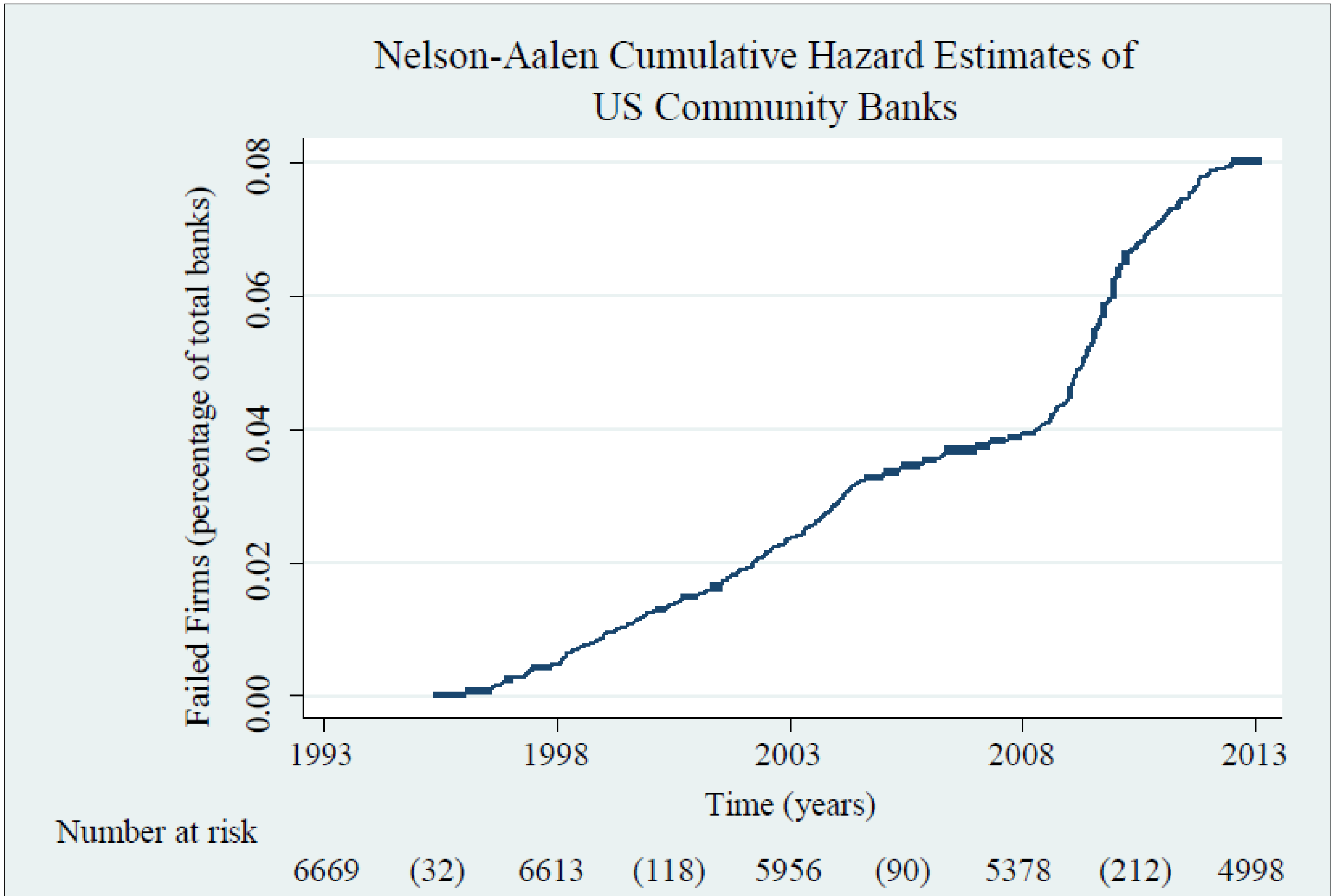

Nelson-Aalen Cumulative Hazard Estimates of
US Community Banks
Failed Firms (percentage of total banks)
0.00
0.02
0.04
0.06
0.08
1993
1998
2003
2008
2013
Time (years)
Number at risk
6669 (32) 6613 (118) 5956 (90) 5378 (212) 4998

Figure 1

failed community banks relative to total community banks over the entire sample period. Overall, approximately 0.07% of the community banks in the entire sample have failed by the end of 2013. Below the x-axis of Figure 1 the number of community banks at risk and the number of failed community banks (listed in parenthesis) for a given time interval are displayed. For instance, between the years 1998 and 2003 approximately 118 community banks failed while 6,613 were at risk at the beginning of the period. The largest failure rate occurs between 2008 and 2013 with a total of 212 banks succumbing to failure. This observation coincides with the far reaching consequences of the 2008 global financial crisis on the US financial and banking sectors.

Table 5 reports the regression results of conditioning community bank failure on balance sheet information. Model I shows that the balance sheet variables Asset, Riskcapt1, Loanci, and Loancon are all significant at better than the 1% level, while Loanre is only significant at the 10% level. The negative coefficients on Riskcapt1, Loanci, and Loancon suggest that the risk of community bank failure (i.e. the hazard rate) increases (or alternatively the survival likelihood decreases) as these covariates decrease. Thus, as the level of Tier-one core capital (a measure of a bank's financial strength) decreases we see an increase in the rate of community bank failure; similarly, as the level of both commercial and industrial loans and loans to consumers decreases a corresponding increase in the community bank failure rate occurs. The positive coefficients on Asset and Loanre indicate that the risk of bank failure increases (or alternatively the survival likelihood decreases) as these covariates increase. This indicates that larger community banks (per the covariate Asset) are relatively more likely to fail than smaller ones. Taken together with the results of Tables 3 and 4 which show that the average size of failed community banks is larger than non-failed banks and that this difference is statistically significant, we infer that the smallest community banks are not necessarily the most susceptible to failure. Prior work, such as Wheelock and Wilson (2000) and Cole and Wu (2009), generally document that smaller banks are more likely to fail in the universe of US commercial banks. Our finding complements this line of literature in that the size effect is not fully linear—thus, while community banks may be more susceptible to failure relative to non-community banks, the smallest community banks are not the most at risk. We postulate that these smaller institutions likely maintain more risk-averse profiles relative to their larger counterparts, hence insulating them more from outside financial shocks. Likewise, the positive coefficient of Loanre indicates that an increase in real estate loans for an institution means an increased likelihood of failure. The strong link between the 2008 global financial crisis and the high default rates in the real estate lending market provide a reasonable explanation for this finding.

The corresponding hazard ratios (i.e. exponentiated coefficients) in the table provide an estimate of the rate of failure for a one-unit increase in the respective covariate. Hazard ratios larger than 1.0 indicate the increase in the rate of failure occurring for a one-unit increase in the associated covariate, after controlling for other factors in the model. Conversely, hazard ratios smaller than 1.0 indicate the increase in the rate of failure occurring for a one-unit decrease in the associated covariate, after controlling for other factors in the model. Due to the small magnitude of the covariate coefficients all of the hazard ratios in Model I are very close to one. The likelihood ratio reports the test of whether the covariates are jointly equal to zero. In the case of Model I, and all subsequent models for that matter, we soundly reject the null hypothesis that the coefficients are jointly zero. The overall model fit is reported below via the AIC, BIC, and log likelihood statistics. Finally, the total number of banks, the total number of bank failures, and total number of bank-year observations are reported in the latter third of the table. For Model I we utilize all 6,669 banks, including all of the 452 failed banks, for analysis—yielding a total of 130,507 bank-year observations.

Model II of Table 5 reports the regression results of conditioning community bank failure on both balance sheet information and the macroeconomic liquidity measure. The inclusion of the covariate Ted_Spread yields similar results for the bank-specific variables as in Model I, except that Lonre is now statistically insignificant. Interestingly, however, the Ted_Spread is highly significant and has a coefficient (102.0856) of high economic significance relative to the much smaller coefficients of the bank-specific covariates. The positive coefficient and correspondingly large hazard ratio associated with Ted_Spread indicates that a one-unit increase in the macroeconomic covariate is associated with a relatively strong increase in the rate of failure for community banks; that is, as market liquidity decreases the hazard rate for community banks markedly increases (or alternatively the survival likelihood decreases). Moreover, the AIC, BIC, and log likelihood statistics validate the use of the macroeconomic covariate in the model as all measures improve relative to Model I. Lastly, Model III provides an alternative specification to that of Model II as we substitute the covariate Loanlease in and remove the covariate Asset due to the high correlation between the two.[44] The inclusion of the variable Loanlease yields a highly significant, economically small negative coefficient (-0.000010) similar to the other bank-specific variables. In general, the covariates Riskcapt1, Loancon, and Ted_Spread retain their coefficient sign, magnitude, and statistical significance. However, in the alternative model the covariate Loanci switches sign and Loanre now becomes significant at better than the 1% level. While this alternative specification shows some importance for the covariate Loanlease and Loanre, it provides an overall worse fit of the data relative to Model II.

[44] Wheelock and Wilson (2000) note that loans and leases are typically the least liquid and most risky portion of a bank's assets; in some cases loans and leases can represent the largest portion of a bank's assets.

Table 5

Cox Survival Model Conditioned on Balance Sheet and Macroeconomic Covariates

	Model I Balance Sheet			Model II Balance Sheet & Macro			Model III Balance Sheet & Macro		
Variables	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio
Asset	0.000006	0.00	1.000006	0.000006	0.00	1.000006			
Riskcapt1	-0.000101	0.00	0.999899	-0.000104	0.00	0.999896	-0.000101	0.00	0.999899
Loanci	-0.000008	0.00	0.999992	-0.000009	0.00	0.999991	0.000011	0.00	1.000011
Loancon	-0.000070	0.00	0.999930	-0.000070	0.00	0.999930	-0.000040	0.00	0.999960
Loanre	0.000002	0.10	1.000002	0.000002	0.12	1.000002	0.000020	0.00	1.000020
Loanlease							-0.000010	0.00	0.9999899
Ted_Spread				102.0856	0.00	2.16E+44	97.5489	0.00	2.32E+42
Likelihood Ratio	1,334.03	0.00		1,347.72	0.00		1,316.44	0.00	
AIC		6,296.17			6,284.47			6,315.76	
BIC		6,330.20			6,325.31			6,356.59	
Log Likelihood		-3,143.09			-3,136.24			-3,151.88	
Banks		6,669			6,669			6,669	
Failures		452			452			452	
Obs. (bank-year)		130,507			130,507			130,517	

Note. The Cox regression models are based on balance sheet and macroeconomic information. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether β_x=0, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

Table 5 provides important information about community bank failure conditional on asset size and loan specifics. Moreover, it shows that macroeconomic liquidity conditions play a significant role in the conditional failure risk of the institutions. Yet, conditioning on only balance sheet information is not without its shortfalls; the small economic significance of the associated coefficients implies that the sole use of balance sheet information to predict community bank failure, or hazard rates, is a relatively inefficient approach.

Table 6 reports the results of conditioning community bank failure on income statement information. Model I shows the income statement variables Intinc, Intexp, Nonintinc, Netinc, and LLLP are all highly significant. The coefficients on Intinc, Nonintinc, and Netinc are negative indicating that a decrease in any source of income increases the risk of community bank failure. The covariates Intexp and LLP are both positive which means that an increase in interest expenses and expenses set aside for bad loans subsequently increases the rate of bank failure. Similar to conditioning on balance sheet information in Table 5, we find that the covariates, while statistically significant, tend to be economically small suggesting that conditioning community bank failure on income statement covariates is not the most informative approach.

In Model II we again augment the bank-specific variables with the macroeconomic covariate Ted_Spread. However, in this instance we find the addition of the covariate to be statistically insignificant; this result comes in stark contrast to that of Table 5.[45] The other covariates in the model remain unchanged from the specification in Model I. Lastly, in Model III, we remove the covariate Netinc and insert the covariate Incext, as the two variables are very similar income measures, to see which specification performs better overall. The results strongly

[45] Though the sign of Ted_Spread is the opposite of what theory would dictate a 95% confidence interval of the insignificant coefficient broadly ranges from -84.46 to 64.46.

Table 6

Cox Survival Model Conditioned on Income Statement and Macroeconomic Covariates

	Model I Income Statement			Model II Income Statement & Macro			Model III Income Statement & Macro		
Variables	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio
Intinc	-0.000145	0.00	0.999855	-0.000145	0.00	0.999855	-0.000145	0.00	0.999855
Incext							-0.000044	0.00	0.999956
Intexp	0.000162	0.00	1.000162	0.000162	0.00	1.000162	0.000162	0.00	1.000162
Nonintinc	-0.000046	0.00	0.999954	-0.000046	0.00	0.999954	-0.000046	0.00	0.999954
Netinc	-0.000044	0.00	0.999956	-0.000044	0.00	0.999956			
LLLP	0.000022	0.00	1.000022	0.000022	0.00	1.000022	0.000022	0.00	1.000022
Ted_Spread				-1.6341	0.97	0.195136	-1.6537	0.97	0.191342
Likelihood Ratio	603.98	0.00		603.99	0.00		603.96	0.00	
AIC		7,026.10			7,028.09			7,028.12	
BIC		7,060.12			7,068.93			7,068.95	
Log Likelihood		-3,508.04			-3,508.05			-3,508.06	
Banks		6,669			6,669			6,669	
Failures		452			452			452	
Obs. (bank-year)		130,495			130,495			130,495	

Note. The Cox regression models are based on income statement and macroeconomic information. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether $\beta_x=0$, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

mirror those from Model II, in fact the coefficient on Incext (-0.000044) is the same as Netinc in Model II; moreover, both Model II and III provide very similar best-fit measures.

Overall, Table 6 provides useful information about community bank failure conditional on income sources and uses. We find a dichotomous result, relative to Table 5, pertaining to the usefulness of macroeconomic liquidity information. Yet, similar to Table 5, we find economically small covariate coefficients from conditioning community bank failure solely on income statement information suggesting that such an approach is likely not the optimal methodology to determining why community banks fail.

Given the low economic significance of using balance sheet and income statement information as conditioning variables of community bank failure, we appeal to prior literature and examine ratios based on the CAMELS ratings. Tables 7 through 10 utilize the CAMELS financial ratios which prior work has found to be quite relevant for predicting aggregate US commercial bank failure. Table 7 reports the results of conditioning community bank failure on capital adequacy financial ratios. Model I shows that Eq_Asset and Eq_RWA are both negative and highly statistically significant. Moreover, the magnitude of the coefficients suggests a more meaningful economic impact from the covariates than those in Tables 5 and 6. The negative coefficients on Eq_Asset and Eq_RWA align with our expectations and prior findings regarding US commercial banks (see Kolari et al., 2002; Cole and Wu., 2009; Wheelock and Wilson, 2000; Alali and Romero, 2013). Clearly, a bank which has less equity has less protection against unforeseen loan losses and declines in asset values. Accordingly, results show that community banks with higher equity as a percentage of total assets (or risk-weighted adjusted assets) are less likely to fail. Contrasting with prior literature on US commercial banks we find that the Eq_Loanlease ratio is highly insignificant in predicting community bank failure.

Table 7

Cox Survival Model Conditioned on Capital Adequacy Financial Ratios and Macroeconomic Covariates

	Model I Capital Adequacy			Model II Capital Adequacy & Macro			Model III Capital Adequacy & Macro		
Variables	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio
Eq_Asset	-23.1528	0.00	8.81E-11	-26.5347	0.00	2.99E-12	-50.5253	0.00	1.14E-22
Eq_Loanlease	0.000343	0.99	1.000343	0.000394	0.99	1.000394			
Eq_RWA	-17.6261	0.00	2.21E-08	-15.7521	0.00	1.44E-07			
Ted_Spread				88.2394	0.00	2.10E+38	114.1555	0.00	3.78E+49
Likelihood Ratio	2,704.43	0.00		2,719.88	0.00		2,677.95	0.00	
AIC		4,918.66			4,905.22			4,946.29	
BIC		4,939.06			4,932.42			4,959.90	
Log Likelihood		-2,456.33			-2,448.61			-2,471.14	
Banks		6,640			6,640			6,669	
Failures		452			452			452	
Obs. (bank-year)		129,904			129,904			130,517	

Note. The Cox regression models are based on capital adequacy financial ratio and macroeconomic information. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether $\beta_x=0$, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

In Model II, as before, we incorporate the macroeconomic covariate into the specification. Results show a substantial impact from its integration; the (positive) coefficient on Ted_Spread is 88.2394 and it has a P-value better than the 1% significance level. Additionally, all of the best-fit criteria show an improvement by including the macroeconomic factor into the model. This outcome provides solid evidence in favor of macroeconomic liquidity conditions strongly impacting the hazard rate of community bank failure. Lastly, Model III employs only the Eq_Asset covariate and the macroeconomic measure. The majority of prior work generally just incorporates this single bank-specific measure into their multivariate analysis. Our findings in Model III strongly mirror those of prior studies in terms of sign, significance, and magnitude for Eq_Asset. The overall results of Table 7 underlie the importance of capital adequacy (as measured by Eq_Asset and Eq_RWA) in predicting community bank failure. Additionally, the findings support the importance of macroeconomic conditions in the plight of predicting community bank failure.

Table 8 reports the results of conditioning community bank failure on asset quality and liquidity financial ratios. The most interesting part of Model I pertains to the loan-to-asset covariates we analyze. In Model I we include four loan-to-asset ratios: Loanci_Asset, Loancon_Asset, Loanre_Asset, and Loanlease_Asset.[46] We find the coefficients on Loanci_Asset, Loanre_Asset, and Loanlease_Asset are positive, which is consistent with the findings of Wheelock and Wilson (2000). However, we find only the covariate Loanre_Asset is statistically significant, while Wheelock and Wilson (2000) find that the other two covariates, Loanci_Asset and Loanlease_Asset, are significant. Alali and Romero (2013) similarly report the covariate Loanre_Asset to be positive but insignificant. However, they also find that the

[46] We also analyze the agricultural-based ratio Loanag_Asset. In general, we find the covariate to be negative and highly significant. However, because this data is unavailable for many banks we are restricted in its use for reporting results.

covariate Loanlease_Asset is negative and significant which is an interesting contrast to both our results and Wheelock and Wilson (2000). Nonetheless, we interpret the positive significance of the real estate loans to total assets ratio as a reflection of the niche lending market which community banks often fill. Moreover, the fact that an increase in the Loanre_Asset covariate increases failure risk comes in sharp contrast to reports that the percentage of default rates are generally lower for community banks than non-community banks. Yet, the strong ties of the real estate market to the most recent financial crisis combined with the extensive real estate loans smaller institutions held on their balance sheets caused deep financial losses and suffering from this lending segment. Thus, we observe that community banks with higher real estate loans as a percentage of total assets have increased failure risk. Additionally, in line with Ali and Romero (2013) we find Loancon_Asset to be negative and significant, though the coefficient we find (-12.3806) is much larger than what they document (-0.52). This large coefficient magnitude and subsequent hazard rate are a reflection of community banks' strong position in the consumer lending market; hence, we observe that as consumer loans decrease as a percentage of total assets there is an increase in community bank failure risk.

In moving from Model I to Model II we again see, as in Table 7, that the Ted_Spread variable is both positive and highly significant. Moreover, its addition to the model improves the overall fit. This finding, taken together with the prior tables, shows that while bank-specific variables (specifically financial ratios) play a crucial role in determining community bank failure, the contribution of market liquidity as proxied by the TED spread is ultimately very important to whether community banks fail or survive. Model III merely removes the insignificant covariate Loanlease_Asset from the analysis. The new specification improves the fit to the data and renders the covariate Loanci_Asset statistically significant at better than the 5%

Table 8

Cox Survival Model Conditioned on Asset Quality and Liquidity Financial Ratio and Macroeconomic Covariates

	Model I Asset Quality and Liquidity			Model II Asset Quality and Liquidity & Macro			Model III Capital Adequacy & Macro		
Variables	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio
Loanci_Asset	2.2632	0.13	9.6137	2.0856	0.17	8.0491	2.4808	0.04	11.9513
Loancon_Asset	-12.3806	0.00	4.20E-06	-12.0713	0.00	5.72E-06	-11.5350	0.00	9.78E-06
Loanre_Asset	5.3348	0.00	207.4388	5.2723	0.00	194.8677	5.628771	0.00	278.3199
Loanlease_Asset	0.417512	0.81	1.518179	0.725293	0.67	2.0653			
Chargeoff_Asset	4.7343	0.29	113.7821	6.4821	0.15	653.3248	6.2314	0.17	508.4688
LLLP_Asset	-25.3005	0.00	1.03E-11	-25.7622	0.00	6.48E-12	-25.5917	0.00	7.69E-12
LLLP_Loanlease	19.9335	0.00	4.54E+08	20.8230	0.00	1.10E+09	20.8859	0.00	1.18E+09
Chargeoff_Loanlease	9.4587	0.00	12,818.89	9.0360	0.00	8,399.87	9.1278	0.00	9,208.03
Loanlease_Dep	-3.5544	0.00	0.028599	-3.8106	0.00	0.022135	-3.5568	0.00	0.028531
Riskcapt1_RWA	-26.9329	0.00	2.01E-12	-26.7728	0.00	2.36E-12	-26.8533	0.00	2.18E-12
Ted_Spread				65.4359	0.00	2.62E+28	64.9422	0.00	1.60E+28
Likelihood Ratio	3,012.33	0.00		3,022.57	0.00		3,022.39	0.00	
AIC		4,624.77			4,616.53			4,614.70	
BIC		4,692.78			4,691.34			4,682.71	
Log Likelihood		-2,302.38			-2,297.26			-2,297.35	
Banks		6,640			6,640			6,640	
Failures		452			452			452	
Obs. (bank-year)		129,900			129,900			129,900	

Note. The Cox regression models are based on asset quality and liquidity financial ratio and macroeconomic information. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether β_x=0, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

level. A couple of other interesting points worth mentioning concerning Table 8, recent work by Alali and Romero (2013) find that the covariate Chargeoff_Asset is seemingly very useful in explaining US commercial bank failure; however, we find the variable to be largely insignificant in each model.[47] Additionally, they document a positive relationship between Loanlease_Dep, whereas we find a consistently negative one for community banks.

Table 9 reports the results of conditioning community bank failure on earnings financial ratios. Model I shows that Incext_Asset, Netopinc_Asset, Wage_Asset, and ROE are all of the expected negative sign, and highly significant. The coefficient on Wage_Asset has the largest magnitude of -267.8146 while ROE has the smallest at -0.042289. Though we do not explicitly include management ratios in this study we argue that the Wage_Asset variable indirectly proxies for management information in the CAMELS ratings system. Accordingly, we reason that better more efficient managers require higher compensation for their efforts, thus we should see that an increase in the Wage_Asset covariate results in a reduction in bank failure risk. Furthermore, the covariate also encompasses the fact that community banks are heavily reliant on retaining quality employees with strong ties to the community and superior relationship building skills. In order for the small banks to retain such employees they must pay reasonable wages and benefits to avoid high employee turnover and consequential detrimental effects to the institution. As such, a similar effect in the Wage_Asset covariate should also be observed as previously described. The results of Table 9 strongly verify our reasoning—as the wage component increases as a proportion of total assets we see a vast decrease in the likelihood of

[47] They report a coefficient of approximately 647.11 with a hazard ratio of 1.09E+13, and a P-value of 0.01.

Table 9

Cox Survival Model Conditioned on Earnings Financial Ratio and Macroeconomic Covariates

	Model I Earnings			Model II Earnings & Macro			Model III Earnings & Macro		
Variables	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio
Incext_Asset	-4.8288	0.00	0.007997	-4.8316	0.00	0.007974			
Netopinc_Asset	-10.5121	0.00	0.000027	-10.4892	0.00	0.000028			
Wage_Asset	-267.8146	0.00	4.90E-117	-267.8107	0.00	4.90E-117	-236.9618	0.00	1.20E-103
ROA							-15.1905	0.00	2.53E-07
ROE	-0.042289	0.00	0.958592	-0.042451	0.00	0.958437	-0.033989	0.00	0.966582
Ted_Spread				14.1004	0.61	1,329,615	0.042538	0.88	1.043456
Likelihood Ratio	1,332.97	0.00		1,333.22	0.00		1,244.59	0.00	
AIC		6,295.27			6,297.02			6,383.65	
BIC		6,322.49			6,331.04			6,410.87	
Log Likelihood		-3,143.64			-3,143.51			-3,187.83	
Banks		6,669			6,669			6,669	
Failures		452			452			452	
Obs. (bank-year)		130,517			130,517			130,517	

Note. The Cox regression models are based on earnings financial ratio and macroeconomic information. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether β_x=0, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

community bank failure. Given this line of thought we underline the importance of quality management and employees in the community banking industry.[48]

The addition of the macroeconomic liquidity factor in Model II yields a positive but insignificant Ted_Spread. Interestingly, this is the same result we found when evaluating income statement information in Table 6 and the earnings ratios at hand are similarly based on income statement information. The AIC, BIC, and log likelihood criterion show a dramatic drop off in the fit of the model when compared to Tables 7 and 8; this is the same pattern witnessed between Tables 5 and 6 when comparing balance sheet and income statement information, respectively, as well. Overall, it seems that community bank failure predicated on income statement information, even when evaluated as financial ratios, provides less pertinent information to predicting bank failure relative to balance sheet data. Lastly, Model III incorporates the ROA covariate and removes the covariates Netopinc_Asset and Incext_Asset due to high correlations between the covariates. In doing so we find that both Wage_Asset and ROE remain largely unchanged and Ted_Spread still remains statistically insignificant. The new variable ROA is negative and highly significant, which is not too surprising considering its relation to Netopinc_Asset. However, the overall fit of Model III is markedly worse than that of the other two specifications.

Given the relative importance of the CAMELS financial ratios we utilize the covariates analyzed in Tables 7 through 9 to form a single aggregate model. Table 10 reports the regression results of conditioning community bank failure on the aggregate financial ratios. In Model I of the aggregate analysis we find that the covariate Eq_Asset is still the most valuable capital adequacy measure in predicting community bank failure. Moreover, we still see that an increase

[48] Alali and Romero (2013) document a positive but insignificant coefficient for their salary and wages/total assets earnings ratio.

in the ratios Loanci_Asset and Loanre_Asset result in an increase in community bank failure risk. The large negative coefficient on Loancon_Asset once again highlights the important unique relationship between consumer lending and community banks. Reviewing the asset quality and liquidity ratios we find that the covariate Chargeoff_Asset becomes statistically significant in the aggregate regression; however, its importance in the community banking industry is still dwarfed by that of US commercial banks. The importance of the covariate LLLP_Asset really emerges in Table 10, and at first glance the result seems somewhat counterintuitive. Given that loan loss provisions are expenses set aside for bad debt, it seems that a decrease in this variable relative to total assets should decrease the risk of bank failure. However, loan loss provisions are quite pro-cyclical and subject to considerable managerial earnings management. As such, the provisions for loan and lease losses are generally lower during good economic times; yet, during subsequent economic downturns, as the value of assets decline and more capital is needed to buffer against potential future losses, capital is at its most expensive and becomes largely unattainable for smaller banks altogether—forcing a "credit crunch" and a risk in bank failure risk. Accordingly, the LLLP_Asset variable is capturing the strong cyclicality of loan loss provisioning in the community banking industry. In other words, those community banks which don't build adequate capital buffers prior to an economic downturn have a much higher risk of failure. Lastly, the earnings ratios show that income before extraordinary items relative to total assets (Incext_Asset) is the best income-based measure to capture the predictability of community bank failure. Moreover, the traditional measure of bank ROE, while significant, is economically very small after controlling for other available covariates. Our indirect proxy of bank management, Wage_Asset, still has the largest coefficient of all ratios, and as such we again stress the importance of community banks in possessing effective management teams to remain financially healthy and thriving.

Model II again incorporates the Ted_Spread covariate into our analysis and, as in the majority of prior cases, shows an important contribution to the community bank failure specification. The covariate is again positive and highly significant; moreover, the coefficient (97.1341) still possesses a momentous economic impact. Additionally, the best-fit statistics show a sizable improvement in the model fit with the addition of the macroeconomic covariate. Overall, we conclude that US macroeconomic liquidity conditions play a vitally important role in the failure and survival rates of the community bank industry. Finally, in Model III we provide an alternative specification in which we supplement Model II with several more potentially relevant covariates. We include the covariates Eq_RWA, Loanlease_Asset, and Netint_Asset (which was found to be relatively unimportant in Table 9 and was consequently left out of the analysis) into the model. The addition of the covariates provides only a slight improvement in the models overall fit; the variable Eq_RWA is significant at better than the 5% level, though Eq_Asset losses its significance at the 1% level suggesting both variables are reasonably similar proxies. Loanlease_Asset is still statistically insignificant and actually causes Loancon_Asset to become insignificant with its addition. The most interesting result is the addition of the net interest margin covariate (Netint_Asset) which is both statistically and economically significant. This supplementary finding places additional emphasis on the need to monitor community bank investment and debt decisions.

Table 10

Cox Survival Model Conditioned on Aggregate Financial Ratio and Macroeconomic Covariates

	Model I Aggregate			Model II Aggregate & Macro			Model III Aggregate & Macro		
Variables	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio	Coef.	P-value	Hazard Ratio
Eq_Asset	-19.7574	0.00	2.63E-09	-22.91245	0.00	1.12E-10	-11.7910	0.03	7.57E-06
Eq_RWA							-10.3821	0.02	0.000031
Loanci_Asset	3.3866	0.00	29.5639	3.4317	0.00	30.9296	3.2183	0.05	24.9855
Loancon_Asset	-7.5150	0.00	0.000545	-6.7472	0.01	0.001174	-4.5029	0.11	0.011077
Loanre_Asset	4.6877	0.00	108.5991	4.8761	0.00	131.1229	4.2903	0.00	72.9873
Loanlease_Asset							1.3391	0.45	3.8154
Chargeoff_Asset	10.8211	0.04	50,063.70	12.7626	0.02	348,935.90	9.4555	0.08	12,778.60
LLLP_Asset	-31.8804	0.00	1.43E-14	-32.3875	0.00	8.60E-15	-32.5584	0.00	7.25E-15
LLLP_Loanlease	12.6740	0.00	319,325.80	13.0029	0.00	4.44E+05	14.0180	0.00	1.22E+06
Chargeoff_Loanlease	3.9003	0.00	49.4157	3.4894	0.06	32.7677	4.0008	0.04	54.6425
Loanlease_Dep	-2.5013	0.04	0.081979	-2.6768	0.00	0.068780	-3.1483	0.00	0.042926
Riskcapt1_RWA	-13.8300	0.00	0.000001	-12.8127	0.00	0.000003	-8.7675	0.00	0.000156
Incext_Asset	-7.2707	0.01	0.000696	-7.0849	0.01	0.000838	-9.6459	0.00	0.000065
Netopinc_Asset	0.5995	0.75	1.821137	1.6549	0.38	5.2328	3.0405	0.13	20.9167
Netint_Asset							-23.9150	0.00	4.11E-11
Wage_Asset	-176.8089	0.00	1.63E-77	-176.3705	0.00	2.53E-77	-157.2648	0.00	5.02E-69
ROE	-0.025705	0.00	0.974623	-0.027222	0.00	0.973145	-0.028236	0.00	0.972159
Ted_Spread				97.1341		1.53E+42	70.8544	0.00	5.91E+30
Likelihood Ratio	3,376.01	0.00		3,398.96	0.00		3,429.89	0.00	
AIC		4,269.08			4,248.13			4,223.21	
BIC		4,364.29			4,350.15			4,345.62	
Log Likelihood		-2,120.54			-2,109.07			-2,093.60	
Banks		6,640			6,640			6,640	
Failures		452			452			452	
Obs. (bank-year)		129,900			129,900			129,900	

Note. The Cox regression models are based on aggregate financial ratio and macroeconomic information. For each full set of M bank-specific and macroeconomic variables we compare the M regressors and M-1 regressor models and retain the appropriate model based on the following: (1) the significance of the covariates based on the P-values, (2) the likelihood ratio which tests whether β_x=0, (3) the Akaike Information Criteria (AIC), (4) the degrees-of-freedom-adjusted Bayesian Information Criterion (BIC), and (5) the log likelihood ratio best-fit criterion.

5. CONCLUDING REMARKS

The last 40 years have seen the community banking sector markedly shrink. In particular, extensive changes in regulation and industry practices over the last 20 years have played a critical role in the downsizing of community banks in the US financial system. Moreover, the far-reaching roots of the 2008 global financial crisis had severe consequences for the banking industry as a whole. Yet, in spite of the institutional decline community banks continue to play a vital role in the US economy. Due to their economic and financial importance this paper conducts an exploratory investigation, utilizing survival analysis, to determine which accounting variables aid in predicting community bank failure. Prior work largely focuses on aggregate US commercial banks; however, we feel that the indicators and early warning signs of community bank failure should recognize the distinct differences and risk profiles of community and non-community banks.

We utilize a broad set of FDIC bank-specific accounting data over the period 1992-2013 to examine the characteristics of community banks which failed between the years 2000-2013. We incorporate information from balance sheets, income statements, and financial ratios based on the CAMELS ratings. We also incorporate a well-known US market liquidity component into our models to try and link how macroeconomic liquidity shocks, proxied by the TED spread, contribute to community bank failure. Overall, we utilize 452 failed community banks consisting of 6,350 bank-year observations and 6,217 non-failed community banks consisting of 124,167 bank-year observations, for a total sample of 6,669 community banks and 130,517 bank-year observations.

Our empirical results indicate that ordinary balance sheet and income statement information are a relatively ineffective way to predict community bank failure, notwithstanding balance sheet information offers an informational edge over income statement information and highlights that smaller banks (based on total assets) are actually less likely to fail than their larger community bank counterparts. Financial ratio information, in particular ratios which encompass capital adequacy, asset quality and liquidity, and earnings, provide a dramatic increase in the ability to predict community bank failure risk; this finding is in line with prior literature which examines firm and commercial bank failures. We find that community banks which have declining amounts of equity and loan loss provisions as a percentage of total assets (i.e. Eq_Asset and LLLP_Asset) have substantially increased failure rates. We also find that banks with more commercial and industrial loans as well as real estate loans relative to total assets (i.e. Loanci_Asset and Loanre_Asset) have increased failure rates. Interestingly, community banks which reduce their proportion of consumer lending as a percentage of total assets (i.e. Loancon_Asset) are more likely to fail—emphasizing the importance of community banks and their smaller, rural lending practices. Income-based ratios such as net operating income to total assets and return on equity (Netint_Asset and ROE) seem to have less relative importance in predicting community bank failure. The most fascinating result pertains to the salary and wage expenses to total assets ratio (i.e. Wage_Asset) which we argue indirectly proxies as a manager effectiveness/efficiency and quality employee retention measure. The covariate is both highly statistically and economically significant. Provided that "better" managers and employees require increased compensation for their efforts, we find that a reduction in salary and wages as a proportion of total assets results in a dramatic increase in community bank failure risk. Thus, managerial effectiveness and efficiency in addition to employee quality seems to be a vital part of financially healthy community banks. Of particular noteworthiness, we also document that the use of a macroeconomic indicator of liquidity provides a substantial improvement in modeling predictive community bank failure. Prior studies have indicated that macroeconomic variables play a secondary role to firm- or bank-specific characteristics, and while we fundamentally agree with this assessment, our results provide further impetus to explore the impact of liquidity on the community banking sector. Specifically, we document that as macroeconomic liquidity conditions deteriorate (i.e. as the TED spread widens) there is a considerable rise in the failure rate of community banks, especially for institutions which are already experiencing financial difficulties.

The results of this paper have important practical implications at an institutional and governance level for protecting community banks, such as serving as early warning signals of impending problems. Furthermore, our analysis serves as an important cue that more research in the area of community banks is vital to the enhancement of the industry; we provide motivation for future research to continue to investigate the importance of macroeconomic effects on the US community banking industry.

REFERENCES

Aalen, O.O., 1978. Nonparametric inference for a family of counting processes. *Annals of Statistics* 6(4), 701-726.

Alali, F. and Romero, S., 2013. Characteristics of failed U.S. commercial banks: an exploratory study. *Journal of Accounting and Finance* 53(4), 1149-1174.

Arena, M., 2008. Bank failures and bank fundamentals: A comparative analysis of Latin America and East Asia during the nineties using bank-level data. *Journal of Banking and Finance* 32(2), 299-310.

Beaver, W.H., McNichols, M.F., and Rhie, J., 2005. Have financial statements become less informative? Evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies* 10(1), 93-122.

Bell, T.B., 1997. Neural nets or the logit model? A comparison of each model's ability to predict commercial bank failures. *Intelligent Systems in Accounting, Finance and Management* 6(3), 249-264.

Berger, A.N. and Udell, G.F., 2001. Small Business Credit Availability and Relationship Lending: The Importance of Bank Organizational Structure. *FEDS Working Paper No. 2001-36*.

Bharath, S.T., and Shumway, T., 2008. Forecasting default with the Merton distance to default model. *Review of Financial Studies* 21(3), 1339-1369.

Board of Governors of the Federal Reserve System, 2014. Economic Research and Data, http://www.federalreserve.gov/econresdata/default.htm.

Bonfim, D., 2009. Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking and Finance* 33(2), 281-299.

Box-Steffensmeier, J. and Jones, B.S., 2004. *Event History Modeling: A Guide for Social Scientists* (Cambridge University Press, Cambridge, ENG).

Brewer, E., Jackson, J., and Moser, J., 1996. Alligators in the Swamp: The impact of derivatives on the financial performance of depository institutions. *Journal of Money, Credit and Banking* 28(3), 478-496.

Campbell, J., Hilscher, J., and Szilagyi, J., 2008. In search of distress risk. *The Journal of Finance* 63(6), 2899-2939.

Carling, K., Jacobson, T., Linde, J., and Roszbach, K., 2007. Corporate credit risk modeling and the macroeconomy. *Journal of Banking and Finance* 31(3), 845-868.

Chava, S. and Jarrow, R.A., 2004. Bankruptcy prediction with industry effects. *Review of Finance* 8(4), 537-569.

Cleves, M., Gutierrez, R.G., Gould, W., and Marchenko, Y.V., 2010. *An Introduction to Survival Analysis Using Stata* (Stata Press, College Station, TX).

Cole, R.A. and Gunther, J.W., 1995. A CAMEL rating's shelf life. Federal Reserve Bank of Dallas, *Financial Industry Studies*, 13-20.

Cole, R.A. and Gunther, J.W., 1995. Separating the Likelihood and Timing of Bank Failure. *Journal of Banking and Finance* 19(6), 1073-1089.

Cole, R.A. and Wu, Q., 2009. Predicting bank failures using a simple dynamic hazard model. *In 22nd Australian Finance and Banking Conference*.

Cox, D.R., 1972. Regression Models and Life-Tables. *Journal of the Royal Statistical Society Series B (Methodological)* 34(2), 187-220.

Critchfield, T., Davis, T., Davison, L., Gratton, H., Hanc, G., and Samolyk, K., 2004. Community Banks: Their Recent Past, Current Performance, and Future Prospects. *FDIC Banking Review* 16(3), 1-56.

Curry, T.J., Elmer, P.J., and Fissel, G.S., 2007. Equity market data, bank failures and market efficiency. *Journal of Economics and Business* 59(6), 536-559.

Demirguc-Kunt, A., 1989. Modeling large commercial-bank failures: a simultaneous equations analysis. *Federal Reserve Bank of Cleveland*.

DeYoung, R., Hunter, W.C., and Udell, G.F., 2004. The Past, Present and Probable Future of Community Banks. *Journal of Financial Services Research* 25(2/3), 85-133.

Federal Deposit Insurance Corporation, 2012. FDIC Community Banking Study, www.fdic.gov/regulations/resources/cbi/study.html.

Federal Reserve Bank of Kansas City, 2003. The Role of Community Banks in the U.S. Economy. *Economic Review*, http://www.kc.frb.org/.

Francis, J. and Schipper, K., 1999. Have financial statements lost their relevance? *Journal of Accounting Research* 37(2), 319-352.

Fraser, D., 1976. The determinants of bank profits: An analysis of extremes. *Financial Review* 11(1), 69-87.

Gady, R., 1972. Anatomy of profitable medium-size banks in the fourth district. Federal Reserve Banks of Cleveland, *Economic Review*, 20-32.

Gomez-Gonzalez, J.E. and Kiefer, N.M., 2009. Bank Failure: Evidence from the Colombian Financial Crisis. *The International Journal of Business and Finance Research* 3(2), 15-31.

Hassan, M.K. and Hippler, W., 2014. National and Regional Trends in Community Banking. *Working Paper*.

Hein, S.E., Koch, T.W., and MacDonald, S.S., 2005. On the Uniqueness of Community Banks. Federal Reserve Bank of Atlanta, *Economic Review*, 15-36.

Hosmer Jr., D.W., Lemeshow, S., and May, S., 2008. *Applied Survival Analysis: Regression Modeling of Time to Event Data* (Wiley, New York, NY).

Kalbfleisch, J.D. and Prentice, R.L., 2002. *The Statistical Analysis of Failure Time Data* (Wiley, New York, NY).

Kaplan, E.L. and Meier, P., 1958. Nonparametric Estimation from Incomplete Observations. *Journal of the American Statistical Association* 53(282), 457-481.

Kaplan, T. and Camelia, M., 2013. Balance Sheet Strength and Bank Lending Duing the Global Financial Crisis. *IMF Working Paper 13-102*.

Kolari, J., Glennon, D., Shin, H., and Caputo, M., 2002. Predicting large US commercial bank failures. *Journal of Economics and Business* 54(4), 361-387.

Lane, W.R., Looney, S.W., and Wansley, J.W., 1986. An Application of the Cox Proportional Hazards Model to Bank Failure. *Journal of Banking and Finance* 10(4), 511-531.

Lux, M. and Greene, R., 2015. The State and Fate of Community Banking. Harvard Kennedy School, *M-RCBG Associate Working Paper Series No. 37*.

Mannasoo, K.A. and Mayes, D.G., 2009. Explaining Bank Distress in Eastern European Transition Economies. *Journal of Banking and Finance* 33(2), 244-253.

Molina, C.A., 2002. Predicting bank failures using a hazard model: the Venezuelan banking crisis. *Emerging Markets Review* 3(1), 31-50.

Nelson, W., 1972. Theory and applications of hazard plotting for censored failure data. *Technometrics* 14(4), 945-966.

Pettway, R. and Sinkey, J., 1980. Establishing On-Site Bank Examination Priorities: An Early-Warning System Using Accounting and Market Information. *The Journal of Finance* 35(1), 137-150.

Ravi Kumar, P. and Ravi, V., 2007. Bankruptcy prediction in banks and firms via statistical and intelligent techniques—A review. *European Journal of Operational Research* 180(1), 1-28.

Sales, A.S. and Tannuri-Pianto, M.E., 2007. Explaining Bank Failures in Brazil: Micro, Macro and Contagion Effects (1994-1998). Central Bank of Brazil, *Working Paper 147*.

Shumway, T., 2001. Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business* 74(1), 101-124.

Swicegood, P. and Clark, J.A., 2001. Off-site monitoring systems for predicting bank underperformance: a comparison of neural networks, discriminant analysis, and professional human judgment. *International Journal of Intelligent Systems in Accounting, Finance and Management* 10(3), 169-186.

Thomson, J.B., 1992. Modeling the regulator's closure option: A two-step logit regression approach. *Journal of Financial Services Research* 6(1), 5-23.

Wall, L., 1985. Why are some banks more profitable than others? *Journal of Bank Research* 15(4), 240-256.

Whalen, G., 1991. A proportional hazards model of bank failure: an examination of its usefulness as an early warning model tool. *Economic Review* 27(1), 21-30.

Wheelock, D.C. and Wilson, P.W., 2000. Why do banks disappear? The determinants of U.S. bank failures and acquisitions. *Review of Economics and Statistics* 81(1), 127-138.

APPENDIX A

Table A1

Best Fit Copulas based on Log-likelihood

Full Sample Period (Oct. 1992 - Oct. 2013)	Energy	Foods & Fibers	Grains & Oilseeds	Livestock	Precious Metals
Normal Copula	-23.7330	-9.2894	-19.7170	-4.9685	-19.0050
Student's t Copula	-27.9620	-11.0520	-20.7880	-3.7519	-20.3250
Rotated-Gumbel Copula	-34.7900	-15.3100	-22.7990	-5.7289	-24.5840

Note. This table provides the best fit measure for the copula functions based on the log-likelihood criteria, for each sub-sector, over the full sample period.

APPENDIX B

Table B1

Correlation Matrix of Balance Sheet and Macroeconomic Conditioning Variables

	Asset	Liab	Eqtot	Riskcapt1	Loanci	Loancon	Loanre	Loanag	Totalloan	Ted_ Spread
Asset	1.0000									
Liab	0.8651	1.0000								
Eqtot	0.8798	0.7644	1.0000							
Riskcapt1	0.6959	0.9052	0.7702	1.0000						
Loanci	0.5951	0.6865	0.5117	0.6068	1.0000					
Loancon	0.3227	0.2783	0.2919	0.2273	0.1733	1.0000				
Loanre	0.7754	0.9374	0.6860	0.8692	0.5661	0.1404	1.0000			
Loanag	0.1769	0.1785	0.1411	0.1391	0.1626	0.0701	0.0578	1.0000		
Loanlease	0.9576	0.8357	0.8113	0.6585	0.6201	0.3241	0.8184	0.1828	1.0000	
Ted_Spread	-0.0123	-0.0105	-0.0146	-0.0115	0.0076	0.0093	0.0134	-0.0072	0.0151	1.0000

Note. Pairwise correlations based on balance sheet and macroeconomic conditioning variables.

Table B2

Correlation Matrix of Income Statement and Macroeconomic Conditioning Variables

	Intinc	Incext	Intexp	Nonintinc	Netinc	Netopinc	LLLP	Ted_ Spread
Intinc	1.0000							
Incext	0.3020	1.0000						
Intexp	0.8448	0.1343	1.0000					
Nonintinc	0.1802	0.5456	0.1011	1.0000				
Netinc	0.3018	0.9973	0.1346	0.5443	1.0000			
Netopinc	0.2995	0.9899	0.1338	0.5499	0.9896	1.0000		
LLLP	0.4514	-0.1772	0.3410	0.1029	-0.2028	-0.2046	1.0000	
Ted_Spread	0.0670	-0.0408	0.1996	-0.0077	-0.0410	-0.0301	0.0220	1.0000

Note. Pairwise correlations based on income statement and macroeconomic conditioning variables.

Table B3

Correlation Matrix of Aggregate Financial Ratio and Macroeconomic Conditioning Variables

	Eq_Asset	Eq_ Loanlease	Eq_RWA	Loanci_ Asset	Loancon_ Asset	Loanre_ Asset	Loanag_ Asset	Loanlease_ Asset	Lossallow_ Asset	Chargeoff_ Asset	LLLP_ Asset	LLLP_ Loanlease	Chargeoff_ Loanlease	Loanlease_ Dep	Riskcapt1_ RWA	Incext_ Asset	Netopinc_ Asset	Netint_ Asset	Wage_ Asset	ROA	ROE	Ted_ Spread
Eq_Asset	1.0000																					
Eq_Loanlease	0.1512	1.0000																				
Eq_RWA	0.0174	0.2201	1.0000																			
Loanci_Asset	0.7888	-0.0172	-0.0025	1.0000																		
Loancon_ Asset	0.1848	-0.0132	-0.0294	0.1841	1.0000																	
Loanre_Asset	0.8161	-0.0322	-0.0024	0.8080	0.1928	1.0000																
Loanag_Asset	-0.0264	-0.0094	-0.0187	-0.0296	-0.0164	-0.4487	1.0000															
Loanlease_ Asset	0.8911	-0.0562	-0.0362	0.8049	0.1906	0.9485	0.0077	1.0000														
Lossallow_ Asset	0.8162	-0.0210	-0.0021	0.8077	0.1941	0.8455	0.0481	0.7973	1.0000													
Chargeoff_ Asset	0.2867	-0.0030	-0.0033	0.4885	0.1041	0.2451	-0.0412	0.2971	0.2443	1.0000												
LLLP_Asset	0.7936	-0.0028	-0.0022	0.8630	0.1967	0.8334	-0.0416	0.7950	0.9833	0.2911	1.0000											
LLLP_ Loanlease	-0.2103	0.1436	0.0243	-0.0263	0.0018	-0.1569	0.0369	-0.1158	-0.1573	0.0091	-0.1274	1.0000										
Chargeoff_ Loanlease	0.0101	-0.0055	0.0040	0.0138	0.0170	0.0115	-0.0712	0.0274	0.0111	0.0302	0.0121	0.0036	1.0000									
Loanlease_ Dep	0.0396	-0.0028	0.0004	0.0116	0.0193	0.0406	-0.0054	0.0616	0.0212	0.0101	0.0123	-0.0094	0.0041	1.0000								
Riskcapt1_ RWA	0.0179	0.1591	0.9979	-0.0015	-0.0278	-0.0013	-0.0179	-0.0344	-0.0010	-0.0029	-0.0011	0.0235	0.0036	0.0005	1.0000							
Incext_Asset	0.7311	0.0366	-0.0034	0.8072	0.1605	0.7054	0.0164	0.6636	0.7055	0.3924	0.7514	0.0636	0.0083	0.0109	-0.0027	1.0000						
Netopinc_ Asset	0.0067	0.0365	0.0226	0.0055	0.0720	0.0040	0.0747	-0.0082	0.0043	-0.0226	0.0035	-0.0108	-0.0435	-0.0099	0.0211	0.1520	1.0000					
Netint_Asset	-0.0217	-0.0120	-0.0514	-0.0085	0.3308	-0.0167	0.0493	0.0305	-0.0164	0.0136	-0.0144	0.0161	0.0113	0.0132	-0.0506	0.0082	0.1268	1.0000				
Wage_Asset	0.1319	0.0840	-0.0007	0.8252	0.1873	-0.0297	-0.0224	-0.0042	0.0401	0.2527	0.6867	-0.1880	0.0102	-0.0005	0.0003	0.6986	0.0073	-0.0150	1.0000			
ROA	0.0006	0.0365	0.0226	-0.0006	0.0682	-0.0021	0.0714	-0.0150	-0.0018	-0.0243	-0.0036	0.0005	-0.0438	-0.0102	0.0214	0.1492	0.9420	0.1246	0.0000	1.0000		
ROE	-0.0005	0.0005	0.0001	-0.0009	0.0286	-0.0011	0.0218	-0.0033	-0.0010	-0.0150	-0.0015	-0.0072	-0.0362	-0.0011	0.0003	0.0158	0.1503	0.0512	-0.0007	0.1565	1.0000	
Ted_Spread	0.0009	-0.0032	-0.0018	0.0000	0.0269	-0.0005	0.0101	0.0194	-0.0007	0.0008	0.0000	-0.0035	0.0104	0.0061	-0.0018	0.0000	-0.0168	-0.0032	-0.0002	-0.0247	-0.0009	1.0000

Note. Pairwise correlations based on aggregate financial ratios and macroeconomic conditioning variables.

Vita

The author was born in Bend, Oregon in May, 1986. He received his Bachelor of Arts degree in Business Administration: Finance and Economics from George Fox University (GFU) in April, 2008. He joined the University of New Orleans (UNO) Financial Economics graduate program to pursue his Ph.D. in August, 2011. He received his Master of Science in Financial Economics from UNO in May, 2013 and became an active teaching assistant at the university. He obtained his Ph.D. in Financial Economics from UNO in May, 2015.