Abstract

We study aggregates of corporate value using a method largely unexplored in the literature: valuation multiples. Possibilities illuminated by the valuation multiple perspective are examined using conventional and self-developed econometric methods, aiming to understand and predict the relationship between macroeconomic factors and multiples. The methods are applied to a data set of ~2,000 listed European firms. Volatility, inflation, term structure spread and industrial production growth emerge as principal variables and are used for modeling multiple aggregates and cross-sectional multiple distributions. The models unanimously forecast a future decline of multiple levels, as the current levels are unjustifiable by the prevailing macroeconomic environment.

Samandrag

Me analyserer aggregert prising av verksemder med hjelp av ein lite nytta framgangsmåte mellom forskarar innanfor finans: verdsetjingsmultiplar. Multippelvinklinga gjev høve til å skjøna og å føreseia samanhengen mellom makroøkonomiske faktorar og verdsetjing. Me brukar konvensjonelle og eigenutvikla økonometriske metodar på eit dataset over om lag 2000 europeiske ålmennaksjeselskap. Volatilitet, inflasjon, differansen mellom lange og korte renter og industriell tilverking synest å vera vektige forklåringsvariablar, både for indeksar avleidde frå fordelinga og for fordelinga i seg sjølv. Me greier ikkje å rettferdiggjere det noverande multippelnivået gjeve den makroøkonomiske stoda i dag, og går difor ut ifrå at verksamdsprisinga vil gå attende.

Preface

This combined project and master thesis is written as the concluding part of the Industrial Economics and Technology Management Master's programme at the Norwegian University of Science and Technology (NTNU) with specialization in Financial Engineering. The combined thesis is written as a coherent document. However, if a separation of the documents is required, Chapter 4 forms the project thesis, and Chapter 5 and 6 are exclusive for the master thesis, while Chapter 1, 2, 3 and 7 encapsulate the combined thesis and are relevant for all chapters.

We would like to thank our supervisor, Associate Professor at NTNU and OsloMet, Einar Belsom, for his extensive feedback and guidance during the whole process. Einar was highly engaged and flexible and asked relevant, and much appreciated questions. The thesis has also been written in collaboration with FSN Capital, and we appreciate the discussions and the data access we have gained through them. As we have a non-disclosure agreement with FSN Capital, the thesis will be withheld from publication for three years. No programming code will be attached, but the analyses are documented in the main text and in an extensive appendix. The thesis was written at campus in Trondheim and at Rystad Energy's office in Oslo to whom we are grateful.

During the final weeks working with the thesis, a substantial correction occurred in the financial markets, but this is unseen to our data and has not affected our conclusions.

Oslo, October 24, 2018

Gard Rystad | Carl Gustaf Lenngren

Table of Contents

Abbreviations

Symbol	=	*Definition*
Abbreviations		
Capex	=	Capital expenditures
CoGS	=	Costs of Goods Sold
D&A	=	Depreciation and Amortization
DCF	=	Discounted Cash Flow
EV	=	Enterprise Value
EBIT	=	Earnings Before Interest and Tax
EBITDA	=	Earnings Before Interest, Tax, Depreciation and Amortization
FCF	=	Free Cash Flow
MAD	=	Mean Absolute Deviation
MCap	=	Market Capitalization
NOPLAT	=	Net Operating Profit Less Adjusted Tax
NPVGO	=	Net Present Value of Growth Opportunities
Opex	=	Operational expenditures
RMSE	=	Root Mean Square Error
SG&A	=	Sales, General and Administrative costs
Models and Variables		
ARIMAX(p,q,d)	=	AutoRegressive(p-lags) Integrated(d differentiations) Moving Average(d-lags) with eXogenous variabels
vST50(lr)(BCt)	=	Implied volatility
mEUCPI(BCt)	=	Eurozone CPI inflation
iEUspd101	=	Term structure spread
mEUpYOY	=	Year on year EU industrial production growth
sEUCC(d)	=	EU Condumer Confidence Index
mEUFAYOY	=	EU Fixed Assets growth year on year
Ad	=	Regression model with I(1)-variables selected from the Diverse data set
Ap	=	Regression model with I(1)-variables selected from the Predictable data set
Dd	=	Regression model with I(0)-variables selected from the Diverse data set
Dp	=	Regression model with I(0)-variables selected from the Predictable data set
Rd	=	Autoregressive regression model with I(0) and I(1)-variables selected from the Diverse data set
Rp	=	Autoregressive regression model with I(0) and I(1)-variables selected from the Predictable data set
qPCA	=	Quantile PCA-model
qOLS	=	Quantile OLS-model
qVAR	=	Quantile VAR-model
Parametric regression	=	Parametric distribution-based regression
Frequency regression	=	Frequency based distribution regression model

Models are generally denoted in italic. Note that this list is non-exclusive and serves as a overview.

Mathematical Symbols		
$\mathbb{1}_w$	=	Indicator function - is 1 if w is true, 0 else
$\forall$	=	The universial quantifier
b	=	Index of a "bucket"
B	=	Number of buckets
B_t	=	Book equity at time t
c	=	The index of a specific company
C	=	The number of companies
d_t	=	Differentiated (log return) of price index y_t at time t
D_{ct}	=	Dividend for company c at time t
$\mathscr{D}_t^q$	=	log-difference between two quantiles
e, exp()	=	Euler's number
$\exists$	=	Existential quantifier
g	=	Dividend growth rate
h	=	Forecast horizon index
H	=	Maximum forecast horizon
$H(\boldsymbol{x}_t)$	=	Function mapping from $\boldsymbol{x}_t$ to a distribution of multiples m
i	=	Inflation rate
$I(i)$	=	Integrated process of order i - see Alexander (2008a)
$L(\boldsymbol{\theta};\boldsymbol{x}_{1:T})$	=	Likelihood function for parameters $\boldsymbol{\theta}$ given exogenous variables $\boldsymbol{x}_{1:T}$
$\mathscr{L}(\boldsymbol{\theta};\boldsymbol{x}_{1:T})$	=	$\log L(\boldsymbol{\theta};\boldsymbol{x}_{1:T})$
m_{ct}	=	Valuation multiple measure of company c at time t ($\frac{V_{ct}}{\pi_{ct}}$)
M_t	=	Stochastic variable from which multiples are drawn
$\mathscr{N}(\mu,\sigma)$	=	Normally distributed with mean μ and standard deviation σ
r_{ct}	=	Expected rate of return / discount rate at time t for company c
r_c	=	Expected constant rate of return / discount rate for company c
r_f	=	Risk free rate of return
R	=	Matrix of coefficients
s	=	Simulation index
S	=	Number of simulation
t	=	Time in months index
T	=	Time series length (223)
$\mathscr{T}(\mu,\sigma,\nu)$	=	Student t-distributed with parameters μ, σ, ν
V_{ct}	=	Value measure of company c at time t
V_t	=	Stochastic variable from which valuations are drawn at time t
$\boldsymbol{x}_t$	=	Independent variables at time t
y_t	=	Valuation index at time t
y_{qt}	=	Quantile index at time t
α	=	Regression constant/intercept
$\boldsymbol{\beta}$	=	Regression coefficients
γ	=	Autoregressive coefficient
$\Gamma(x)$	=	Legendre's notation of the Gamma function (Davis 1959)
Δ_b	=	The size of bucket b
$\boldsymbol{\Delta x}_t$	=	Differentiated independent variables at time t
ϵ_t	=	Residual process at time t
η_t	=	Moving Average error term at time t
$\Theta(n)$, $\Omega(n)$, $O(n)$	=	Big-O notation for algorithm space and time complexity
Λ	=	Loss functions
μ	=	Distribution location parameter
ν_{bt}	=	Frequency of bucket b at time t
ν	=	Distribution degrees of freedom parameter
π_{ct}	=	Profit measure of company c at time t
π_t	=	Stochastic variable from which profits are drawn at time t
ρ	=	Pearson correlation coefficient
$\boldsymbol{\rho}_t$	=	Distribution parametes at time t
σ	=	Standard deviation, volatility, scale parameter
Σ	=	Covariance matrix
Ψ_l	=	Autoregressive matrix for lag l
$\omega \in \Omega_t$	=	Valuation multiple observation
Ω_t	=	Set of valuation multiples at time t
$\aleph$	=	Aggregation function (median, average, quantiles etc.)
$\times$	=	Cartesian product

Mathematical notation is introduced whenever used, and this list is intended to clarify some conventions. Estimators are denoted using hat, vectors in bold, matrices in capital letters, and indices are used in subscript and superscript. All logarithms are with base e, and set notation from Russell & Norvig (2003) is used: $x_{1:T} = \{x_t | t \in \{1,...,T\}\}$.

Chapter 1

Introduction

The European stock market is currently at all-time high and has more than doubled since 2009[1]. This kind of stock market development is unsurprising, and the essential question becomes: does the valuation reflect underlying value creation prospects? Current earnings is a gauge of value creation, as the S&P500 Price/Earnings index in Figure 1.1 acknowledges. However, in order to effectively unwind the relationship between value and value creation, one must comprehend the system of underlying drivers that affects both factors: The macroeconomic environment.

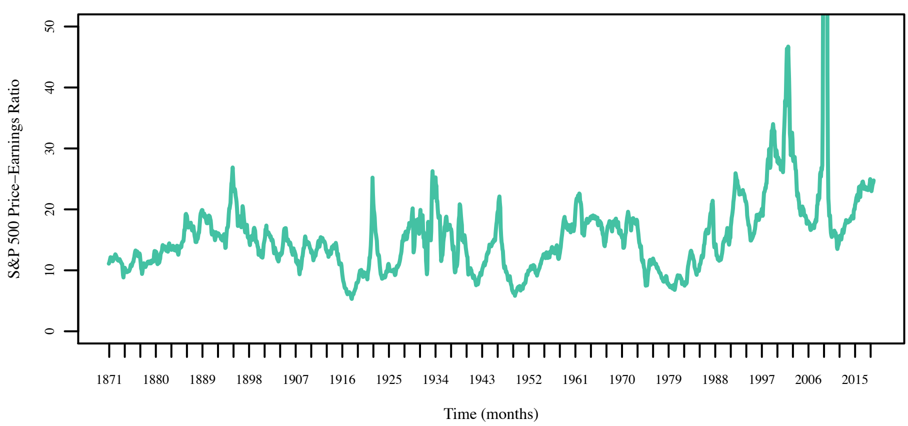


Figure 1.1: S&P 500 Price/Earnings from 1871 to 2018[2]

Our underlying aspiration is to understand and predict the relationship between general firm valuation levels and the macroeconomic environment. We aggregate time series of individual firms

[1]The broad STOXX Europe 600 Index had an open price of EUR 172.71 on March 31, 2009, and an open price of EUR 364.40 on October 11, 2018, an increase of 111%.

[2]The S&P 500 Price/Earnings index aggregates Price/Earnings multiples for S&P 500 firms and is retrieved from Robert Shiller's website, www.multpl.com.

valuation multiples to represent overall valuation levels and use familiar factors to represent the macroeconomic environment. We also model the full distribution of valuation multiples in each cross-section explicitly to extract information about the variability of pricing levels. Finally, we predict future multiples using the developed representations and a selection of forecasting techniques. The conventional and self-developed methods are tested on a dataset of $\sim 2,000$ listed European firms from most industries.

While valuation multiples are widely used in the financial industry, the academic counterpart is primarily concerned with price variables and equity returns.[3] Studies using multiples tend to focus on the market cross-section at particular instances of time.[4] Campbell & Shiller (1988) use time series of multiples adjusted for cyclicality and Kane et al. (1996), White (2000) and Shamsuddin & Hillier (2004) regress price/earnings time series indices on macroeconomic factors. Basu (1977) combine time series- and cross sectional data, testing the Efficient Market Hypothesis. Like Basu, we use two-dimensional data, but we ignore individual stock movements and consider distributional shifts. Methodically, we build on literature from econometrics, statistics, mathematics and machine learning, particularly Alexander (2008*a*) and Tsay (2010),[5] A substantial share of the methodology is self-developed, perhaps with latent applications for other researchers. Contrary to most literature, we use multiples accounting for debt, and the perspective is macroeconomic.[6] With this, we enter a largely unexplored area of the research field.

Chapter 2 is a contemplation of valuation, including a discussion on drivers of valuation multiples. Hypotheses are formulated explicitly in Chapter 3 after assembling an empirical framework. In Chapter 4 we regress indices on macroeconomic variables unearthing the dependency relationships. This understanding is enhanced in Chapter 5, in which we develop techniques to regress the cross-sectional distribution of multiples on macroeconomic factors. In Chapter 6, we propose three methods of forecasting based on the results from Chapter 4 and 5. Finally, we conclude and propose areas of further research in Chapter 7.

[3] See e.g., Ding et al. (1993), Thorbecke (1997), Gultekin (1983), Kwon & Shin (1999), Maio & Philip (2015), Chen (2009) and Cheng (1995) for use of stock market return variables.

[4] See L Little & L Little (2000), Nicholson (1960), Beaver & Morse (1978) and Schreiner et al. (2007).

[5] Other notable textbook sources used for methodical development include Russell & Norvig (2003), Mitchell (1997), Roman (2007), Walpole et al. (1993), Lundgren et al. (2010) and Kreyszig (2010).

[6] A comprehensive discussion on why we use multiples including debt is found in Section 3.2.

Chapter 2

Valuation

All discussions related to valuation are fundamentally associated with the notion of value. Value theory[1] is concerned with how and why subjects value objects, and has been studied in philosophy since the Antiquity. Plato distinguishes between *instrumental* and *intrinsic* value - referring to whether the object is a means or an end by itself (Plato 2008). *Intrinsic* values were largely justified theologically before Kant introduced *personal inviolability* in his categorical imperative (Kant 1949). Value-nihilists deny the existence of Kant's intrinsic values and reduce perception of value to an emotional response, and Nietzsche (1910) claims that this realization will have grave consequences for Europe. In economics, the primary concern is the value of objects[2] as opposed to moral value, and aspects of this value may be empirically observable. The value of an object may also be clearly definable, tangible, predictable and stable or none of mentioned, and its nature has serious consequences for everyday life. The value we will be concerned with here is the value of a firm.

The economic value of a good is often defined as the value an agent would be willing to exchange for the good, while market value is the value for which exchanges actually occur. Many theories on how economic value arises have been proposed, such as intrinsic value theories[3] often associated with Adam Smith (1776), and subjective theories of value proposed by Jevons (1871), Menger (1871) and Walras (1874). In the former, an objective value exists and is usually related to a production factor, while in the latter, the determinant of value is subjective consumer needs. The subjectivist theory is influential for neoclassicistic economic theory, which dominates mainstream economics today. Here, supply and demand functions reflect the economic value of a good for all actors in the market, and the equilibrium in their intersection represents the market value. In valuation of a firm, the equilibrium is formed between supply and demand of the firm as an investment opportunity and will vary the firm's attractiveness relative to other investment opportunities.

The value of a firm is usually defined as the monetary value of equity and net debt. Investors normally acquire equity or debt or both to get an economic return,[4] and the economic value of the

[1] See Hartman (1967) or Edwards (2010) for elaborate discussions of axiology (value theory).

[2] In this discussion, the term object includes anything that may be said to have a monetary value, such as physical objects, services, rights, reputation and others.

[3] Prominent examples of intrinsic value theories are Marx' Labor Theory of Value (Marx 1867), the Land Theory of Value proposed by the Physiocrats (Landreth & Colander 2002) and the Exchange Theory of Value proposed by Rubin in 1927.

[4] Motivations for equity ownership may include realization of transactional synergies, corporate control, market control, and philanthropic purposes.

firm is effectively a valuation of the stochastic cash flow from future dividends and interest. The equity value may be estimated using the discounted dividend model (Williams 1938):

$$V_c = \sum_{t=1}^{\infty} \frac{D_{ct}}{(1+r_c)^t} \tag{2.1}$$

in which t denotes time, V_c is an approximation of the equity value for firm c, D_{ct} is the dividend and r_c is the equity cost of capital. If $\forall t(D_{ct} = (1+g)D_{c,t-1})$, where g is the dividend growth rate, the Gordon Growth model is obtained (Gordon 1962):

$$V_c = \frac{D_{c1}}{r_c - g} \tag{2.2}$$

The monetary value of the dividend growth may be estimated directly by subtracting the value of the no-growth dividend in Equation 2.1, resulting in a dividend Net Present Value of Growth Opportunities (NPVGO) model: $D_{c1}(\frac{1}{r_c-g} - \frac{1}{r_c}) = \frac{gD_{c1}}{r_c^2 - r_c g}$.

Much value creation from equity, however, is realized through capital gains, and dividends may be hard to estimate.[5] Dividing D_t from Equation 2.1 into parts that might be easier to measure, Miller & Modigliani (1961) show that an equivalent measure of equity value is:

$$V_c = \sum_{t=1}^{\infty} \frac{e_{ct} - dB_{ct}}{(1+r_c)^t} \tag{2.3}$$

where e_{ct} is equity earnings for period t and $dB_t = B_t - B_{t-1}$ is the change in book equity, reflecting retained earnings for firm reinvestments (e.g., solidity strengthening, working capital changes and capital expenditures (capex)). Free cash flow (FCF) reflects cash-effects not captured by net income and gives rise to the well-established Discounted Cash Flow model (DCF) (Koller et al. 2010):

$$V_c = \sum_{t=1}^{\infty} \frac{\pi_{ct}}{(1+r_c)^t} \tag{2.4}$$

here, π_{ct} is the FCF at time t, and V_c is value of equity or total firm value depending on FCF definition. DCFs are frequently used in the financial industry and are considered accurate if r_c and FCFs are well known. However, the esoteric risk measurement r_c may be difficult to interpret and estimate, particularly in the presence of uncertainty or complex dynamics, such as options (Dixit & Pindyck 1994). In this case, real options theory may be more adequate, but the frameworks have yet to be adopted by financial professionals as it requires voluminous data on cash flow probability distributions. The value of a real option is usually based on a calculation originating in the continuous Bellmann equation, in which dt denotes the continuous time differential and dV_c the value variable differential:[6]

$$r\,dt = \mathbb{E}(dV_c) \tag{2.5}$$

The mentioned fundamental valuation models (Equation 2.1 to 2.5) estimate value based on a postulate of intrinsic value in future cash flows. Contrarily, valuation methods may be constructed based on empirical equity valuations, widely available from stock exchange data. Based on peers p with associated value V_p and financial information F_p, a function h may be learned, mapping

[5] Dividends are effects of leadership policy and capital strategy. Earnings are often reinvested, and dividend payouts strategies are considered signaling effects. See Black (1976) for an illuminating perspective on the "Dividend Puzzle".

[6] If the underlying value is assumed to follow a Gaussian process, the more familiar differential equation is obtained: $\frac{1}{2}\sigma^2 V^2 \frac{\partial^2 F}{\partial V^2} + \alpha V \frac{\partial F}{\partial V} - \rho F = 0$, in which σ is the stochastic parameter and α is the drift parameter in the Gaussian process of V.

financial data to a value approximation for firm c being valued: $V_c = h(F_c)$. A conventional way to create h is to use a multiple $m_p = V_p/\pi_p$ based on some profit measure π_p and to apply an aggregation[7] to obtain a multiple m. With this, a value approximation for company c is obtained:

$$V_c = \pi_c m \tag{2.6}$$

Multiples represent a convenient, pragmatic and empirical valuation method, and are extensively used in business, usually in combination with intrinsic methods. Businesses use multiples because they are relatable and provide an intuition on what one pays relative to what one gets. The core of our study does not concern valuation of individual enterprises, but the general market valuation levels. Development of the aggregate V_c from listed firms represents a valuation of the market portfolio of firms but is associated with some fundamental issues: portfolio composition changes, firm profitability is neglected, different markets are incomparable, and the time series are unit-root nonstationary, volatile, and lack intuitive interpretation. We seek a valuation metric relative to the benefit obtained by investors, and the Copernican Turn of this thesis is done to Equation 2.6, obtaining:

$$m_c = \frac{V_c}{\pi_c} \tag{2.7}$$

This measure will be our primary concern because it provides insight into the market value relative to corporations' abilities to generate profits. m_c reflects a view on the future and serves as a point of reference vis-a-vis intrinsic valuation. Growth expectations, risk profiles, and sentiments of investors are implicitly included in V_c, and the inclusion of π_c allows for comparison between enterprises, despite differences in size. The design of Equation 2.7 may be criticized for being arbitrary, but it turns out that it largely captures the statistical aspects of size independence.[8] While the distributions of V_c and π_c are expected to adhere to Gitbrat's law (Gibrat 1931), the distribution of multiples is a fundamental representation of a state of the market valuation, and aggregated multiple indices are valuable for investors as a gauge of the current market valuation, suggesting an answer to the question:

What is the price for a unit of profits?

The remainder of this chapter is a cursory presentation of how the macroeconomic environment can be expected to affect multiples, but specific hypotheses will not be formulated until an empirical framework is established in Chapter 3. For the sake of transparency in the discussion, we may reveal that a multiple accounting for value creation for *all* investors will be used in the trunk of our analyses: *median EV/EBITDA*. It is instructive to distinguish between effects on the numerator, V_c, and the denominator π_c. The former is expected to respond to changes in variables reflecting the investor perception of the future, while the latter mirrors the current business environment.

The numerator, V_c, is driven by the demand for investment opportunities and the supply of asset classes, both factors varying with the underlying macroeconomic environment.[9] The intersection of supply and demand determines the valuation of a firm:

The overall demand for investment opportunities is governed by the amount of capital seeking to be invested in portfolios of a specific risk. Savings trends are affected by political, social and macroeconomic factors, such as low interest rates that incentivize gearing of investments. Bernanke (2005) argues that savings capital from emerging markets has contributed to a "global savings glut"

[7] A range of techniques for aggregating a set of multiples to an individual value is discussed in Section 3.2.2.

[8] In the introduction to Chapter 5, we show that the distribution of m_c is roughly independent of V_c and π_c.

[9] The focus of this discussion and our thesis is rational factors, although impacts of irrational and behavioral aspects may be significant. See Shiller (2000), Barberis & Thaler (2002), Baker & Wurgler (2007), Barberis et al. (1999) and Frazzini & Lamont (2005).

in developed markets, but the empirical foundations are controversial (Chinn & Ito 2007). Savings trends may also be associated with inequality (Bourguignon 1981), and Piketty (2015) shows that income equality has risen persistently the last fifty years. However, the neoclassicist approach of Bourguignon is criticized for being simplistic, and Alvarez-Cuadrado & El-Attar Vilalta (2012) argue that the opposite is the case. Policymakers affect demand through central bank interest rates and schemes such as Quantitative Easing (Joyce et al. 2011). The current levels of dry powder in the economy have increased demand for riskier asset classes as investors are chasing higher yields in a low interest rate environment.

The overall supply of investment opportunities constitutes substitutes[10] competing primarily on two factors: risk and return. These are intimately linked in the risk-return spectrum and the Sharpe-ratio - representing the fundamental trade-off in investment theory (Sharpe 1964). Risk is incorporated in valuation as a firm's uncertainty surrounding future cash flows and its cost of capital, which is the price investors pay for bearing risk. Uncertainty is often measured in terms of standard deviation,[11] and is associated with both idiosyncratic and systemic risk. According to Sharpe (1964), the risks affecting cost of capital are primarily systemic risks reflecting, e.g., political, social and environmental conditions, which cannot be diversified away like idiosyncratic risks (Markowitz 1952). Investor fear barometers, such as the CBOE Volatility Index, reflect risk expectations and are closely followed by the markets. In many valuation methods, such as Equation 2.4, risk is included explicitly in the discount rate r_t. When firms are valued at Enterprise Value level, WACC may be used as discount rate (Reilly & Wecker 1973):[12]

$$r_{\text{WACC}} = \frac{\text{E}}{\text{D+E}} r_e + \frac{\text{D}}{\text{D+E}} r_d (1-\tau) \tag{2.8}$$

Consequently, interest rate hikes raise costs of capital. Sharpe (1964) and Lintner (1965) propose CAPM[13] as a measure of r_e - making the equity return-interest dependency explicit.[14]

Besides risk, factors affecting the prospects of future cash flows, π_{ct} from Equation 2.4, are determinants of the supply of investment opportunities. A natural assertion is to associate π_{ct} growth with GDP growth, because GDP measures overall domestic value creation (Mankiw 2014).[15] As expected from The Efficient Market Hypothesis (Fama 1970), GDP growth is not an important driver for stock market returns (Ritter 2005). However, a variable reflecting GDP growth expectations would be expected to correlate with concurrent returns,[16] and this is established empirically (Lee 1992). Inflation improve future earnings prospects π_{ct}, but Fisher (1930) argue that cost of capital r_{ct} is increased correspondingly, netting out the effect. In fact, Modigliani & Cohn (1979) and Ritter & Warr (2002) argue that investors suffer from a "money illusion" (term coined by Fisher in 1919) rendering them unable to distinguish between real and nominal growth rates, resulting in

[10] The substitutes are different asset classes like private and public equities, fixed income, and real estate investments but also individual companies or even projects.

[11] Conventional risk measures include q%-VaR, Expected Tail Loss, Absolute deviations and Downside risk measures. See Alexander (2008*a*) for elaboration.

[12] r_{WACC} is the Weighted Average Cost of Capital, $\frac{\text{E}}{\text{E+D}}$ is the target Equity to Market Asset Value-ratio, r_e is the return on equity, r_d is the interest payment and τ is the corporate tax rate. For companies with other securities (e.g., preferred stock) additional terms have to be added, each including security's expected rate of return and as a percentage of total value.

[13] Capital Asset Pricing Model is the traditional model for estimating r_e using $r_e = r_f + \beta(r_m - r_f)$.

[14] Fama & French (1993) propose a three-factor expansion of CAPM incorporating size and value and a five-factor model in Fama & French (2015) including profitability and investment level, but these risk-affecting factors are related to individuals firms.

[15] Firm profits and correlated variables like wages and taxes are constitutes of the Income Approach for GDP.

[16] Some indicators have been recognized as leading variables for GDP growth: Industrial production growth, durable goods orders, business inventories, and Purchasing Managers Index (PMI) (Baumohl 2012).

undervaluation. Another negative impact of inflation is rooted in the inertia of revenue response to increased operational expenditures (opex) and capex (Koller et al. 2010).

The denominator, current earnings π_c, mirrors the prevailing macroeconomic environment and is expected to move relatively independent of V_t because the movement expectations largely will be incorporated in the valuation before they are realized (Fama 1970). If a company delivers on ambitious earnings expectations, the multiple will depreciate ceteris paribus. However, the fluctuations of π_t are smaller than those of V_c, rendering the latter the primary driver of the multiple, while the former provides size independency, see Figure 3.1 in Chapter 3.

The theoretical considerations in this chapter serve as a platform for understanding macroeconomic drivers of firm value multiple aggregates. Many of the macroeconomic factors are interlinked in complex ways, and in Chapter 6, some of these dynamics will be discussed. The economy tends to alternate between periods of expansions and recessions, and the endeavor to predict business cycles and to identify asset price bubbles have kept many researchers occupied.[17] Figure 1.1 is an illustration of how this cyclicality affects multiples.

The notion of value is fascinating both from a philosophical and an economic perspective, and the Copernican Turn from Equation 2.7 places value in the center.

[17] See e.g., Yan (2011), Hamilton (1989), Al-Anaswah & Wilfling (2011), Campbell & Mankiw (1987) and Beveridge & Nelson (1981).

Chapter 3

Data

In order to explore the relationship between enterprise pricing and the underlying macroeconomic environment, we need data describing both. A comprehensive and representative data set allows us to explore a wide range of hypotheses and to obtain solid statistical foundations for our conclusions. In this chapter, the process of extracting, rinsing, modifying, transforming, selecting and allocating the data is described, resulting in the sets of data used for our models.

There are always issues related to the use of empirical data. Any statistical hypothesis entails assumptions about the underlying model and the sampling process from it, and any forecast will inexorably be subject to The Problem of Induction (Hume 1779). Even the very existence of a statistical model is an intellectual construct (Edgeworth 1884). We will be transparent on which assumptions we make, giving the reader an equitable impression of the validity of our inference. In order to answer our research goals, we will define what we mean by firm pricing and macroeconomic environment before retrieving suitable data.

3.1 Source and Data Scope

FactSet[1] is a for-pay data source, providing quality-assured financial data along with extensive descriptions and metadata on a granular level. Because we explore historical relationships between variables, all data used is time series data. The pricing variables are computed from a large set of publicly listed firm data,[2] while the exogenous variables are indices, rates, and other time-series data extracted directly from FactSet. Transparent and consistent data preprocessing is conducted by the FactSet team and will be emphasized where it is relevant. The final data extraction was conducted the 16th of August 2018.

The data horizon is 223 steps of monthly time-series data spanning from January 2000 to July 2018. This range includes business cycle expansions and recessions, and is intended to capture underlying equilibrium relationships in the economy,[3] but is kept relatively short to be representative of

[1]From FactSet website: FactSet creates flexible, open data and software solutions for tens of thousands of investment professionals around the world, providing instant access to financial data and analytics that investors use to make crucial decisions.

[2]Listed firms are assumed to be a relatively good approximation of overall market valuation levels and quality and availability are superior to private firm data.

[3]See references in Chapter 2 and Keynes (1936) for an elaborate discussion of business cycles.

the prevalent economic structure. Data availability, validity of resulting inference and conventional approaches (e.g., see Shamsuddin & Hillier (2004) for similar approach) were also key in determining range. Monthly, rather than daily or annual data, was considered to be the correct balance of smoothing out short-term fluctuations and providing a data set of appropriate size for statistical inference. A European scope is chosen as a trade-off between global relevance, data availability and data manageability.[4] Data on sectors and geography of individual firms was available, but not included in the scope.[5]

3.2 Measuring Firm Valuation

We seek to create a measure of the valuation of individual firms relative to their profit-generating abilities and to use the distributions of these measurements (multiples) and resulting indices in modeling. Letting V_{ct} be a valuation metric for enterprise c at time t and π_{ct} a profit metric, the valuation multiple is defined as:

$$m_{ct} = \frac{V_{ct}}{\pi_{ct}} \tag{3.1}$$

A common measurement of value V_{ct} is Market Capitalization (MCap) - the value of all outstanding shares and thus, the control of the firm. However, this valuation is dependent on capital structure and fails to account for debtors as investors in the firm. As discussed in Chapter 2, the capital structure-independent value of a firm is determined by its future dividends and future interest payments. For a listed firm, the market value of future dividends is observable as the market capitalization, while the market value of the interest payments is the net debt and may be obtained from floated bonds or approximated by reported values. Henceforth, the valuation metric used will primarily be Enterprise Value (EV), when not specified otherwise:

$$\text{EV}^{6} = \text{MCap} + \text{Interest bearing debt} - \text{Cash \& Other Liquid Assets} \tag{3.2}$$

Monthly time series of EV for each listed European firm are preprocessed by FactSet and include year-end net debt values from the most recent annual report at the time. Data on floated bonds is practically unavailable, but balance sheet values approximate market values (See IRFS 13 "Fair Value Measurement").

The profit measurement, π_{ct}, should approximate the underlying value creation at the moment of observation. Net income would be natural if V_{ct} was market capitalization since it measures the profit attributable to equity owners after tax, but EBITDA, EBIT, and NOPLAT are more natural as they include profits to debt holders and eliminate the distorting effect of nonoperating assets and nonoperating income statement items. Depreciation and amortization mirror capital expenditures in the long term, but are not cash effects, tend to fluctuate and may be incomparable across borders in Europe.[7] Excluding D&A, EBITDA is obtained - a more stable measurement and a good proxy of actual cash-generating abilities before investments. Its values are consistently higher than EBIT

[4]The listed firm data primarily reflects large Western European economies. By firm headquarter: United Kingdom (22%), France (12%), Germany (11%), Sweden (9%), Italy (6%), Switzerland (6%), Norway (4%), Spain (4%), Finland (4%) and Poland (3%).

[5]The five largest industries by number of firms are producer manufacturing (12%), health technology (9%), technology services (8%), electronic technology (7%) and consumer services (7%). Note that the financial industry is excluded from the sample because their different balance sheet structure (deposits as debt and loans as assets) make enterprise value multiples invalid.

[6]Preferred stock and minority interest are also added to EV in FactSet.

[7]In particular amortization is affected by extraordinary events such as impairment of intangible assets. Thus, one could argue that depreciation tends to mirror a more stable capex, and EBITA would be superior. However, data availability is poor.

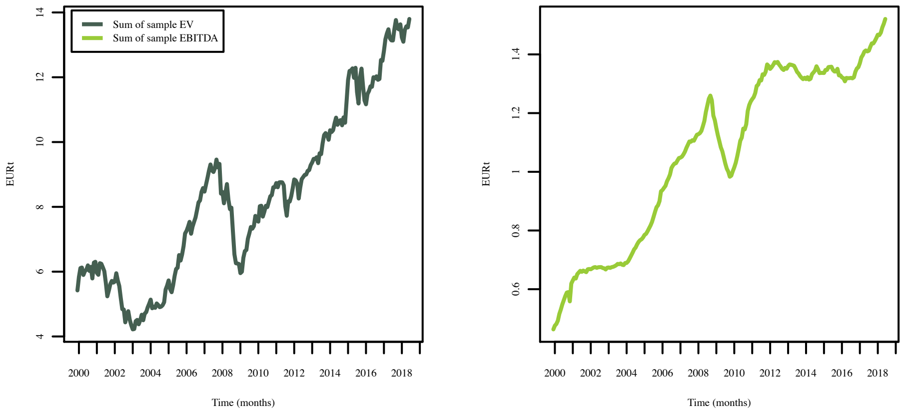


Figure 3.1: Time development of total EV and total EBITDA in trillion EUR

and NOPLAT, rendering these multiples defined for more firms. However, it creates a slight bias in valuation of companies with large investments or heavy balance sheets. A reasonable adjustment is to subtract capex from EBITDA, as expected investments impact expected returns (Miller & Modigliani 1961). However, capex data is volatile as it includes irregular investments like project development capex, and does not necessarily represent value creation in the concurrent year. EBITDA has a strong rationale, performs well (Kaplan & Ruback 1995), is conventional and easily interpretable,[8] matches EV well and has large data availability. From this point on, it will be the primary π_{ct} measurement:

$$\text{EBITDA} = \text{Sales} - \text{CoGS} - \text{SG\&A} \tag{3.3}$$

Monthly time weighted LTM EBITDA is extracted directly from FactSet for each company.[9] Figure 3.1 displays the sum of EV and sum of EBITDA for the sample of firms. The EV decrease and the subsequent EBITDA decrease during the Global Financial Crisis (GFC) are clearly visible. The EV/EBITDA multiple is computed for every firm c at every time step t and will always be the multiple referred to, unless otherwise specified. The cross-sectional median multiple development is displayed in Figure 3.2 along with the number of companies in sample at each time $|\Omega_t|$.

3.2.1 Data Synthesis

There are a few issues related to the computation of multiples; they are only defined for positive V_{ct} and π_{ct},[10] and small π_{ct} leads to extreme valuation multiple outliers. The issues apply to $\sim 5.4\%$, $\sim 0.8\%$ and $\sim 0.5\%$[11] of all observations, respectively. In the majority of the analyses, we will exclude these observations, although noting that it creates a minor bias.

Figure 3.3 illustrates the fraction of observations taken out due to issues with undefined multiples. The maximum share of excluded companies over the time period under consideration is

[8] Fernandez et al. (2001) suggests that EV/EBITDA is the 2nd most used valuation multiple after P/E.

[9] Factset have estimated monthly LTM (Last Twelve Months) earnings π_{ct}^{m} based on yearly annual report earnings from the preceding year π_{cp}^{y} and the current year π_{cc}^{y} as $\pi_{ct}^{m} = \frac{1}{12}((12-t)\pi_{cp}^{y} + t\pi_{cc}^{y})$. For 2018 values, median broker consensus estimates are used.

[10] Negative EBITDAs result in negative multiples, i.e., a low valuation, when in fact they are valued highly compared to their weak earnings.

[11] Extreme outliers are here defined as multiples with value larger than 50x.

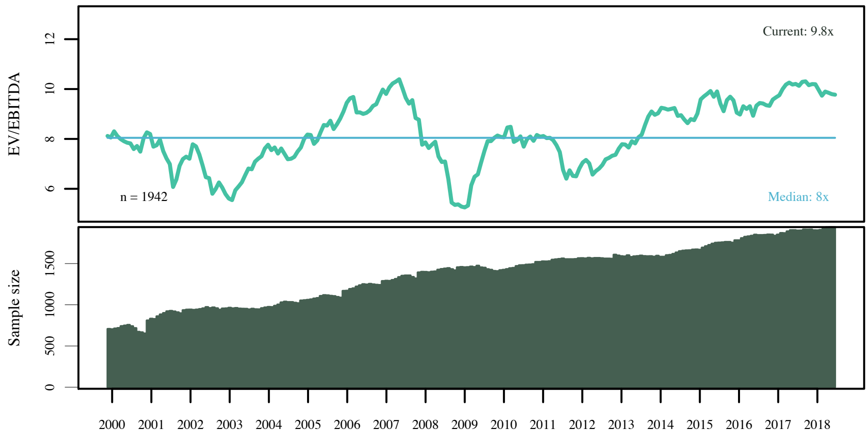


Figure 3.2: Median EV/EBITDA development and number of firms in the dataset

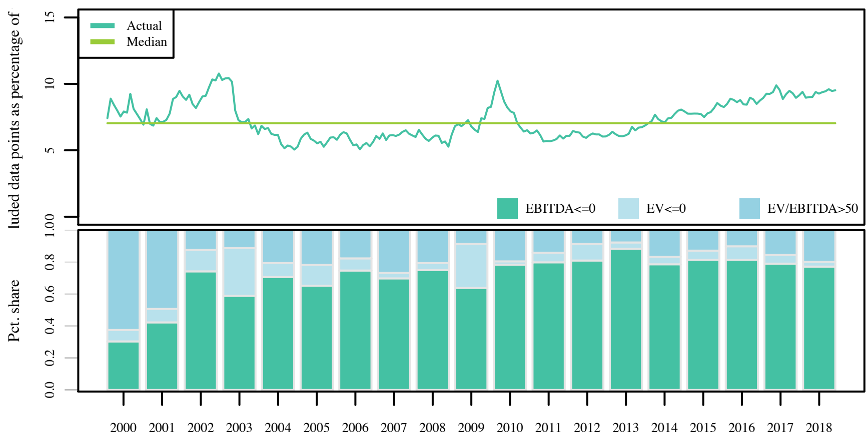


Figure 3.3: Upper: Excluded data points as % of total, Lower: Reason for exclusion

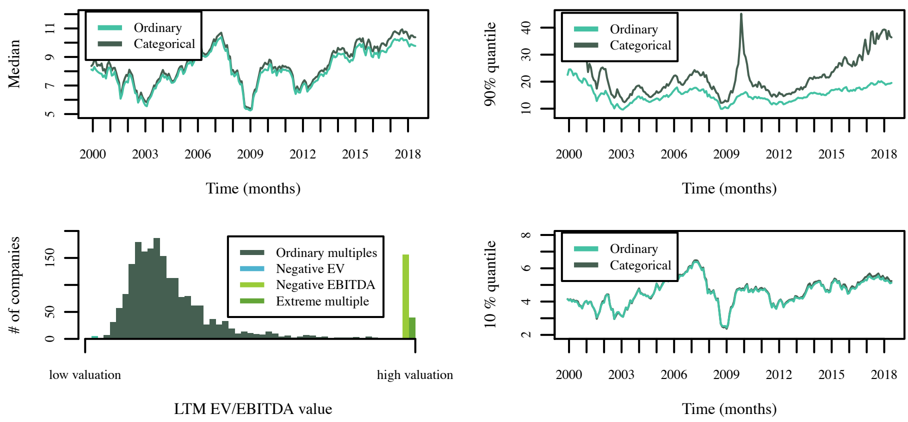


Figure 3.4: Median, quantiles and intersection of July 2018 distribution of the "categorical model"

approximately 10% - primarily due to to the effects of the tech stock mania in 2000 and GFC in 2007/2008. We hypothesize that the bias is too small to impact results significantly, but we have developed an algorithm to minimize the effects of the issue and will use this to test the hypothesis in Subsection 4.1.3.

The algorithm is conducted in the following manner: Allocate EV<0 values to a low value-category, EBITDA<0 to a high value-category and EV/EBITDA>50 to another high-value category. Now, a transitively ordered set of firm multiples may be obtained,[12] on which aggregations, such as quantiles, may be conducted to obtain numerical or categorical values. This model is referred to as the "categorical model", and median, quantiles and an intersection of the distribution are plotted in Figure 3.4. The extreme quantiles are higher than the ordinary quantiles as these are proportionally more affected by the inclusion of otherwise truncated data.[13] Generally, the categorical model and the model with omitted extreme values are similar, but the latter is easier to interpret and to work with, less sensitive to outliers and is numerically defined for all variables and thus, will be the main model.

3.2.2 Index Data

Chapter 4 is concerned with analyses on time-varying index variables aggregated from the cross-sectional distribution of multiples. Individual company multiples are computed according to Equation 3.1, before applying an aggregation function $\aleph$ to the set of multiples at time t, $\Omega_t = \{m_{c\tau} | \tau = t\}$:[14]

$$\aleph(\Omega_t) = y_t \qquad \text{i.e.} \quad \aleph : \mathbb{R}^{|\Omega_t|} \to \mathbb{R} \tag{3.4}$$

The "aggregation problem" referred to in Chapter 2 is the problem of finding an an appropriate $\aleph$. Natural definitions of $\aleph$ include median, mean, trimmed mean (Stigler 1973), π_{ct}-weighted aver-

[12] The transitive ordering is such that all values in the high-value bin are equally high and higher than values in other bins, and correspondingly, all values in the low-value bin are equally low and lower than all other values.

[13] If the qth quantile in the excluded $Q = F^{-1}(q)$, the qth quantile in the categorical model will be $Q_{cat} = F^{-1}(q\frac{1+o}{m})$ if there are o omitted variables, m included variables and $q < \frac{m}{m+o}$.

[14] The constraints we impose on $\aleph$ are relatively strict. See Shiller (2000) adjusting for earnings cyclicality by composing an index from π-weighted rolling average of S&P500 P/E- ratio for more creative $\aleph$s.

Aggregation	Numerator	Denominator	Calculation	Mean	Median	Standard Deviation	Max	Min	ADF test statistic (p-value)
Median	EV	EBITDA	$\text{Median}_C(m_{ct})$	8.15	8.04	1.27	10.4	5.25	-2.80 (0.24)
Mean	EV	EBITDA	$\frac{1}{\lvert\Omega_t\rvert}\sum_{c=1}^{C} m_{ct}$	9.38	9.39	1.43	11.9	6.07	-3.11 (0.11)
V_{ct}-weighted	EV	EBITDA	$\frac{V_c}{\sum_{c=1}^{C} V_c}\sum_{c=1}^{C}\frac{V_c}{\pi_c}$	9.43	8.82	1.85	15.5	6.29	-3.86 (0.02)
π_{ct}-weighted	EV	EBITDA	$\frac{\sum_{c=1}^{C} V_c}{\sum_{c=1}^{C} \pi_c}$	7.85	7.61	1.36	12.7	5.04	-3.05 (0.13)
Median	Price	Earnings	$\text{Median}_C(m_{ct})$	15.6	16.1	2.37	19.6	8.44	-2.89 (0.20)
Median	EV	EBIT	$\text{Median}_C(m_{ct})$	11.85	11.80	1.64	15.01	7.33	-2.86 (0.21)
Median	EV	Sales	$\text{Median}_C(m_{ct})$	1.12	1.10	0.22	1.59	0.67	-2.65 (0.30)

Table 3.1: Overview of indices and statistical properties

age[15] and V_{ct}-weighted average.[16] Because the distribution is skewed, the mean and median have rather different values. The multiple distribution yields extreme values making the mean unsuited, and the weighted averages are driven by large firms and are incomparable to the distribution. The median is representative, invariant to many distributional transformations[17] and generalize naturally to quantiles and the rest of the distribution, and will hence be the primary aggregation function, leading to the primary index: *median EV/EBITDA*, which will be denoted as y_t unless otherwise specified. No logarithmic transformation of y_t is conducted, because y_t is relatively close to $\sim \mathcal{N}(\mu,\sigma)$[18], y_t has causal relationship with the exogenous variables (see Subsection 3.3) and a transformation would be unconventional (see Ramcharran (2002) and Zorn et al. (2008)).

Table 3.1 displays some statistical properties of seven indices: the standard *median EV/EBITDA*, three based on variations of $\aleph$ and three based on variations of V_{ct} and π_{ct}. As expected, EV/Sales < EV/EBITDA < EV/EBIT because Sales > EBITDA > EBIT. P/E yields numerically higher values as earnings are net of taxes. A unit-root cannot be rejected by an ADF-test for most indices, rendering a differentiation natural:[19]

$$d_t = \log(\frac{y_t}{y_{t-1}}) \tag{3.5}$$

The differentiated indices confidently reject unit root hypotheses for all models, and it is natural to assume $I(1)$ for the undifferentiated indices. The historical development of the seven indices resembles that of *median EV/EBITDA*, as seen in Figure 3.5.

3.2.3 Distribution Data

The full data we are facing m_{ct} is panel data and could be subject to panel analysis methods.[20] However, our focus is on the macroeconomic picture and systematic risk, and not on individual company stories and idiosyncratic risk. The objects of our study are the cross-sectional distributions, Ω_t. Figure 3.6 shows a snapshot of Ω_t at month-end July 2018. The distribution is skewed and has a fat high multiple tail. Figure 3.7 shows the log distribution at the same instance for the underlying EV and EBITDA - bell-shaped, as expected from Gibrat (1931)'s law. The underlying distribution

[15] The π_{ct}-weighted average is equivalent to the construction of an artificial firm: $\frac{1}{\sum_{c=1}^{C}\pi_c}\sum_{c=1}^{C}\pi_c(\frac{V_c}{\pi_c}) = \frac{\sum_{c=1}^{C} V_c}{\sum_{c=1}^{C}\pi_c}$, and is the only aggregation that is unaffected by $\pi_{ct} < 0$ or $V_{ct} < 0$.

[16] Strictly speaking, the π_{ct}- and V_{ct}-weighted index-$\aleph$s are not only functions of Ω_t, but also of $\{V_c\}$ and $\{\pi_c\}$.

[17] For any monotonous transformation T, the following follows trivially: $\aleph(T(\Omega_t)) = T(\aleph(\Omega_t))$ if $\aleph$=median.

[18] A JB normality test gives p-value of $p = 0.05$ (Jarque & Bera 1980), and the Box-Cox transformation parameter is ~ 1.4 (Box & Cox 1964). See Section 4.1 and Section 3.3 for respective elaborations.

[19] An elaborate discussion of unit root stationarity, ADF-tests and consequences will follow in Chapter 4. See Alexander (2008*a*) for notational remarks.

[20] See further research, Appendix E.I, for elaboration on panel data and panel analysis methods.

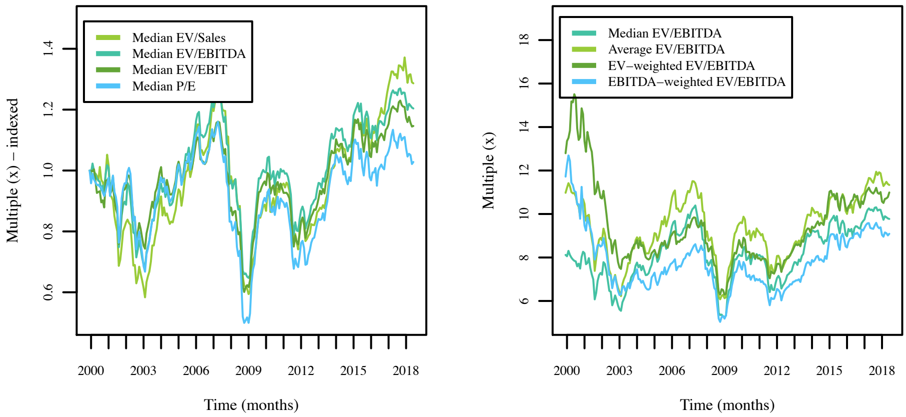


Figure 3.5: Historical development of seven indices based on four aggregation methods and four sets of variables. Note that varying variable indices are normalized

from which Ω_t is drawn, is the quotient distribution of dependent drawings from the two underlying distributions. See Chapter 5 for a thorough discussion on this topic.

The indices provide information about the overall multiple level development, but no *relative* information about the multiples like the distribution development. In Chapter 5, we develop analyses involving Ω_t in three representational forms: quantile data, raw sample data and frequency data. Quantiles y_{qt} are computed for every distribution of multiples,[21] giving an easily interpretable measure of the cross-sectional development of multiples historically. In Figure 3.8, the 10% to 90% quantiles are plotted. Frequency data $\bar{v}_{bt}$ is obtained by counting the number of companies within a certain multiple range (bucket), $[m, m+\Delta_b)$, for each time t, and dividing by $|\Omega_t|$. This set of data, as displayed for $t = 223$ in Figure 3.6, will be utilized for analyses on the dynamics of the complete distribution. Naturally, the frequencies adhere to $\sum_{b=1}^{B} \bar{v}_{bt} = 1$ for B buckets. The bucket ranges can be defined by fixing Δ_b or $|\Omega_{bt}|$, or by optimizing with respect to some other statistic. Raw sample data is used in Chapter 5 for MLE-estimation, and the sample moments of Ω_t are used for comparison against fitted distributions.[22] Figure 3.9 shows a 3D-plot of the development of distributional frequencies over time, illustrating the complete dimensionality of the distribution data.

3.3 Exogenous Variables

The exogenous variables describe features of the macroeconomic environment and will be used as independent variables in the regressions. We select variables using a structured funnel approach illustrated in Figure 3.10 - an essential part of obtaining robust models with sensible predictions in Chapter 4, 5 and 6. The selection approach narrows down from a universe of macroeconomic variables to six portfolios of variables, associated with specific models presented in Chapter 4. The funnel is comprised of filters, expansions, separations and merger steps. Subsection 3.3.2 contains a brief description of the resulting models and variables and offers intuition on why the variables may

[21] y_{qt} for the q'th quantile is defined by: $y_{qt} = \bar{\omega}_{\lceil q|\Omega_t|\rceil,t}$ where $\bar{\omega}_{nt}$ are sorted multiples at t, indexed by n.
[22] See Appendix C.IX for an overview of distributional moment development of Ω_t

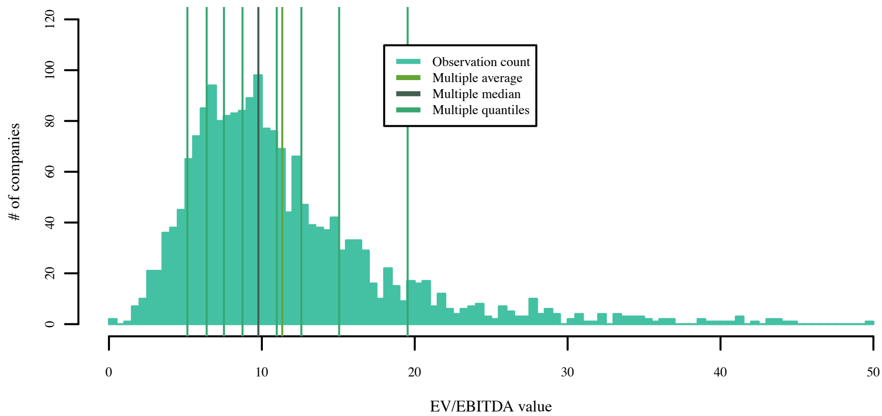


Figure 3.6: Distribution of company multiples July 2018

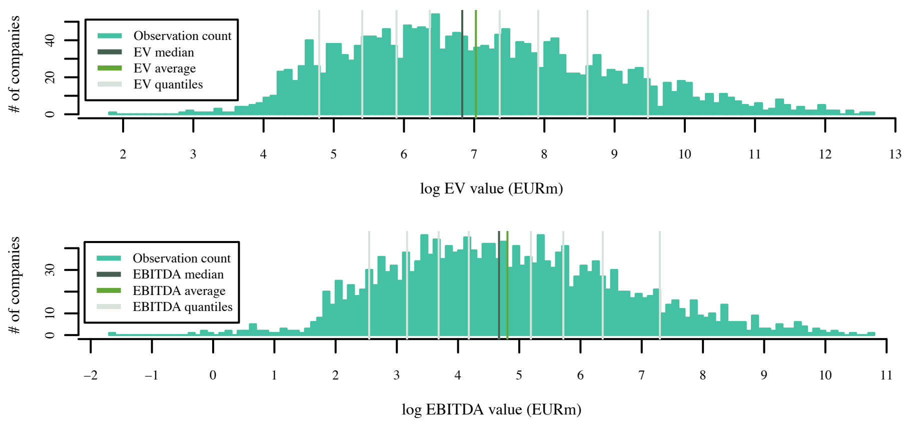


Figure 3.7: Distributions of EV and EBITDA values as of July 2018

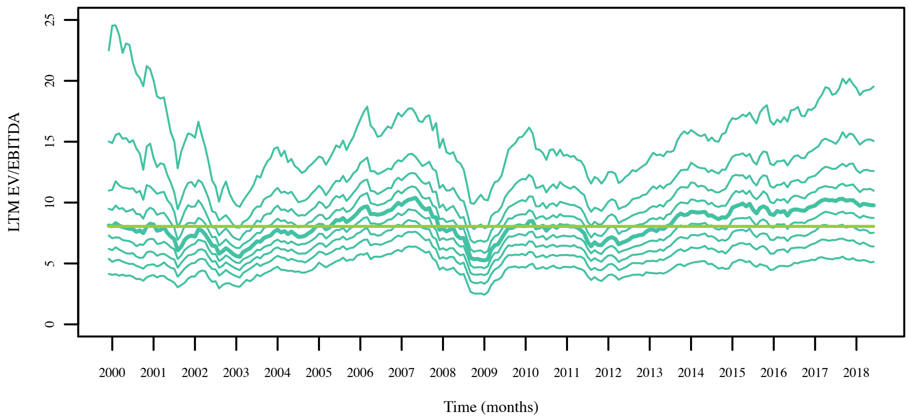


Figure 3.8: 10% to 90% quantiles of EV/EBITDA multiple

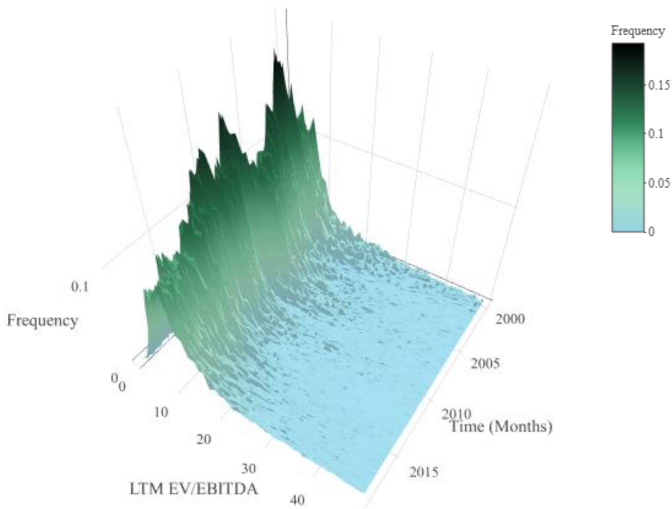


Figure 3.9: Empirical EV/EBITDA distribution development from 2000 to 2018

	Transformation[24]
Interest rates	Absolute differentiation
Financial balance sheets	Log returns
Labor market	Absolute differentiation
Inflation	Absolute differentiation & log return
Survey data	Absolute differentiation & log return
Retail sales	Log return
PMI	Absolute differentiation
Industrial production	Delta log return
Volatility	Log return

Table 3.2: Transformations performed on each category of variables

be relevant for explaining multiple development.

3.3.1 Variables Selection Funnel

The data in scope of the initial selection is a range of macroeconomic factors and other data that may reflect the state of the European economy, obtained from FactSet. For the regressions to be fruitful, the independent variables must have a causal exogeneity to the multiples. Consequently, financial variables representing substitute investment opportunities, such as commodity prices, derivatives and fixed income, are excluded. Failure to include the complete universe of variables inevitably induce an omitted variable bias, but we start with a broad selection of categories and variables, illustrated in Figure 3.11. A complete list of the 65 initially selected variables is provided in Appendix A.I.

The first step of the selection process is a funneling based on availability and quality. Variables with missing values, wrong granularity, flawed or invalid values in the relevant time frame are removed, [23] and some variables were deemed unrealistic or irrelevant. Consequently, the number of variables is reduced from 65 to 40.

The order of integration $I(i)$ is not conspicuous for the individual exogenous variables but was estimated based on ADF testing and interpretation of the variables' nature. The variables are mapped into $I(1)$ and $I(0)$ datasets by being differentiated the appropriate number of times. The differentiation method depends on the nature of the variables: interest rates, inflation rates, and other variables that may be negative were differentiated arithmetically, while variables of an exponential nature, such as central bank balance sheets and volatility, were logged and differentiated. For variables with an ambiguous nature, both methods were utilized. Table 3.2 exhibits the differentiation conducted on each class of variable.

$I(1)$ dependents will be regressed on $I(1)$ independents and correspondingly for the $I(0)$ variables - a necessary condition for regression validity[25] and intuition[26]. The two sets of independent variables comprise a set of 55 $I(1)$ variables for analysis of absolute pricing indices and a set of 44 $I(0)$ variables for analysis of differentiated price indices.

[23] The time frame considered is the same as for the valuation multiple data: Monthly data between January 31, 2000, and July 31, 2018.

[24] For variable x_t and x_{t-1}, differentiations are defined as: absolute differentiation - $d_t = x_t - x_{t-1}$, log return - $d_t = \log(\frac{x_t}{x_{t-1}})$, delta log return - $d_t = \log\frac{x_t}{x_{t-1}} - \log\frac{x_{t-1}}{x_{t-2}}$.

[25] The Engle-Granger-cointegrating regression with independent and dependent $I(i)$-variables and stationary residuals is in fact the *only* regression for which integrated variables may be subject to OLS (Alexander 2008*a*).

[26] It is natural to expect that two economic variables are in a long-term equilibrium or that changes in one affect the other, but not the combination.

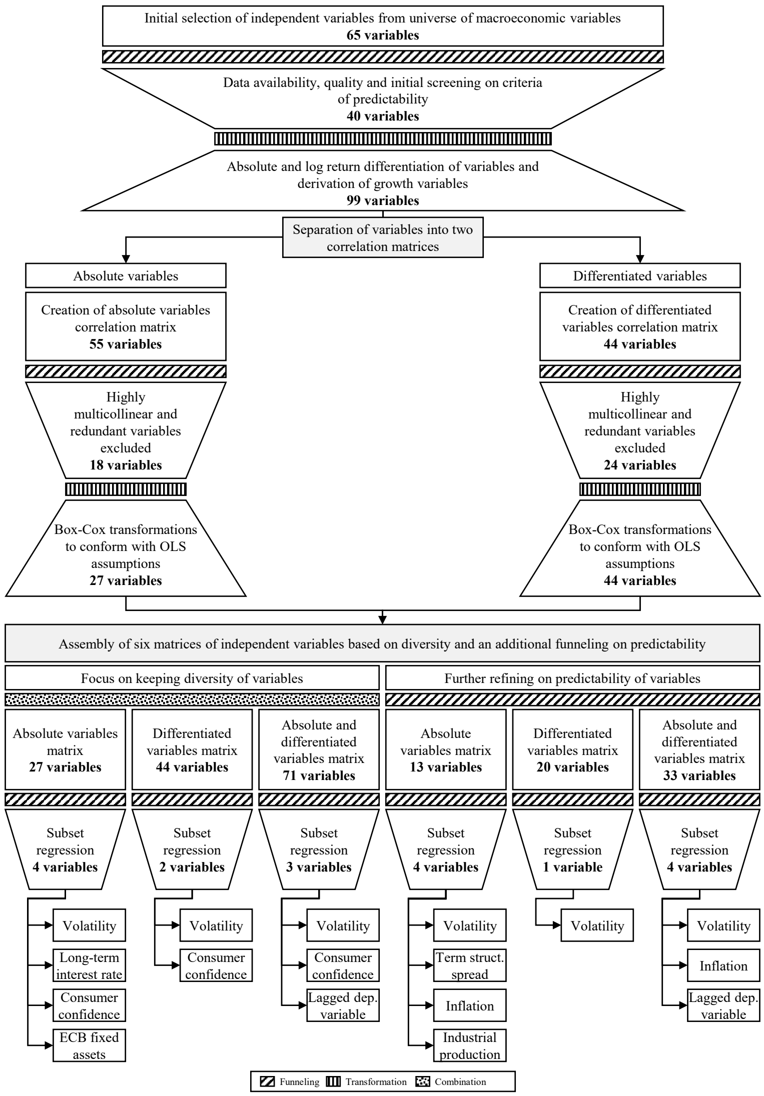


Figure 3.10: Illustration of steps in independent variables selection funnel

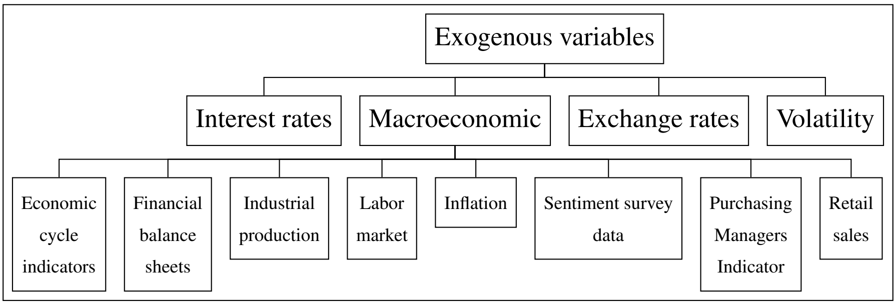


Figure 3.11: Categories of independent variables

In an attempt to avoid issues caused by multicolinearity in the multivariate regressions,[27] a filtering based on inter-variable correlation was conducted by observing the correlation matrices in Figure 1 and 2 in Appendix A.II. For each set of multicolinear variables, individual variables are removed sequentially until all variable pair-correlations are satisfyingly low.[28]. In a stepwise procedure, the data sets were reduced to 18 $I(1)$ variables and 24 $I(0)$ variables, see Figure 3 and 4 in Appendix A.II for final correlation matrices.

OLS estimation results are invariant to linear transformations, while normality-inducing transformations, such as the Box-Cox transformation, may improve accuracy (Box & Cox 1964). Box-Cox transformations are designed to make distributions more similar to the normal distribution, and parameters are estimated using an MLE-method or using a normality test statistic (Sakia 1992). The Box-Cox transformation is defined as follows:

$$x^{(\lambda)} = \begin{cases} \frac{x^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0 \\ \log x, & \text{if } \lambda = 0 \end{cases} \tag{3.6}$$

However, as several of the time series include negative values, we use a transformation proposed by Yeo & Johnson (2000) without restrictions on variable domain:.

$$x^{(\lambda)} = \begin{cases} \frac{(x+1)^{\lambda}-1}{\lambda}, & \text{if } \lambda \neq 0, x \geq 0 \\ \log(x+1), & \text{if } \lambda = 0, x \geq 0 \\ -\frac{(-x+1)^{2-\lambda}-1}{2-\lambda}, & \text{if } \lambda \neq 2, x < 0 \\ -\log(-x+1), & \text{if } \lambda = 2, x < 0 \end{cases} \tag{3.7}$$

λ is estimated using the *car*-package in R, and parameters are rounded to nearest multiple of $\frac{1}{2}$. The post-transformation funnel contains 27 $I(1)$-variables and 44 $I(0)$-variables.

Interest rates, market volatility, and GDP are variables investors often are interested in. They might not have a specific view on the development prospects, but very often, they wish to understand pricing sensitivity to these factors. These kinds of variables will be referred to as "predictable", while a range of other variables, including Consumer Confidence Indicators and Purchasing Manager Indices are less intuitive and predictable. In the next funnel step, we distinguish between these two types of variables and create a data set comprised solely of predictable variables and a diverse dataset

[27] Multicolinearity makes coefficient estimators unstable and inference about drivers flawed (Alexander 2008*b*).

[28] A rule of thumb is used: $\rho(a,b) < \sqrt{R^2}$ for the corresponding regression. For example: a set of interest rate variables was reduced to one long term rate and one short term rate, as initial rates were highly correlated.

containing all variables. The diverse sets contain 27 $I(1)$ variables and 44 $I(0)$ variables, respectively, while the predictable sets contain 13 $I(1)$ variables and 20 $I(0)$ variables, respectively.

3.3.2 Final Variable Description

The objective of the data funnel is ultimately to output variables that adequately explain pricing indices in regression models. Consequently, the final funnel step is a subset regression optimizing variable selection based on explanatory power, R^2, in a given model. Six linear regression models, that will be elaborated in Section 4.1, were used for the subset regression: two cointegration regressions using predictable $I(1)$ and diverse $I(1)$, two stationary regressions using predictable $I(0)$ and diverse $I(0)$, and two autoregressive regressions using predictable $I(0)$ and $I(1)$ and diverse $I(0)$ and $I(1)$. Finding the optimal R^2 for $n \in [1,10]$ variables selected out of $m \in \{13,20,27,40,44,64\}$ possible variables using brute force has time complexity $\Theta\binom{m}{n} = \Theta(n^m)$,[29] which quickly becomes intractable when m increases. To make the problem tractable, we use the *leaps* R package implementation of the branch and bound-algorithm proposed by Furnival & Wilson (1974), significantly reducing the search tree.

After performing subset regressions, we arrived at six models with 1 to 4 independent variables, and these models will be denoted *Ad, Ap, Dd, Dp, Rd* and *Rp* (i.e., Absolute (A), Differentiated (D), AutoRegressive (R), predictable (p) and diverse (d)). The benefit of the inevitable increase in R^2 with increasing number of variables is evaluated against risk of overfitting, significance of variables, Bayesian information criterion (BIC) and the cost of adding a variable to be predicted. See Figure 3.12 for plot of R^2 versus number of variables. A detailed explanation of the variable selection process from the subset regression is included in Appendix A.III.

The final set of variables crystallizing from our funnel turn out to be familiar variables with strong causal rational supporting the effect on multiples. See time series plots in Figure 3.13. Descriptive statistics for each of the final variables in the six models are displayed in Table 3.3, and it is evident the Box-Cox transformed variables are more normal, as the skew and excess kurtosis are close to 0. A critique of funnel approaches with such a large set of variables, is that the selected variables might end up being variables that have a large R^2s by coincidence, but will fail to generalize well, i.e., the models overfit,[30] but our variables have solid foundations in economic theory from Chapter 2:

Implied Volatility (vST50, vST50lr, vST50BCt) The Euro STOXX 50 Volatility (VSTOXX) is an index measuring implied volatility based on blue chip-option pricing of the fifty largest European stocks, obtained from Deutsche Borse AG and SIX Group. The index represents annualized 30-day forward-looking investor volatility expectation and is computed based on the option pricing model from Black & Scholes (1973). It is known to be a measure of investor fear and mirrors the U.S. market VIX. Reformulating the Gordon Growth model from Equation 2.2 and assuming that r_c for the market is defined by a constant Sharpe ratio $s = \frac{r_c - r_f}{\sigma_i}$, introduced in Chapter 2, we use a Taylor series expansion to obtain:

$$m_c = \frac{V_{ct}}{\pi_{c1}} = \frac{1}{r_c - g} = \frac{1}{s\sigma_i + r_f - g} = \frac{1}{r_f - g} - \frac{s}{(r_f - g)^2}\sigma_i + ... \tag{3.8}$$

[29] For elaboration on Big-Θ, Big-O and Big-Ω notation for algorithm time and space complexity, see Cormen et al. (2009). For $n = 10$, $m = 64$, the problem is of the order 10^{11}, and if one regression takes $1ms$, the problem is solved in ~5 years.

[30] A learned regression function h overfits if in-sample accuracy is better than out of sample. See Russell & Norvig (2003) for a definition within the "Probably Approximately Correct" framework, and discussions on consequences.

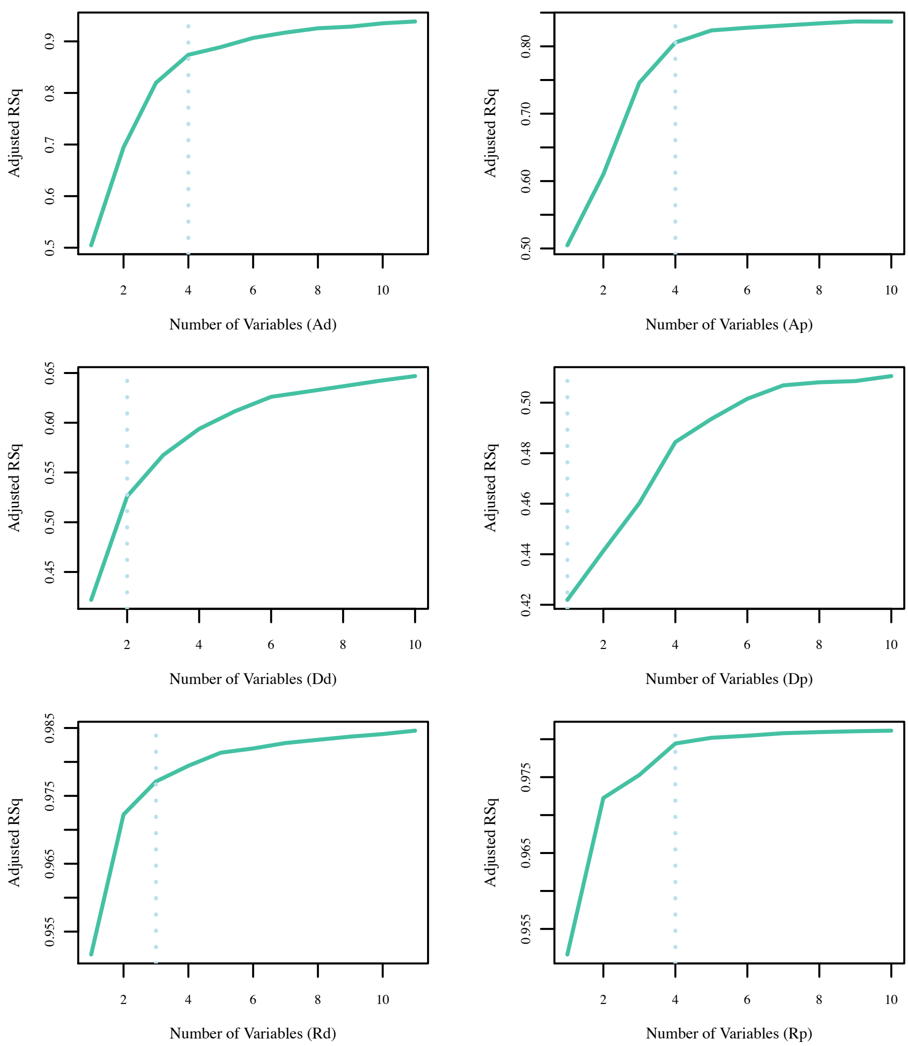


Figure 3.12: R^2 against number of variables in the six models

Variable	n	Mean	Standard Deviation	Median	MAD	Min	Max	Range	Skew	Excess Kurtosis	S.E
Ad											
vST50	223	23.99	9.01	22.11	7.17	11.99	61.34	49.35	1.48	2.54	0.60
sEUCC	223	-11.58	7.63	-11.30	7.41	-34.70	2.00	36.70	-0.39	-0.07	0.51
iEU10	223	2.84	1.67	3.23	2.00	-0.14	5.63	5.76	-0.27	-1.28	0.11
mEUFAYOY	223	-0.14	11.80	0.63	3.81	-42.02	25.51	67.53	-1.97	5.83	0.79
Ap											
vST50BCt	223	1.57	0.07	1.57	0.07	1.42	1.74	0.32	0.12	-0.43	0.00
mEUCPI	223	1.74	0.96	2.00	0.74	-0.70	4.10	4.80	-0.47	-0.22	0.06
iEUspd101	223	1.24	0.70	1.27	0.68	-0.16	2.72	2.88	0.01	-0.67	0.05
mEUpYOY	223	1.00	4.91	1.80	3.13	-21.40	9.46	30.85	-2.09	5.97	0.33
Dd											
vST50lr	223	0.00	0.18	-0.02	0.15	-0.52	0.65	1.17	0.36	0.57	0.01
sEUCCd	223	0.01	1.50	0.20	1.33	-5.30	3.90	9.20	-0.42	0.90	0.10
Dp											
vST50lr	223	0.00	0.18	-0.02	0.15	-0.52	0.65	1.17	0.36	0.57	0.01
Rd											
MedianEVEBITDAn1	223	8.14	1.26	8.04	1.44	5.25	10.39	5.14	-0.17	-0.72	0.08
vST50lr	223	0.00	0.18	-0.02	0.15	-0.52	0.65	1.17	0.36	0.57	0.01
sEUCCd	223	0.01	1.50	0.20	1.33	-5.30	3.90	9.20	-0.42	0.90	0.10
Rp											
MedianEVEBITDAn1	223	8.14	1.26	8.04	1.44	5.25	10.39	5.14	-0.17	-0.72	0.08
vST50lr	223	0.00	0.18	-0.02	0.15	-0.52	0.65	1.17	0.36	0.57	0.01
vST50	223	23.99	9.01	22.11	7.17	11.99	61.34	49.35	1.48	2.54	0.60
mEUCPIBCt	223	2.51	1.51	2.80	1.23	-0.61	7.01	7.62	-0.01	-0.08	0.10

Notes: vST50 - Volatility, iEUspd101d - Term strucure spread, mEUCPI - Inflation, mEUpYOY - Industrial production growth year on year, sEUCC - European Consumer Confidence, iEU10 - Eurozone 10 year interest rate, mEUFAYOY - EU Fixed assets growth, MedianEVEBITDAn1 - Lagged multiple, BCt - Box Cox transformed, lr - log-return, d - differentiated.

Table 3.3: Statistical properties of the independent variables

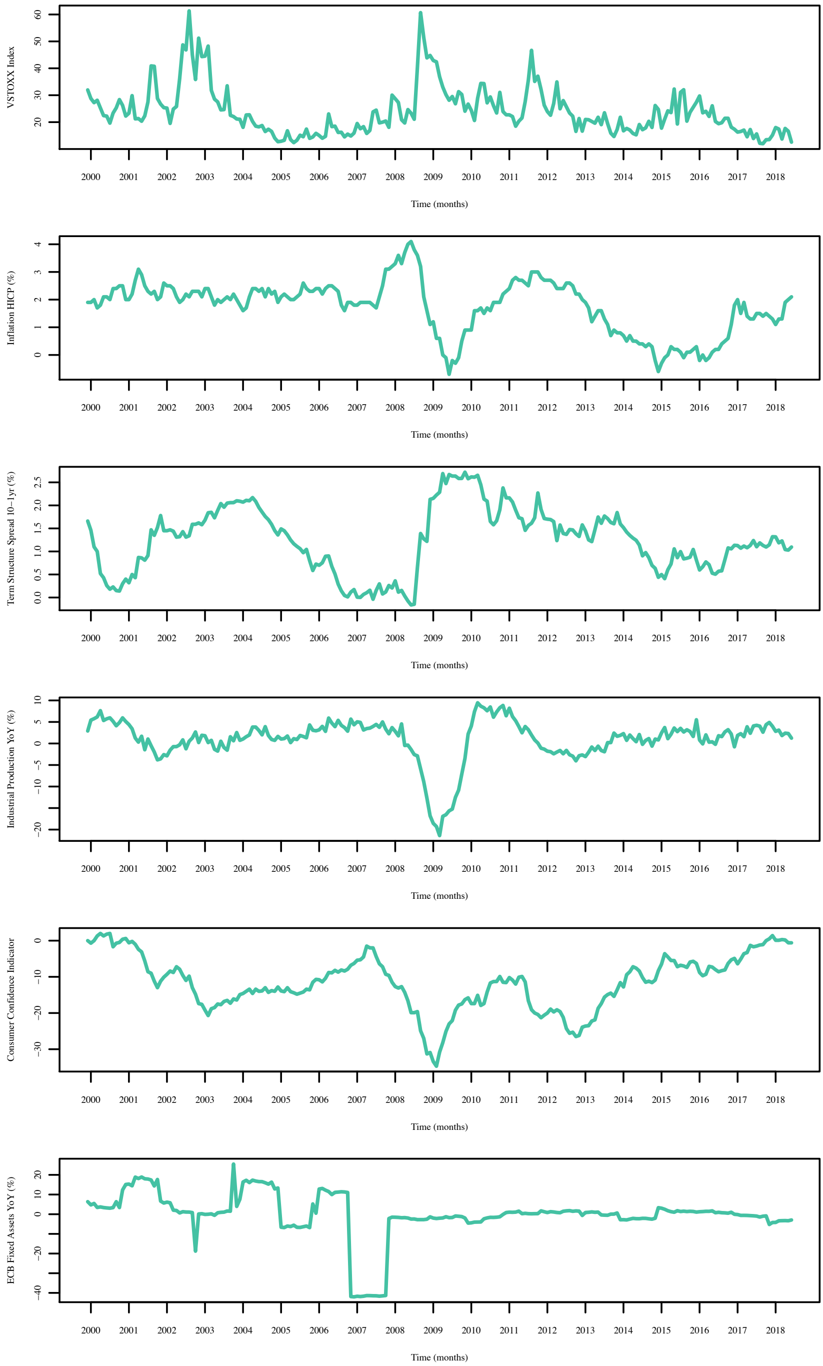


Figure 3.13: Time series plots of macroeconomic variables

The first term may be interpreted as a risk-free multiple valuation, while the second and consecutive perturbation terms represent the effect of implied volatility going forward. Hence a negative linear relationship may be expected,[31] an assertion that is supported by Kane et al. (1996) finding P/E-multiples to be closely associated with volatility.

Inflation (mEUCPI, mEUCPIBCt) In the Eurozone, consumer price inflation is measured by the Harmonised Index of Consumer Prices (HICP) reported by the European Central Bank (ECB). Most European policymakers (including ECB and Bank of England) attempt to maintain a target inflation through the monetary policy instruments, and although it is hard to control directly, targeting makes inflation a variable that investors may have an opinion on (Bernanke & Mishkin 1997). Assuming that the "money illusion" argument proposed by Modigliani & Cohn (1979) holds, Equation 2.4 may be augmented with a constant actual inflation rate i_a and a perceived inflation rate i_p. The multiple m_{in} of a firm will be valued as follows:[32]

$$m_{in} = \sum_{t=1}^{\infty} \frac{(1+g+i_p)^t}{(1+r_c+i_a)^t} = \frac{1}{r_c - g + i_a - i_p} = \frac{1}{r_c - g} + \frac{i_p}{(r_c-g)^2} - \frac{i_a}{(r_c-g)^2} + \ldots \tag{3.9}$$

Where the first term represents no-inflation value, the second term represents value increase from perceived cash flow increase, and the third component represents actual value decrease due to increased cost of capital from inflation. With this, the multiple is negatively impacted by inflation if the investor perception of inflation is incorrect, an assertion empirically supported by White (2000) and Zorn et al. (2008). The balance sheet inertia argument proposed by Koller et al. (2010) impacts short-term cash flows directly, as depreciation remains constant while investments increase, and the effect on multiples is trivially negative.

Term Structure Spread (iEUspd101) The spread between the yield of the 10-year and 1-year euro benchmark bond ($r_{spd} = r_{10} - r_1$) serves as a proxy for the slope of the yield curve and is provided by Tullett Prebon Information. The slope of the yield curve is a gauge for the investor expectation on the development of the economy (Kessel 1971). In high-activity business cycles, the yield will be high. Hence if the slope of the yield curve is steep, it is an indication that investors believe that we are moving towards a high-activity business cycle, and inverted or flat yield curve has been seen as recession predictors (Ang et al. 2004). The term structure spread is a proxy for the expected difference between the future and current interest rates and thus, is expected to be proportional to the cost of capital and negatively correlated with the multiple. A high term structure spread indicates that the future investment environment will be characterized by high yields and thus, increasing cost of capital and lowering multiples.

Industrial Production Growth (mEUpYOY) Annual industrial production growth in the Eurozone reported by Eurostat[33] measures changes in value added at factor cost of industry and construction. The variable is strongly connected to GDP and particularly future GDP and is an important variable in economics and monetary policymaking. If the current industrial production growth g_{ip} is assumed to be a leading variable for future GDP growth g_{GDP}: $GDP_t(1+g_{GDP}) = GDP_{t+1}$, and GDP growth is used as a proportional proxy for growth in Equation 2.2, the following

[31] vST50BCt is Box-Cox transformed with $\lambda = -0.5$, and a Taylor expansion of the equivalent, more complicated expression would yield a negative coefficient σ_{bcT} term. This argument is also valid for the box-cox transformed inflation-variable mEUCPIBCt with $\lambda = 1.5$.

[32] Note that g, i_p and i_a were assumed sufficiently small, such that $(1+g)(1+i) = 1+g+i+gi \approx 1+g+i$.

[33] Directorate of the European Commission providing statistical information to EU institutions.

relationship is obtained:

$$m_{ct} = \frac{1}{r_c - \alpha g_{ip}} = \frac{1}{r_c} + \frac{\alpha g_{ip}}{r_c^2} + ... \tag{3.10}$$

The first term is the no-growth multiple, while the second and consecutive terms represent the value of growth - illustrating that growth affects multiples positively.

Consumer Confidence Indicator (sEUCC) The consumer confidence indicator is a leading indicator of consumer confidence based on consumer surveys performed by Eurostat. It measures the financial situation of a household based on unemployment expectations, savings, and consumption, and multiples are expected to correlate positively with it because consumer spending drives revenues and profits.

ECB Fixed Assets Growth (mEUFAYOY) Annual growth of the aggregated balance sheet of fixed assets for the Eurozone reported by ECB measures changes in holdings of land and buildings, furniture and equipment. The Fixed assets constitute ~ 0.1% of total assets held by ECB as of August 2018. Central Bank balance sheets have been associated with returns in financial markets in previous studies (Meaning & Zhu 2011) and represent investment opportunity demand side - central bank security acquisition policies will increase demand for investment opportunities, raising multiples.

In Chapter 6, forecasts of variables from FactSet are used to build a base case scenario. Inflation is estimated based on a Eurozone analyst consensus estimate for 2018-2022. Industrial production growth estimates for the Eurozone were only available for 2018, 2019 and 2020 and were extrapolated linearly after 2020. Spread forecasts were calculated as the difference between a 10-year Government Bonds consensus estimate weighted by the share of total companies in each country, and the consensus estimate short-term interest rate. No volatility forecasts were available, and the base case was set to be at a slow increase from the current historically low level.

Chapter 4

Valuation Multiple Indices

In this chapter, we aim to understand how macroeconomic factors affect pricing indices by examining historical time series and applying relevant statistical models. These include linear regressions, quantile regressions, nonlinear regressions, machine learning methods and autoregressive moving average models with exogenous variables. The data sets are separated into dependent and independent variables, based on a perception of causality rooted in economic theory from Chapter 2.

4.1 Linear Regression Models

Linear regression is the bedrock in our analyses aiming to estimate the linear impact of chosen macroeconomic variables on valuation indices. Linearity is often effective, but might be constraining if the underlying relationships are complex. The models specified in Subsection 4.1.1 were used to select exogenous variables in the subset regression from Subsection 3.3.1, and those data sets, along with the indices, are the ones that will be scrutinized: *Ap*, *Ad*, *Dp*, *Dd*, *Rp* and *Rd*. In Subsection 4.1.3, the same underlying variables will be used on other indices than the chosen main index: *median EV/EBITDA*.

4.1.1 Model Specification

Pertinent unit-root stationarity considerations are imperative to avoid spurious regression, as stressed by Granger & Newbold (1974). Table 4.1 displays ADF-test statistics conducted on *Ap* variables in their absolute and differentiated forms.[1] For the differentiated variables, unit root stationarity is uncontroversial. For the undifferentiated variables, however, the unit root hypothesis cannot be rejected. As noted by Kwiatkowski et al. (1992), unit roots are hard to reject if processes are nearly integrated, and the stationary nature of macroeconomic variables is eagerly debated - see Nelson & Plosser (1982), Perron (1989) and Enders & Granger (1998). Consequently, an assumption of unit root stationary for the absolute variables would be controversial. Trend stationarity without structural breaks, however, is rejected by observing the coefficient obtained by including a time variable in the regression. Despite the possibility of structural shifts and more complex nonstationarity effects, one modelling hypothesis emerges as reasonable: The differentiated variables are $I(0)$ and the absolute variables are $I(1)$.

[1] ADF-tests, proposed by Dickey & Fuller (1981), uncover unit roots, i.e., $\gamma = 1$ for $y_t = \alpha + \gamma y_{t-1} + \epsilon_t$.

Variable	ADF test statistic	p-value
y_t	−2.80	0.240
iEUspd101	−2.56	0.339
mEUCPI	−3.34	0.064
mEUpYOY	−4.33	< 0.01
vST50BCt	−2.25	0.471
d_t	−5.74	< 0.01
iEUspd105d	−5.98	< 0.01
mEUCPId	−4.31	< 0.01
mEUpYOYd	−4.43	< 0.01
mST50lrBCt	−7.61	< 0.01
Critical values		
10%	5%	1%
-3.13	-3.42	-3.99

Table 4.1: ADF test statistics for selected variables

Table 4.2: Results from Johansen test

	Johansen Statistic			
	Test	10%	5%	1%
r ≤ 4	4.24	6.50	8.18	11.65
r ≤ 3	17.22	15.66	17.95	23.52
r ≤ 2	44.81	28.71	31.52	37.22
r ≤ 1	83.42	45.23	48.28	55.43
r = 0	127.09	66.49	70.60	78.87
Eigenvalues				
MedianEVEBITDA	iEUspd101	mEUCPI	mEUpYOY	vST50BCt
0.18	0.17	0.12	0.06	0.02

$I(i)$ variables may be regressed on $I(i)$ variables for $i > 0$ *iff* they are cointegrated (Engle & Granger 1987). Table 4.2 displays results from a Johansen cointegration test on Ap variables.[2] The rank of the Error Correction Model cointegration matrix used in the test is not ≤ 2 with $p < 0.01$, indicating that there are multiple linearly independent cointegrating vectors - i.e., the variables are cointegrated. This result is expected from the causal linear relationships between each of the underlying variables and y_t derived in Subsection 3.3.2, and indicates the presence of a long-term economic equilibrium process Alexander (2008*a*). Furthermore, it is natural to expect that the denominator π_{ct} is affected by the long-term environment with an inertia.

In the exogenous variable funnel from 3.3.1, multicollinearity was largely eliminated through selection based on correlation, rendering effects on estimators stability unlikely (Farrar & Glauber 1967). Table 4.3 displays the correlation matrix and the Variance Inflation Factors (VIF) for the variables used in Ap. The assumption that multicollinearity is unproblematic is affirmed by VIFs being < 10 (Farrar & Glauber 1967).

[2]See Johansen (1991) or Alexander (2008*a*) for Johansen test details, and Appendix B.IV for results for Ad. Johansen test for Ad is considerably less confident in determining cointegration, supporting Ap cointegration rationale. The implementation of the Johansen test in R's *urca*-package was used.

Table 4.3: Correlation matrix and Variance Inflation Factors for the independent variables in the *Ap* model

	iEUspd101	mEUCPI	mEUpYOY	vST50BCt
iEUspd101	1.00			
mEUCPI	-0.15	1.00		
mEUpYOY	-0.36	0.22	1.00	
vST50BCt	0.29	0.05	-0.33	1.00
VIF	1.20	1.08	1.28	1.19

OLS regressions are sensitive to outliers, and this gives rise to an inherent trade-off between the resulting estimator bias towards extreme values and ability to represent these events. Here, extreme events are included, as the knowledge about these is an essential part of the scope, but the impact was mitigated through differentiation and transformations in Section 3.3.1.[3]

Based on the preceding discussion we create six multivariate linear models, for each of the six final data sets.[4] Letting y_t be *median EV/EBITDA*, $\boldsymbol{x_t}$ be the independent variables given by *Ap* and *Ad*, ϵ_t be an error process for time t and α and $\boldsymbol{\beta}$ be regression coefficients, the following model is fitted for the *Ap*- and *Ad*- data sets, using OLS:

$$y_t = \alpha + \boldsymbol{\beta} \cdot \boldsymbol{x_t} + \epsilon_t \tag{4.1}$$

Table 4.4 shows the results from tests on normality, stationarity, autocorrelation, and heteroskedasticity for this and subsequent model residuals. An i.i.d.-assumption is clearly invalid, but the residuals are unit-root stationary, indicating that 4.1 is a cointegrating Engle-Granger regression (Tsay 2010). For *Ad*, the residuals are unit-root stationary with $p = 0.06$.[5] The residuals for both models are autocorrelated and heteroskedastic,[6] rendering estimators inefficient and t- and F-tests flawed, and indicating that there might be misspecification, possibly related to perturbation terms excluded in the relationships explored in Subsection 3.3.2. The residual autocorrelation in *Ap* will be studied closely in Section 4.4. The JB-test indicates that residuals of *Ap* fit well with the normality assumptions, while *Ad* residuals do not (Jarque & Bera 1980). This will not lead to any grave consequences, particularly when the residuals already have been shown to be autocorrelated and heteroskedastic. A natural model adjustment to deal with OLS-assumption issues is to use the differentiated data from the *Dd* and *Dp* data sets. The constant term is left out, and the model is fitted with OLS:

$$d_t = \boldsymbol{\beta_d} \cdot \boldsymbol{\Delta x_t} + \epsilon_t \tag{4.2}$$

where $d_t = \log y_t / y_{t-1}$, $\boldsymbol{\Delta} x_t$ are the differentiated exogenous variables from *Dd* and *Dp*-data sets, and ϵ_t is a conventional error process. The *Dx* residuals have no indications of unit roots,[7] but the issue of autocorrelation and heteroskedasitcity persist, especially for *Dp*. The major issue for these models is the loss of information about long-term relationships previously captured by the cointegration and the dependence of an initial value in long-term predictions. Because the differentiated models fail to include information about cointegration, and the absolute models fail to include the

[3]The quantile regression conducted in Section 4.2 is less sensitive towards outliers, but supports the main findings of this section.

[4]Note that these models are the models used in the subset regression in order to obtain the final data sets.

[5]A more confidently stationary series of residuals (ADF statistic = -4.299) is obtainable by using the Johansen test eigenvalues as regression coefficients. See Appendix E.I.

[6]Causes of autocorrelation and heteroskedasticity include omitted variables, structural breaks, inappropriate functional form and over- or under-differencing, and is prevailing in time series models (Alexander 2008*a*).

[7]A notational remark: *Dx*= {*Dp*,*Dd*}, *Ax*= {*Ap*,*Ad*}, *Rx*= {*Rp*,*Rd*}, *Xd* = {*Dd*,*Ad*,*Rd*} and *Xp* = {*Dp*,*Ap*,*Rp*}.

Test	H_0		*Ad*	*Ap*	*Dd*	*Dp*	*Rd*	*Rp*
JB-test	$\epsilon_t \sim N(\mu,\sigma)$	JB-statistic	13.2	1.9	6.5	34.2	2.93	0.01
		p-value	0.01	0.33	0.04	>0.01	0.20	0.99
ADF-test	ϵ_t has a unit root for 6 lags	ADF-statistic	-3.36	- 3.71	-5.55	-5.25	-5.30	-5.86
		p-value	0.06	0.02	<0.01	<0.01	<0.01	<0.01
LB-test	ϵ_t is not autocorrelated for 60 lags	Q-statistic	643.5	678.8	53.2	96.6	48.4	63.6
		p-value	<0.01	<0.01	0.725	<0.01	0.86	0.35
BP-test	ϵ_t is homoskedastic	Breusch-Pagan LM	1.09	6.10	3.62	9.47	2.09	0.39
		p-value	0.30	<0.01	0.06	<0.01	0.15	0.53

R implementation: Jarque Bera test - normtest-package, Augmented Dickey-Fuller test - tseries-package, Ljung-Box test - stats-package using lag = 60(forecast horizon), Breusch-Pagan test - car-package.

Table 4.4: Test statistics and p-values for selected residual tests. Results not compliant with OLS assumptions are underlined.

current information, a natural synthesis is the inclusion of an autoregressive term. Using the *Rp* and *Rd* data and an autoregressive coefficient γ, we propose:

$$y_t = \alpha_r + \gamma y_{t-1} + \boldsymbol{\beta_r} \cdot \boldsymbol{x_t} + \epsilon_t \tag{4.3}$$

The residuals of *Rx* do not have a unit root, and null hypotheses of non-autocorrelation, homoskedasticity and normality have not been rejected in the standard tests, suggesting that the estimators for these models are BLUE (Walpole et al. 2012). Both differentiated and absolute exogenous variables were included in the subset regression on *Rx* to account for the possibility of the nearly-integrated variables. In Section 4.1.2, modeling results and practical applications of the models will be discussed.

4.1.2 Modelling Results

Table 4.5 displays the results from the 6 regression models. The F-test statistics indicate with $p < 0.01$ that the overall R^2 values are significant for all models, but are flawed due to non-i.i.d. residuals. Consequently, a bootstrap estimation was conducted, yielding the confidence intervals for R^2 in square brackets, supporting the F-tests result with a high degree of confidence.[8] R^2s from different model classes are incomparable: *Ax*-R^2 represents ability to explain variance in y_t given $\boldsymbol{x}_t$, *Dx*-R^2 represents ability to explain variance in change of y_t given $\Delta\boldsymbol{x}_t$, and *Rx*-R^2 represents ability to explaining y_t given $\boldsymbol{x}_t$ and y_{t-1}. For prediction, *Ax*-R^2 will be independent on the horizon, while *Rx*-R^2 and *Dx*-R^2 will propagate forward in the prediction horizon. This incomparability necessitates other test methods for comparison.

For a further assessment of parameter robustness and prediction accuracy, a k-fold out-of-sample test was developed.[9] Figure 4.1 displays in-sample, 10-fold out-of-sample and 4-fold chronological out-of-sample results for *Ap*.[10] The 10-fold-prediction mitigates the impact of autocorrelation in

[8]The bootstrapping was done using methods from Efron (1979) and Efron & Tibshirani (1993), as implemented in the *boot*-package in R. The bootstrapping of R^2 and coefficient estimators are bias-corrected and accelerated (BCa), based on 1,000 simulations. This method does not account for residual autocorrelation, but more sophisticated techniques will be examined in Section 4.4.

[9]The k-fold testing algorithms are conducted by allocating the data into k partitions, before training the algorithm on $k-1$ partitions and predicting the final partition, providing a complete out-of-sample accuracy test on the sample. The allocation is random in the 10-fold test and chronological in the 4-fold test. Kohavi et al. (1995) review k-fold testing against bootstrapping.

[10]See Figure 13 to 17 in Appendix B.I for corresponding plots for *Dx*, *Rx* and *Ad*.

Table 4.5: Regression model coefficients for *Ap*, *Ad*, *Dp*, *Dd*, *Rp* and *Rd*. t-test standard error in parenthesis and BCa bootstrapped 95% confidence intervals in brackets

	Dependent variable:					
	Median EV/EBITDA		Log return median EV/EBITDA		Median EV/EBITDA	
	Ad	*Ap*	*Dd*	*Dp*	*Rd*	*Rp*
Median EVEBITDA 1-lag					0.98^{***}	0.90^{***}
					(0.01)	(0.02)
					[0.96, 1.00]	[0.87, 0.93]
vST50	-0.06^{***}					-0.01^{***}
	(0.004)					(0.002)
	[-0.06, -0.05]					[-0.02, -0.01]
vST50BCt		-8.99^{***}				
		(0.60)				
		[-9.92, -7.22]				
vST50lr			-0.13^{***}	-0.13^{***}	-0.98^{***}	-0.69^{***}
			(0.01)	(0.01)	(0.07)	(0.08)
			[-0.16, -0.12]	[-0.16, -0.12]	[-1.14, -0.83]	[-0.83, -0.54]
sEUCC	0.08^{***}					
	(0.004)					
	[0.08, 0.09]					
sEUCCd			0.01^{***}		0.06^{***}	
			(0.001)		(0.01)	
			[0.01, 0.01]		[0.04, 0.08]	
mEUCPI		-0.62^{***}				
		(0.04)				
		[-0.69, -0.56]				
mEUCPIBCt						-0.06^{***}
						(0.009)
						[-0.08,-0.04]
iEU10	-0.27^{***}					
	(0.02)					
	[-0.31, -0.25]					
iEUspd101		-0.49^{***}				
		(0.06)				
		[-0.61, -0.39]				
mEUpYOY		0.09^{***}				
		(0.01)				
		[0.07, 0.10]				
mEUFAYOY	-0.03^{***}					
	(0.003)					
	[-0.03, -0.02]					
Constant	11.26^{***}	23.90^{***}			0.14	1.34^{***}
	(0.10)	(0.92)			(0.09)	(0.17)
	[11.18, 11.51]	[21.06, 25.27]			[-0.03, 0.32]	[1.00, 1.71]
Observations	223	223	223	223	223	223
R^2	0.88	0.81	0.53	0.42	0.96	0.96
R^2 BCa conf. int.	[0.84, 0.91]	[0.77, 0.85]	[0.43, 0.63]	[0.31, 0.53]	[0.97, 0.98]	[0.97, 0.98]
Adjusted R^2	0.87	0.81	0.53	0.42	0.98	0.98
Residual Std. Error	0.45	0.56	0.03	0.03	0.19	0.18
F Statistic	385.69	231.10	124.82	163.91	3,161.16	2,643.51
Note:					*p<0.1; **p<0.05; ***p<0.01	

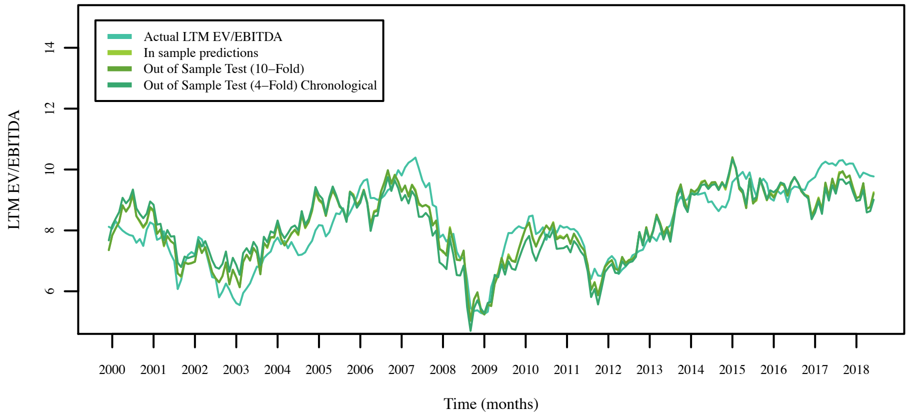


Figure 4.1: In and out-of-sample predictions for *Ap*-model

the explanatory and dependent variables through random sampling and thus, have negligible deviations from the in-sample testing, while the 4-fold chronological tests expose the bias caused by autocorrelation and is a more realistic representation of prediction accuracy. Nevertheless, deviations between 4-fold testing and in-sample prediction are small - an indication that the coefficients obtained are robust and that no prodigious trend shift has occurred during the time interval under consideration.

The k-fold-tests provide insight into the robustness of the coefficients, but do not assess the actual predicative capability. The required data to provide prediction of horizon h from time t differs: *Ax* requires $\boldsymbol{x}_{t+h}$, *Dx* requires $\boldsymbol{x}_t$ and y_{t+h}, while *Rx* is path-dependent[11] and requires y_t and $\boldsymbol{x}_{t:t+h}$.[12] An in-sample prediction test is developed to assess the models' abilities to predict y_{t+k} provided y_t and $\boldsymbol{x}_{t:t+k}$ using the original OLS coefficients.[13] For *Ax*, the predictions are simply:

$$\hat{y}_{t+k} = \hat{\alpha} + \hat{\beta} \cdot \boldsymbol{x}_{t+k} \tag{4.4}$$

while *Dx* predictions are initial value dependent:

$$\hat{y}_{t+k} = y_t e^{\hat{\beta}_d \cdot \Delta \boldsymbol{x}_{t+1:t+k}} \tag{4.5}$$

where $\Delta \boldsymbol{x}_{t+1:t+k} = \sum_{i=t+1}^{t+k} \Delta \boldsymbol{x}_i = \boldsymbol{x}_{t+k} - \boldsymbol{x}_t$. *Rx* are path- and initial value dependent:

$$\hat{y}_{t+k} = \hat{\alpha}_r \frac{1-\hat{\gamma}^k}{1-\hat{\gamma}} + \hat{\gamma}^k y_t + \sum_{i=1}^{k} \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+i} \hat{\gamma}^{k-i} \tag{4.6}$$

See Appendix B.II for proof by induction of Equation 4.6.

[11] Note that *Rx* may be made path-independent by making assumptions on the transition from $\boldsymbol{x}_t$ to $\boldsymbol{x}_{t+h}$.

[12] The notation $\boldsymbol{x}_{a:b} = \{\boldsymbol{x}_{t'}\}|(a \leq t' \leq b)$, borrowed from Russell & Norvig (2003) will be used consistently.

[13] The in-sample coefficients are used because we isolate predicative nature of the models from the coefficient-robustness issue, as tested in the bootstrapping and k-fold tests.

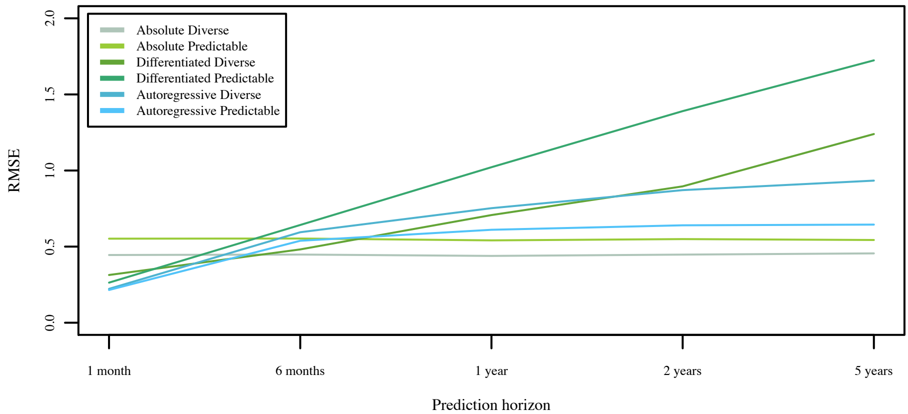


Figure 4.2: In-sample prediction RMSE given independent variable in horizon period

Figure 4.2 displays Root Mean Square Errors (RMSE)[14] for the six models on five prediction horizons. The differentiated and regressive models yield better results in short-term predictions as these models take the current state into account explicitly. In the long-run, however, the models fail to account for the development of the underlying variables properly and tend to revert to much to the historical mean of y_t, indicating that the cointegration relationship carries a predictive value. As expected, *Ax* performs equally well for all horizons, as it is in-sample. *Xp* performs well compared to *Xd*, despite fact that the set of variables from which *Xp* variables are selected is a subset of the corresponding variable set for *Xd* - advocating "predictable" dataset rationale. In fact, *Rp* actually performs better than *Rd*.

In Chapter 3, we established economic rationales for the linear dependencies materializing from the regression models, though we can only provide evidence of correlations and never causation. Implied volatility (vST50,vST50BCt,vST50lr) is a constituent of all models except *Rd* and has the largest explanatory power in individual-variable regressions.[15] Across all models, higher implied volatility cause lower valuation multiples. This is consistent with Sharpe (1964) and Equation 3.8, implying that higher risk will lead to higher risk premiums and discount rates and thus, lower valuations. This result is empirically supported by Kane et al. (1996), showing that increase of volatility historically has caused decline in P/E multiples.

Interest rates appear in the models in two forms: The long-term rate (iEU10) in *Ad* and as term structure spread (iEUspd101) in *Ap*. The coefficient of the iEU10 rate is negative, supporting the hypothesis that interest rate increases make fixed income more attractive as investment opportunities, moving demand away from equities and lowering returns, as well as increasing cost of capital. The effect is amplified in the term structure spread because it reflects the expected change relative to the current level (Campbell 1987). Notably, this effect is only present in the models leveraging on cointegration effects, indicating that the level of the interest rate is a driver per se, not only the changes.

[14] RMSE $= \frac{1}{\sqrt{h}}\sqrt{\sum_{t'=t}^{t+h}(\hat{y}_{t'} - y_{t'})^2}$ measures mean deviation, as opposed to MAD $= \frac{1}{h}\sum_{t'=t}^{t+h}|\hat{y}_{t'} - y_{t'}|$ measuring median deviation. RMSE is considered more appropriate because it mirrors the OLS-estimation residual square minimization and weights large deviations more than minor deviations. Both techniques are conventional, e.g., see Shamsuddin & Hillier (2004). MAD-aggregation for the test yields similar results in Appendix B.III.

[15] See Appendix A.III for details about subset regression results, including single variable regressions.

Industrial production growth (mEUpYOY) was proposed as a leading indicator of the GDP in Chapter 2 and appears in *Ap* with a positive coefficient. The causal hypothesis is that increased mEUpYOY leads to increased GDP expectations, which again improves earnings expectations for the firms, hiking valuations - see Equation 3.10. Inflation (mEUCPI) is the variable with the second most explanatory power in the *Ap* regression, and it has a negative coefficient, supporting the "money illusion" argument proposed by Modigliani & Cohn (1979) and concretized in Equation 3.9, and the Koller et al. (2010) argument on the effect of inflation on FCF discussed in Chapter 2.

Two other variables, deemed non-predictable, had high explanatory value in the *Xd*-models: EU Consumer Confidence Indicator (sEUCC, sEUCCd) and ECB Fixed Assets Growth (mEUFAYOY). sEUCC reflects the market sentiment, which is very likely to be associated with multiples. Illustratively, regressing multiples on number of stock broker calls mentioning the words "high multiple" would have a large explanatory power, but similarly to sEUCC, the predictability of this variable is low, and it is unlikely that investors have sophisticated views on development prospects. As indicated in Chapter 2, ECB balance sheet variables have a causal rationale associated with Quantitative Easing policies. The mEUFAYOY, however, reflects 0.1% of the total balance sheet as of August 2018, and the coefficient sign contradicts the intuition that heavier central bank balance sheets should drive up prices. Hence, we believe it is a statistical coincidence.

In terms of OLS assumption compliance, *Rp* is superior to *Dp* and *Ap*, both having significant issues with heteroskedasticity and autocorrelation. *Rp* explains more of the in-sample variance, but is path dependent and fails to generalize well for long-term prediction tests. *Ap* is better for predictions on horizons of 1 year or longer, is path and initial value independent, has sufficient in sample accuracy, is easy to interpret, has strong causal interpretations of the exogenous variable, and has robust coefficients and overall regression significance. Based on this, *Ap* will be the primary model going forward.

4.1.3 OLS on Other Indices

A selection of $\aleph$s and estimations of V_{ct} and π_{ct} were discussed in Section 3.2, and we argued that median EV/EBITDA is well-suited for measuring overall firm valuation levels. In this subsection, we examine how other multiple indices are affected by the four macroeconomic variables surfacing from *Ap*. The cointegration arguments and the corresponding Engle-Granger regressions (4.1) are assumed to hold for these indices, and the residual indecorousness persists. Table 4.6 displays results from normalized OLS regressions and the coefficients are visualized in Figure 4.3.

All F-tests indicate R^2 significance, and t-test significant coefficients are consistent with *Ap* - supporting the fundamental causality. The categorical median EV/EBITDA introduced in Section 3.2.1 yields coefficients practically indistinguishable from the ordinary median, giving us comfort in median EV/EBITDA despite the bias created by invalid multiple omission. The volatility coefficient in EV-weighted and EBITDA-weighted indices are insignificant, indicating that larger companies are less sensitive to changes in volatility. This hypothesis is supported by a supplementary analysis in Appendix B.V that also reveals that larger companies are less disturbed by inflation and more by growth and term structure spread. The elimination of volatility as an explanatory variable reduce R^2 considerably for these models.

The coefficients vary less among the indices based on other V_{ct}s and π_{ct}s. Variations are as expected from previous discussions: EV-multiple coefficient magnitudes vary with how far down on the income statement the denominator is, and P/E coefficients are more sensitive to growth and less sensitive to cost of capital effects, as the latter is mitigated in the long run by the denominator which is net of interest. The indices bridge our results to existing literature (White 2000) and provide some additional nuances into the nature of valuation multiples. However, we maintain that *median EV/EBTIDA* serves as a representative and relevant index, and that our approach in Chapter 5 will reveal a comprehensive and profound perspective on the nuances.

Table 4.6: Regression results for selected indices

	Dependent variable:							
	EV/EBITDA					**Other median multiples**		
	Median	Average	EV Weighted	EBITDA Weighted	Categorical	EV/Sales	EV/EBIT	P/E
iEUspd101	−0.268***	−0.312***	−0.417***	−0.379***	−0.271***	−0.351***	−0.145***	−0.184***
	(0.032)	(0.038)	(0.060)	(0.056)	(0.033)	(0.037)	(0.030)	(0.033)
mEUCPI	−0.475***	−0.422***	−0.275***	−0.342***	−0.490***	−0.394***	−0.500***	−0.378***
	(0.031)	(0.036)	(0.057)	(0.053)	(0.031)	(0.035)	(0.029)	(0.031)
mEUpYOY	0.335***	0.383***	0.314***	0.372***	0.337***	0.300***	0.398***	0.430***
	(0.034)	(0.039)	(0.062)	(0.058)	(0.034)	(0.038)	(0.031)	(0.034)
vST50BCt	−0.489***	−0.397***	0.038	−0.041	−0.473***	−0.451***	−0.518***	−0.511***
	(0.032)	(0.038)	(0.059)	(0.056)	(0.033)	(0.037)	(0.030)	(0.033)
Observations	223	223	223	223	223	223	223	223
R^2	0.81	0.74	0.35	0.43	0.80	0.75	0.83	0.80
Adjusted R^2	0.81	0.74	0.34	0.42	0.80	0.75	0.83	0.80
Res. Std. Error	0.44	0.51	0.81	0.76	0.45	0.50	0.41	0.45
F Statistic	231.10***	155.82***	29.86***	40.67***	223.14***	163.83***	269.42***	222.50***

Note: *p<0.1; **p<0.05; ***p<0.01

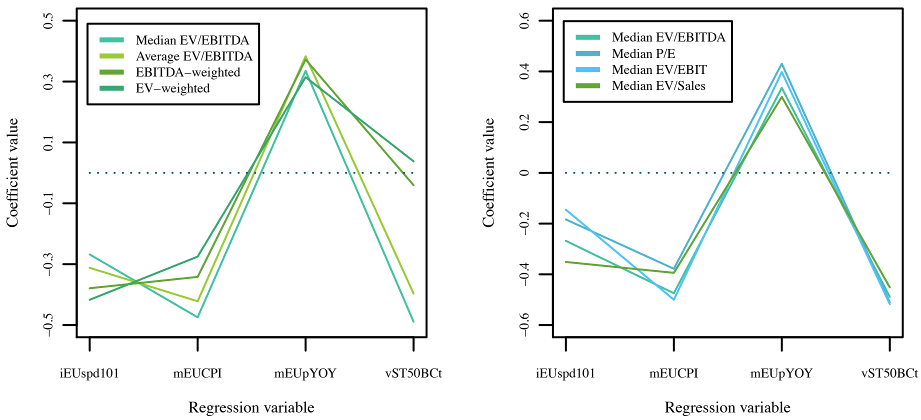


Figure 4.3: Regression coefficients for normalized variables - seven indices

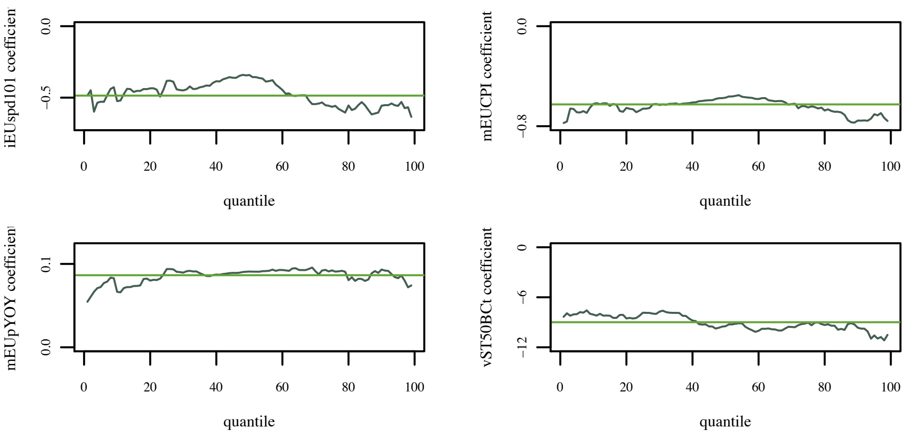


Figure 4.4: 99-Quantile regression coefficients for each of the exogenous variables. Green lines are cofficient from regular *Ap*

4.2 Quantile Regression Models

As shown by Gauß (1821), conditional mean estimated by OLS is the best estimator of the distribution location parameter if the distribution is normal or normal-like. In many real-world applications, however, the distribution is affected by anomalies and outliers, and quantile-based error distributions descriptions may be more appropriate (Koenker & Bassett 1978). Furthermore, a quantile regression allows us to understand the extreme cases (i.e., the quantiles) of development and is less sensitive towards outliers. If the cumulative distribution of y_t is specified by $F(m)$, then the quantile regression model is the following:

$$F^{-1}(q|\boldsymbol{x_t}) = \alpha_q + \boldsymbol{\beta_q} \cdot \boldsymbol{x_t} + F_\epsilon^{-1}(q|\boldsymbol{x_t}) \tag{4.7}$$

We estimate the parameters by solving the corresponding optimization problem using the Barrodale and Roberts algorithm as described by Koenker & D'Orey (1987) and implemented in the R package *quantreg*.[16]

Figure 4.4 displays the quantile regression coefficients for each of the underlying variables for all 99 quantiles. Although data availability, particularly for the extreme quantiles, causes fluctuations, the plots reveal a few things about the distribution of responses to the underlying variables and hence, produce an indication of parameter robustness. The general level of the coefficients corresponds well to the magnitude and sign of the coefficients seen in Section 4.1, which is marked as a green line. There is a slight tendency of higher quantiles having larger coefficient magnitudes, reflecting the fact that the dispersion of the distribution of the error is dependent on the absolute level of the index.

Using these coefficients along with the current observations ($t = 223$) we can construct a proxy for the current conditional cumulative distribution of y_t: $P(y_t \geq m|\boldsymbol{x}_t)$ displayed in Figure 4.5.[17] The cumulative distribution provides valuable insight into the dispersion of y_t, conditional on $\boldsymbol{x}_t$ at

[16]Optimize: $\text{argmin}_{\beta_q,\alpha_q} \sum_{i=1}^{n} \Lambda_\tau(y_i - (\alpha_q + \beta_q \cdot \boldsymbol{x}_t))$ where Λ is a loss function: $\Lambda_\tau(u) = u(\tau - \mathbb{1}_{u<0})$.

[17]The cumulative density function is computed as such: $P(y_t \geq q|\boldsymbol{x}_t) = \alpha_q + \boldsymbol{c}_q\boldsymbol{x}_t$ where $(\alpha_q, \boldsymbol{c}_q)$ is the coefficient vector for the q'th quantile. $P(y_t \geq q|\boldsymbol{x}_t)$ is not generally a valid cdf, as it may be non-monotonic. The issue may be tackled e.g., by fitting a monotonic spline, decreasing quantile resolution or increasing sample size.

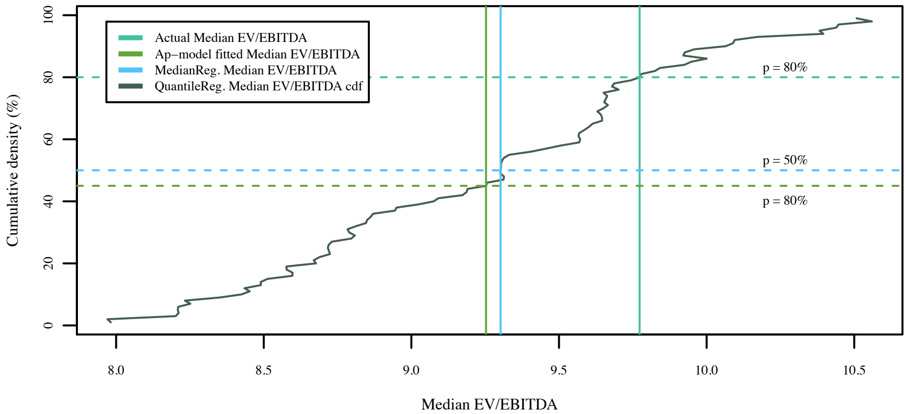


Figure 4.5: Quantile regression approximation of the cumulative density function of median EV/EBITDA at $t = 223$

a given instance of the process. y_t is not drawn independently from this cdf because the residuals are autocorrelated, but will move up and down in the distribution in a structured process. Notably, the current median EV/EBITDA is at the estimated 80% quantile of the cdf approximation, suggesting that the current valuation of enterprises is at a high level given underlying macroeconomic conditions. Assuming that the residuals are unit-root stationary, the multiple is ceteris paribus expected to revert to the median along the cdf. We will elaborate on the nature of the reversion in Section 4.4 and investigate the forecasting aspect in Chapter 6.

4.3 Nonlinear Regression Models

Nonlinear relationships between $\boldsymbol{x}_t$ and y_t would be reasonable, observing the pertubation terms from the Taylor series that were ignored in the linear regressions. There are, however, no conspicuous non-linearities in Figure 4.6, and a linear regression of the residuals on third degree Taylor series[18] and a reset-test[19] fail to reveal any momentous nonlinear relationships. Nevertheless, we do some preliminary testing of conventional machine learning methods aiming to capture other aspects of the relationship.

Figure 4.7 shows results from 4-fold chronological out of sample testing corresponding to Figure 4.1 for selected machine learning algorithms.[21] We tested k-Nearest Neighbors (kNN) with 1 and 20 neighbors, a Support Vector Machine (SVM) and an Artificial Neural Network (ANN). The 4-fold

[18] A subset regression was conducted on the following generalized regression model: $u_t = \sum_{i=0}^{4}\sum_{j=i}^{4}\sum_{k=j}^{4} b_{ijk}x_i x_j x_k + \epsilon_t$ where $x_0 = 1$, $x_{1:4}$ are the exogenous variables and u_t are residuals from *Ap*. Taylor polynomials approximate many nonlinear functions well (Kreyszig 2010). See analysis results in Appendix B.VII.

[19] A Ramsey RESET test was conducted using the R package *lmtest*. See results in Appendix B.VI.

[20] Note that extreme growth values stemming from GFC were omitted from the scatter plot.

[21] See Russell & Norvig (2003) for elaboration on the respective methods. The kNN algorithms are implemented with the *FNN* package. The ANN is trained using backpropagation for a 3-layer implementation with 6 hidden neurons as implemented in *neuralnet* from R. The SVM algorithm is retrieved from the *e1071* package but without sophisticated fine-tuning of the parameters.

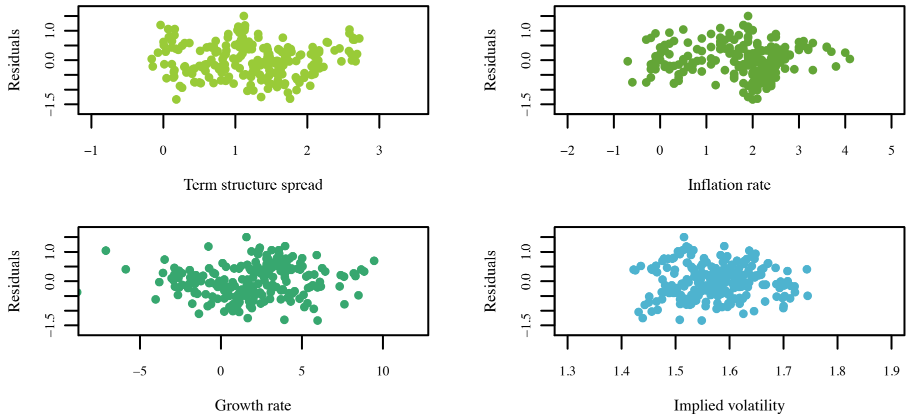


Figure 4.6: Scatter plots: *Ap* regression residuals[20]against independent variables in *Ap*.

out of sample RMSE values are 1.28, 1.21, 0.98 and 0.98, respectively, compared to 0.70 for *Ap*. The models do better than *Ap* in-sample but fail to generalize due to overfitting. The linear regression, with five parameters, is able to capture ~ 81% of the variance, while the machine learning algorithms are more flexible, and with a sample of 223 data points overfitting is hard to avoid. SVM and ANN do better than kNN methods, and by applying appropriate kernel functions and fine-tuning parameters, it might be possible to surpass the performance of linear regression.[22] However, *Ap* is versatile, simple, transparent, and interpretable, and wielding Occam's razor, we are unable to pick any long low hanging AI-fruits (Blumer et al. 1987).[23]

4.4 Mitigating Consequences of Residual Autocorrelation

The problems caused by autocorrelation in the residuals of *Ap* were circumvented in Section 4.1. However, autocorrelation introduce some problems in forecasting and bootstrap estimation. The residual process of Equation 4.1 may be approximated by applying a standard AutoRegressive 1 lag model $AR(1)$ (Tsay 2010):

$$\epsilon_t = \gamma\epsilon_{t-1} + u_t \tag{4.8}$$

in which u_t is assumed to be an i.i.d. error process, and ϵ_t is the residual process from Equation 4.1. Letting $\hat{\epsilon_t} = y_t - \hat{y}_t$ represent ϵ_t, Equation 4.8 can be approximated:

$$\hat{\epsilon}_t = \gamma^a\hat{\epsilon}_{t-1} + u^a_t \tag{4.9}$$

in which γ^a and u^a_t and the associated $\hat{\gamma}^a$ and $\hat{u}^a_t$ are all estimators for γ and u_t. OLS estimation of Equation 4.9 gives $\hat{\gamma}^a = 0.796$ and a set of satisfyingly uncorrelated residuals $\hat{u}^a_t$, see ACF-plot in Figure 30 in Appendix D.IV.

The model in Equation 4.9 may be used to amend bootstrap confidence intervals for the *Ap* coefficients, using the recursive algorithm proposed by Li & Maddala (1997).[24] As seen in Figure 4.8,

[22]The ANN can become exactly equivalent to the linear regression by constructing a zero-hidden layer network with linear activation functions and a constant input node.

[23]See Appendix E.I for elaboration on machine learning opportunities.

[24]For $s \in [1:10000]$: Draw $T-1$ random $\hat{u}^a_t$. Generate correlated residuals for $\hat{\epsilon}^s_{1:T}$, and let $y^s_{1:T} = \hat{y}_{1:T} + \hat{\epsilon}^s_{1:T}$. Conduct OLS on $y^s_{1:T}$ with $\boldsymbol{x}_{1:T}$, obtaining coefficient estimates $\boldsymbol{\beta}_s$. Finally, $\boldsymbol{\beta}_{1:S}$ is the bootstrap sample.

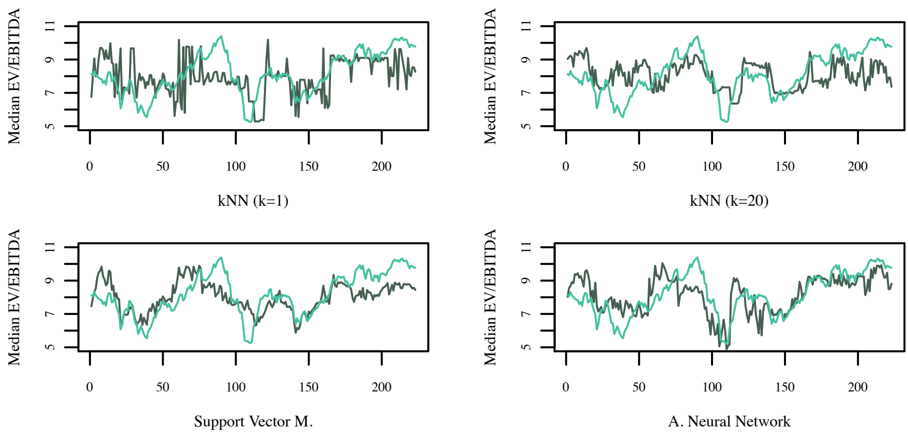


Figure 4.7: 4-fold out of sample test of selected machine learning algorithms. In-sample actual values in turquoise

the coefficient values are still significant ($p < 0.01$) but the estimator variance increases substantially.

Ap fails to account for the current state of the index in forecasting, but using Equation 4.9, the forecast can be adjusted according to the state of the residual process. For h horizon prediction, Equation 4.4 becomes:

$$\hat{y}_{t+k} = \hat{\alpha} + \hat{\beta} \cdot \boldsymbol{x}_{t+k} + (\gamma^{a})^{h} \hat{\epsilon}_t \tag{4.10}$$

and this model serves as the basis for endogenous and Monte Carlo based forecasts in Chapter 6.

Minimizing $\sum_{t=1}^{T} (\hat{u}_t^{a})^2$ given $\sum_{t=1}^{T} \hat{\epsilon}_t^2$-optimal parameters do not yield the maximum likelihood parameters for the whole system. Cochrane & Orcutt (1949) propose an algorithm resulting in minimal-$\sum_{t=1}^{T} \hat{u}_t^2$ parameters, and the R implementation in the *orcutt* package was used to estimate the parameters. A simpler way to estimate the residual process and the exogenous effects simultaneously is to use a Moving Average with K lags *MA(K)* description:

$$\epsilon_t = \eta_t + \sum_{k=1}^{K} a_k \eta_{t-k} \tag{4.11}$$

in which η_t are error processes and K is an integer constant. The *MA(K)* process is fairly successfully in mitigating the autocorrelation for $K \geq 5$, as seen in seen in Figure 4.9. The MA-eXogenous *MAX(K)* model obtained from combining Equation 4.11 with 4.1 can be estimated using an ordinary MLE approach:[25]

$$y_t = \alpha + \boldsymbol{\beta} \cdot \boldsymbol{x}_t + \eta_t + \sum_{k=1}^{K} a_k \eta_{t-k} \tag{4.12}$$

The resulting coefficients are displayed in Table 4.7. The coefficients are familiar, but the magnitude and significance are weaker, particularly for the *Cochrane-Orcutt* and the high-K MAX-models. The volatility variable is the only variable with consistent $p < 0.01$. Ljung-Box statistics are improved from LB $\chi^2 = 678.8$ for *Ap*, but the RSS has increased from 68.0 to values between 72.7 for $K = 1$ and 268.3 for $K = 20$. The models that mitigate autocorrelation well also mitigate the explanatory power of the underlying variables and render the error process prominent in describing the system. An in-sample fixed-starting point test[26] was conducted, revealing the depleted forecasting

[25] We use the algorithm proposed by Nelder & Mead (1965) as implemented in the *stats* package in R.
[26] Provided y_t and $\boldsymbol{x}_{t+h}$, predict y_{t+h}.

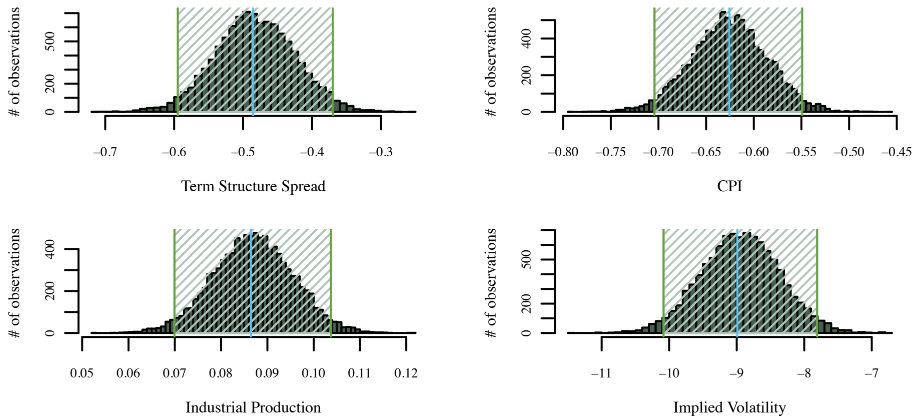


Figure 4.8: Bootstrap 95% confidence intervals for Ap regression coefficients, estimated using recursive bootstrap algorithm for 10,000 samples

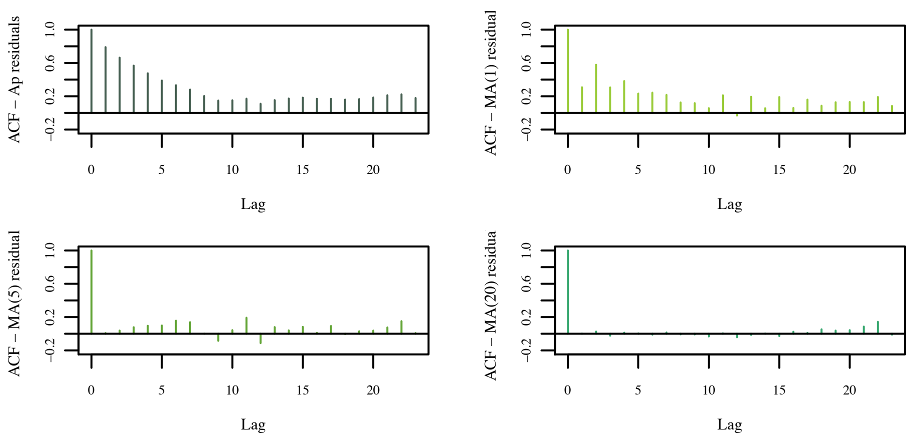


Figure 4.9: ACF values for original residuals in Ap, and $\hat{\eta}_t$-residuals for $K = 1$, $K = 5$ and $K = 20$

ability, see RMSE values in the Table 4.7. The corresponding RMSE for *Ap* is 0.55.

4.5 Summary

The performance of a range of models has been explored, and *Ap* is intuitive, simple, accurate and profoundly rooted in economic theory. The underlying variables are cointegrated, and the model is able to account for the level of the underlying variables as well as the change. Other models, such as the differentiated model, the autoregressive model, and models using other indices than median EV/EBITDA generally support the findings of *Ap*, but they have their respective weaknesses, both quantitatively and intuitively. Quantile regression, machine learning models and vector models contribute in deepening our understanding but do not contradict fundamental discoveries of *Ap*. As a result, we proclaim volatility, inflation, term structure spread and industrial production growth as paramount macroeconomic variables for explaining variation in median valuation multiples historically.

Table 4.7: Resulting coefficients from ARIMAX model estimation

	Model:			
	Cochrane-Orcutt	*MAX-K*		
		$K = 1$	$K = 5$	$K = 20$
ar-1	0.981			
ma-1		0.683***	1.300***	1.147***
		(0.047)	(0.067)	(0.080)
ma[3-5]			[0.501-1.415]***	[1.410-1.661]***
ma[6-19]				[0.502-1.599]***
ma-20				0.192**
				(0.095)
intercept	15.675***	21.440***	16.076***	15.641***
	(0.969)	(0.983)	(0.614)	(0.648)
iEUspd101	−0.006	−0.521***	−0.141*	−0.016
	(0.080)	(0.069)	(0.084)	(0.089)
mEUCPI	0.013	−0.560***	−0.102*	0.035
	(0.058)	(0.048)	(0.061)	(0.051)
mEUpYOY	0.007	0.073***	0.005	−0.011
	(0.009)	(0.010)	(0.010)	(0.009)
vST50BCt	−4.701***	−7.462***	−4.813***	−4.782***
	(0.395)	(0.630)	(0.384)	(0.392)
Observations	223	223	223	223
Log Likelihood		−116.275	−11.332	46.714
σ^2		0.166	0.064	0.037
AIC		246.550	44.665	−41.427
$\sum_{t=1}^{T} \hat{\epsilon}_t{}^2$	251.9	72.7	206.9	268.3
LB χ^2	59.0	352.2	224.5	35.6
LB p-value	0.51	< 0.001	< 0.001	0.995
Prediction RMSE	1.09	0.57	0.98	1.13

Note: LB test is with lag = 6 *p<0.1; **p<0.05; ***p<0.01

Chapter 5

Distribution of Valuation Multiples

In the preceding chapter, drivers of the proposed pricing indices were analyzed. In the creation of these indices, a distribution of data points is condensed to a single index data point at every time step t. In this chapter, we exploit the information available in the complete distribution of multiples and examine how it is affected by the exogenous variables[1]. Understanding the dynamics of the distribution gives a profound understanding of the complete *state* of the valuation market and facilitates in-depth analyses of risks and prospects. Relatively high valued and low valued companies are affected by movements in the underlying variables and by each other in distinct ways, not necessarily affecting the produced indices. Fundamentally, we seek to map a vector of exogenous variables $\boldsymbol{x}_t$ to a distribution, which is a learning problem in which the regression function to be learned is H:

$$H : \mathbb{R}^k \to [\mathbb{R}^+ \to \mathbb{R}^+] \tag{5.1}$$

where $\boldsymbol{x_t} \in \mathbb{R}^k$ and the distribution of a stochastic variable M_t is $f_t(m) \in \mathbb{R}^+$. M_t represents the multiple of any company at time t.[2] H may be described and approximated in multiple ways: e.g., by using a joint probability density function (pdf),[3] by describing the corresponding cumulative density function (cdf),[4] or by providing information about the quantiles or frequencies. The goal here is to develop methods ultimately approximating H, such that we can obtain a complete distribution of multiples requiring only one input vector of macroeconomic variables at time t.

The distribution of M_t is the quotient distribution of two other stochastic variables, the company value V_t and the profit metric π_t. Naturally, the value of a company depends on the size, i.e., $P(V_t|\pi_t) \neq P(V_t)$, so the distributions cannot be modeled separately. Consequently, the linear model in Equation 5.2 is assumed to hold at each intersection t, where α_t and β_t are constants and ϵ_t is a cross-sectional i.i.d. error variable which will turn out to be of immense importance.[5]

$$\log V_t = \alpha_t + \beta_t \log \pi_t + \epsilon_t \tag{5.2}$$

[1] For consistency, the independent variables from *Ap* will be used in this chapter.

[2] $m \in \mathbb{R}^+$ and $\int_0^\infty f(m)dm = 1$. The random multiple variable M_t is defined in $\mathbb{R}^+$.

[3] $H : \mathbb{R}^k \times \mathbb{R}^+ \to \mathbb{R}^+$, i.e., $H(\mathbf{x}_t, m)$ such that for all $\mathbf{x}_t$, $\int_0^\infty H dm = 1$, and $\int_\gamma^\eta H(\mathbf{x}_t, m)dm$ is the fraction of companies having multiples between γ and η.

[4] $H_c : \mathbb{R}^k \times \mathbb{R}^+ \to [0,1]$, i.e., $H_c(\mathbf{x}_t, m)$ such that for all $\mathbf{x}_t$, $\lim_{m\to\infty} H_c(\mathbf{x}_t, m) = 1$, $H_c(\mathbf{x}_t, 0) = 0$ and $H_c(\mathbf{x}_t, \eta) - H_c(\mathbf{x}_t, \gamma)$ is the fraction of companies having multiples between γ and η given $\mathbf{x}_t$. $H_c(\mathbf{x}_t, m) = \int_0^m H(\mathbf{x}_t, m')dm'$.

[5] The model is motivated by seemingly linear bi-variate plots historically, see chart in Appendix C.I.

The parameters are estimated using OLS. The mean β_t is 0.94 with a sample standard deviation of 0.028 over the time horizon under consideration.[6] The fact that $\beta_t < 1$ indicates that large companies have low valuations relative to their π_t. This may be associated with the growth cycle model proposed by Scott & Bruce (1987), indicating that larger companies tend to reach a point where growth prospects are smaller and thus, valuations decline.[7] Now, applying the Copernican Turn from Equation 2.7, we can derive the distribution of M_t based on the regression model:

$$M_t = \frac{V_t}{\pi_t} = \pi_t^{\beta_t - 1} e^{\alpha_t} e^{\epsilon_t} \tag{5.3}$$

If the β_ts were equal to 1, M_t would only be dependent upon the error process ϵ_t. The distribution would be the log distribution of ϵ_t, shifted by the constant e^{α_t}. A JB-test rejects that the residuals, $\hat{\epsilon}_t$ from Equation 5.2, stem from a normally distributed process.[8] $\hat{\epsilon}_t$ is bell-shaped, highly leptokurtic and slightly skewed. The average sample excess kurtosis is $\bar{\hat{\kappa}} = 8.56$ with a standard deviation of 6.23 and the average sample skewness is $\bar{\hat{S}} = 0.50$ with a standard deviation of 0.72. A conventional solution to leptokurticity is to assume that the error process follows a generalized Student-t distribution.[9] By testing the generalized Student-t distribution hypothesis on historical data, we are not able to reject the hypothesis that residuals are t distributed or equivalently that the multiple $(e^{\alpha_t} e^{\epsilon_t})$ is log t distributed.[10]

However, we cannot neglect that $\beta_t \neq 1$ and thus, $P(M_t|\pi_t) \neq P(M_t)$. The complete distribution of M_t as a function of the distributions of π_t and ϵ_t may be found by applying the Fourier Convolution theorem:

$$\begin{aligned} &\log M_t - \hat{\alpha} = (\hat{\beta}_t - 1)\log \pi_t + \epsilon_t \\ &f_M(e^{z+\hat{\alpha}}) = f_{\epsilon_t} * f_{(\beta_t - 1)\log \pi_t}(z) = \mathscr{F}^1\{\hat{f}_{\epsilon_t} \cdot \hat{f}_{(\hat{\beta}_t - 1)\log \pi_t}\} \end{aligned} \tag{5.4}$$

where $\hat{f}_{\epsilon_t}$ and $\hat{f}_{(\hat{\beta}_t - 1)\log \pi_t}$ are the Fourier transformations of the probability distributions f_{ϵ_t} and $f_{(\hat{\beta}_t - 1)\log \pi_t} (= g)$, and $\mathscr{F}$ represents the Fourier transformation operator. Assuming that $f_{\log \pi_t}$ is relatively well behaved,[11] it is easy to see that:

$$\lim_{\beta_t \to 1} g(z) = \delta(z) \tag{5.5}$$

where $\delta(z)$ is the Dirac delta function. Now, the limiting Fourier transform becomes:

$$\lim_{\beta_t \to 1} \int_{-\infty}^{\infty} g(z) e^{-2\pi i \xi z} dz = \int_{-\infty}^{\infty} \delta(z) e^{-2\pi i \xi z} dz = 1 \tag{5.6}$$

[6] See charts in Appendix C.II for an overview of the β_t-coefficients at each time step. All p-values are <0.01.

[7] Also Rostow (1959) and Chandler (1962) are relevant for stages of growth.

[8] See discussion on Jarque & Bera (1980) tests in Section 4.1.1 and Appendix C.III for plot of test statistics. All p-values are <0.01, rejecting the null hypothesis of zero skewness and zero excess kurtosis.

[9] See Alexander (2008*b*) page 98-99 where the leptokurtic character of the generalized Student-t distribution is showcased.

[10] A Kolmogorov-Smirnov (K-S) test, as described by Massey Jr (1951), is performed at each intersection. The null hypothesis that the underlying distribution of $\hat{\epsilon}_t$ stems from a generalized Student-t distribution cannot be confidently (p<0.05) rejected at any of the time steps. The K-S test compares the empirical cdf of $\hat{\epsilon}_t$ to the cdf of a particular Student-t distribution, whose parameters were estimated by solving the corresponding maximum likelihood problem using the *fitdistr* method from the *MASS* package in R. Time series plots of p-values for K-S tests and estimated mean, standard deviation and degrees of freedom is found in Appendix C.IV.

[11] Exotic discontinuous or fractal constructions may not adhere to this property. If certain constraints are applied to the distribution, Equation 5.5 and the consecutive reasoning may be proved rigorously, but we will rely on the reader's intuition.

and consequently, the limiting distribution of M_t as β_t approaches 1 is a solely a derivative of the error term ϵ_t distribution, as expected from Equation 5.3:

$$\lim_{\beta_t \to 1} f_M(m) = \lim_{\beta_t \to 1} \mathscr{F}^1\{\hat{f}_{\epsilon_t} \cdot \hat{g}\}(\log m - \alpha_t) = \mathscr{F}^1\{\hat{f}_{\epsilon_t}\}(\log m - \alpha_t) = f_{\epsilon_t}(\log m - \alpha) \tag{5.7}$$

For $\beta_t = 0.93$, it is thus natural to expect that the distribution of ϵ_t dominates and hence, that the practical consequences of assuming $\beta_t \approx 1$ will not lead to any large errors. This may also be established empirically by assuming a particular form for the $f_{\log \pi_t}$ and using numerical methods for the convolution.

A more radical interpretation of $\beta_t \neq 1$ is that $\frac{V_t}{\pi_t}$ is an inferior multiple because it fails to reflect ϵ_t solely and is dependent on π_t, the size of the enterprise profits. The implication of this argument is that the correct size-adjusted "β-multiple" is:

$$M_t = \frac{V_t}{\pi_t^{\beta_t}} \tag{5.8}$$

The multiple would be harder to interpret, but reasoning about the size of enterprises in discussions about the multiple would become superfluous because the effects already are taken into account. Risking to be accused of theoretical fanaticism, one could argue that 1/1 is as arbitrary as 1/0.93. Appendix E.I contains some discussions on the potential of a "size adjusted" β multiple. Nevertheless, the fact that β_t is close to 1 indicates that the apparently mundane M_t captures size independency effects well.[12]

As seen in Figure 3.8, the level and dispersion of the distribution vary significantly throughout time. It is notable that the Dot-com bubble burst (2000-2002) and the Global Financial Crisis (2007-2008) are different in their nature. The Dot-com bubble burst led to a massive shift in the upper part of the distribution, while the lower parts were less affected. In the Global Financial Crisis, the whole distribution shifted proportionally. Currently, the multiple values are somewhere between the height of the Dot-com bubble and just before the Global Financial Crisis. The 90% quantile is at an all-time high since Dot-com, while the median is still somewhat lower than in 2007.

5.1 Regression on Quantiles

A natural first step in approximating H, is to let the distribution be described by quantiles introduced in Section 3.2.3,[13] $y_{t,q}$ for $q \in \{1, ..., Q\}$[14], and to examine the dependency between the underlying $\boldsymbol{x}_t$ and these quantiles. The issues with stationarity for the quantiles resemble the issues for the median as they cannot be proven with 95% significance not to have unit roots. Nevertheless, an OLS regression is conducted on the basis that each of them is cointegrated with the underlying variables:[15]

$$y_{t,q} = \alpha_q + \boldsymbol{\beta_q} \cdot \boldsymbol{x_t} + \epsilon_{t,q} \tag{5.9}$$

This model capture variations in the whole distribution of multiples rather than just in the median or some other index. Assuming independence between $y_{t,q} \forall q$, the parameters can be estimated

[12] Cross-sectional regressions on the logarithm of EV/EBITDA versus EV and versus EBITDA have been performed at each point in time (see Appendix C.V), and the corresponding coefficients are insignificant or close to zero for the time period under consideration.

[13] Quantile data obtained from the categorical algorithm introduced in Chapter 3 are applicable in this analysis but are considered inappropriate due to incomparability to other models.

[14] A full description of the distribution can be obtained when $Q \to \infty$ i.e., $F(z) = \{q \mid \lim_{|Q| \to \infty} y_q = z\}$.

[15] The $H_0 : r \leq 2$ based on the Johansen test is rejected for all quantiles with $p < 0.01$. See Appendix C.VI.

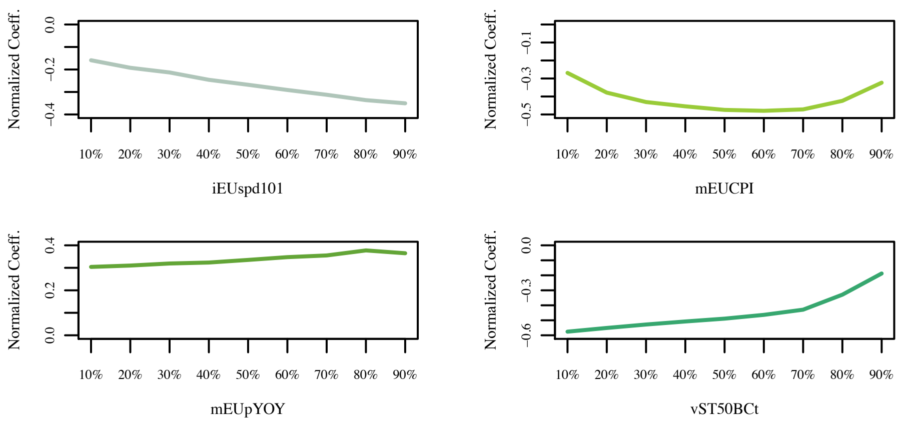


Figure 5.1: Normalized model coefficients for each quantile

by standard OLS methods. The α_q and $\boldsymbol{\beta}_q$ coefficients obtained by OLS estimation on normalized quantiles using underlying variables from *Ap* are displayed in Figure 5.1.[16] The regression coefficients may be used to construct discrete approximations to the distribution given underlying variables,[17] and their signs are consistent with the median multiple regression from Section 4.1.

The coefficients indicate that changes in the term structure spread affect the upper parts more severely than the lower parts,[18] inflation changes have more impact on the center of the distribution, and volatility primarily affects the lower and center parts.[19] The same overall trends are observed in the *Dx* and *Rx* models, although with some inconsistencies.[20] No a posteriori causality interpretation of these findings will be presented here, but we propose the application of System Dynamics to model causality as an area of further research in Appendix E.I. R^2 is lower for extreme quantiles, particularly the 90% quantile with $R^2 = 48\%$, possibly due to outlier sensitivity, as discussed in Section 3.2.1, but also potentially indicating that other variables may be more appropriate for describing the tails of the distribution.

The assumption that dependencies between quantiles are captured by the exogenous variables and their parameters in Equation 5.9 may not hold in order to adequately approximate H. Thus, a natural extension of the OLS model is to postulate explicit dependency relations between quantiles, $y_{t,q}$ and $y_{t,q'}$. A large fraction of variance among the quantiles is common, rendering it logical to

[16] Appendix C.VII contains tabulated coefficients and standard regression statistics for the normalized *Ap* regression (Table 9) along with regression results from corresponding *Rx* and *Dx* models on the quantiles (Table 11 to Table 15). In the non-normalized regressions, the magnitude of the higher quantile coefficients is greater, as the magnitude of their fluctuations is greater. The normalization was conducted to remove this effect, rendering the quantile coefficients comparable.

[17] A simple approximation is: $f_t(m) = 2\sum_{q=1}^{Q} \frac{\mathbb{1}_{\forall(q'\neq q)(|m-q|<|m-q'|)}}{\hat{y}_{t,q+1}-\hat{y}_{t,q-1}}$, where $\hat{y}_{t,q}$ are the estimated quantiles, $\hat{y}_{t,0} = -2\hat{y}_{t,1}$, $\hat{y}_{t,Q+1} = 2m - \hat{y}_{t,Q}$ and $\mathbb{1}$ is an indicator function. Natural extensions include continuous, smooth and second order smooth generalizations.

[18] Note the axis of Figure 5.1, displaying the actual normalized coefficient values, rather than the magnitude.

[19] The marginal change in the Box-Cox transformed volatility v_{BC} is proportional to change in the original volatility measure v: $\frac{\partial v_{bc}}{\partial t} = v^{-1.5}\frac{\partial v}{\partial t}$, and $v^{-1.5} > 0$.

[20] See Table 10 in Appendix C.VII for a complete overview of coefficients from all 6 regression models on normalized quantiles. Implied volatility in *Ad*, like *Ap* affects the center and lower parts, while volatility in the *Rx* and *Dx* models primarily affects the center. Inflation coefficients in *Rp* support mid-distribution trend in *Ap*.

model inter quantile distances, $\Delta_t^q = (y_{t,q+1} - y_{t,q})$, instead of quantiles themselves. We require Δ_t^q to be positive and define $\mathscr{D}_t^q = \log \Delta_t^q$. An ADF-test reject that $d\mathscr{D}_t^q = \mathscr{D}_t^q - \mathscr{D}_{t-1}^q$ has a unit root with $p < 0.01$, rendering $\mathscr{D}_t^q \sim I(1)$ a natural assumption, leading to the model:

$$d\mathscr{D}_t^q = \gamma_q d\mathscr{D}_{t-1}^q + \beta_q \boldsymbol{x}_t + \epsilon_t^q \tag{5.10}$$

The unit root hypothesis cannot be rejected for $\mathscr{D}_t^q$ with an ADF-test, but the $\mathscr{D}_t^q$s are cointegrated, and the following model will take the level of $\mathscr{D}_t^q$ into account directly:

$$\mathscr{D}_t^q = \psi \mathscr{D}_{t-1}^q + \beta_q \boldsymbol{x}_t + \epsilon_t^q \tag{5.11}$$

Now, a VAR(n) model with exogenous variables is an extension of Equation 5.9 that accounts for inter-quantile dependency, as well as being a generalization of the two previous equations:

$$\mathbf{D}_t = \boldsymbol{\alpha} + \Psi_1 \boldsymbol{D}_{t-1} + ... + \Psi_n \boldsymbol{D}_{t-n} + B\boldsymbol{x}_t + \boldsymbol{\epsilon}_t \tag{5.12}$$

where $\boldsymbol{D}_t = (\mathscr{D}_t^{1:Q})^T$, $\boldsymbol{\alpha}$, B, $\Psi_{1:n}$ are vectors and matrices of coefficients and $\boldsymbol{\epsilon}_t$ is a multivariate i.i.d. error process with covariance matrix Σ. The coefficients are estimated using OLS, and n is set to 1 based on information gain.[21] The model is obviously multicolinear, but we have not been able to identify autocorrelation for relevant lags.[22]

Table 5.1 displays the estimated coefficients for a VAR(1) implementation. ψ_{ii} diagonal values and neighboring ψ_{ij} coefficients are the most significant ones, but care must be taken in interpretation of these estimators, because the estimator variance is large due to multicollinearity. Inflation and volatility are the only exogenous variables with significant coefficients, but the cross-quantile estimator fluctuations indicate that also these are unstable. Nevertheless, the eigenvalues of $\hat{\Psi}$ are inside and close to the perimeter of the unit circle, and we can expect that although the process is near integrated, it is not explosive. The quantile VAR model is a-theoretical, but may approximate H adequately by producing quantiles $\boldsymbol{y}_t$ when given a starting set of quantiles $\boldsymbol{y}_{t-1}$ and a fixed point $y_{t,q}$, e.g., median estimate provided by Ap.[23]

The multicollinearity issues in the model from Equation 5.12 may be resolved by decomposing the variance of the quantiles using a principal component analysis as originally proposed by Pearson (1901). Letting K be the $Q \times T$ matrix of quantiles with the covariance matrix Σ_Q, the principal components are the columns of $P = KW$, where $W = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & ... & \boldsymbol{v}_Q \end{pmatrix}$ and $\boldsymbol{v}_i$ are the eigenvectors of Σ_Q.[24]

The eigenvalue problem for $Q = 99$ is solved numerically, noting the objections of Wilkinson (1965).[25] 99.2% of the variance is explained by the first three principal components with eigenvalue shares of 86.1%, 11.4% and 1.7%, respectively. Their historical development is plotted in Figure 5.2. The historical development of the first principal component (*pc1*) resembles the development common to all quantiles, and its correlation with the median multiple is $\rho = -0.99$. *pc2* has $\rho = 0.81$ correlation with the sample standard deviation[26] and can be expected to represent a dispersion effect.

[21] The # of coefficients becomes large for VAR(n) models. Implementation excluding exogenous variables with $n = 1$ gives AIC= −41.1, BIC= −40.1, HQ = −40.7 and $n = 2$ gives AIC= −41.0, BIC= −39.0, HQ = −40.2.

[22] The H_0: "ϵ_t is i.i.d", was tested with a multivariate Portmanteau test, unable to reject significant autocorrelations for $15 \geq \log n$ lags (Hosking 1980). See Table 16 in Appendix C.VII for test results.

[23] The approximation for H is given by $y_{t,q} = y_{t,q-1} + e^{\hat{\alpha}_q + \hat{\boldsymbol{\psi}}_1^q \cdot D_{t-1} + \hat{b}_q \cdot \boldsymbol{x}_t}$ for $q > 5$ and $y_{t,q} = y_{t,q+1} - e^{\hat{\alpha}_q + \hat{\boldsymbol{\psi}}_1^q \cdot D_{t-1} + \hat{b}_q \cdot \boldsymbol{x}_t}$ for $q < 5$, where $\hat{\alpha}$, $\hat{\Psi} = ((\hat{\boldsymbol{\psi}}_1^1)^T ... (\hat{\boldsymbol{\psi}}_1^n)^T)^T$ and $\hat{B} = (\hat{\boldsymbol{b}}_1^T ... \hat{\boldsymbol{b}}_n^T)^T$ are estimated coefficients.

[24] See Alexander (2008*b*) for elaborate description of principal component analysis, interpretation, and terminology.

[25] The eigenvalue problem was solved using the "eigen" function in the "base" package in R, a FORTRAN implementation of a DSQD-based algorithm (differential quotient difference with shifts).

[26] See Figure 24 in Appendix C.IX for time series plot of sample standard deviation, skewness and kurtosis.

Table 5.1: Results from VAR(n) model on quantiles

		Dependent variable:							
		log inter-quantile distances							
		$\mathscr{D}^t_{0.1}$	$\mathscr{D}^t_{0.2}$	$\mathscr{D}^t_{0.3}$	$\mathscr{D}^t_{0.4}$	$\mathscr{D}^t_{0.5}$	$\mathscr{D}^t_{0.6}$	$\mathscr{D}^t_{0.7}$	$\mathscr{D}^t_{0.8}$
	$\mathscr{D}^{t-1}_{0.1}$	0.647^{***}	0.142^{***}	0.031	-0.095^{*}	-0.111^{**}	0.087	0.007	−0.088
		(0.050)	(0.047)	(0.050)	(0.053)	(0.054)	(0.057)	(0.054)	(0.061)
	$\mathscr{D}^{t-1}_{0.2}$	0.199^{***}	0.522^{***}	0.164^{***}	0.104^{*}	−0.059	0.036	0.058	−0.075
		(0.056)	(0.053)	(0.056)	(0.059)	(0.061)	(0.064)	(0.060)	(0.068)
	$\mathscr{D}^{t-1}_{0.3}$	−0.046	0.298^{***}	0.489^{***}	0.113^{**}	0.094	−0.041	−0.011	−0.015
		(0.053)	(0.050)	(0.053)	(0.056)	(0.058)	(0.061)	(0.057)	(0.065)
	$\mathscr{D}^{t-1}_{0.4}$	0.011	−0.059	0.200^{***}	0.429^{***}	0.080	0.004	−0.036	-0.133^{**}
		(0.054)	(0.051)	(0.055)	(0.057)	(0.059)	(0.062)	(0.059)	(0.066)
Ψ_1	$\mathscr{D}^{t-1}_{0.5}$	0.102^{*}	−0.001	0.062	0.291^{***}	0.435^{***}	0.280^{***}	0.074	0.125^{*}
		(0.057)	(0.053)	(0.057)	(0.060)	(0.061)	(0.065)	(0.061)	(0.069)
	$\mathscr{D}^{t-1}_{0.6}$	−0.043	0.020	0.047	−0.018	0.231^{***}	0.427^{***}	0.060	0.062
		(0.057)	(0.053)	(0.057)	(0.060)	(0.061)	(0.065)	(0.061)	(0.069)
	$\mathscr{D}^{t-1}_{0.7}$	0.004	−0.020	0.011	0.063	0.070	0.151^{***}	0.668^{***}	0.213^{***}
		(0.047)	(0.043)	(0.047)	(0.049)	(0.051)	(0.053)	(0.050)	(0.057)
	$\mathscr{D}^{t-1}_{0.8}$	−0.020	−0.0002	−0.020	−0.051	0.047	0.038	0.202^{***}	0.734^{***}
		(0.041)	(0.038)	(0.041)	(0.043)	(0.045)	(0.047)	(0.044)	(0.050)
α^T	Intercept	0.502^{***}	0.270^{**}	0.450^{***}	0.507^{***}	0.264^{*}	0.470^{***}	0.267^{*}	0.863^{***}
		(0.134)	(0.125)	(0.135)	(0.141)	(0.146)	(0.154)	(0.144)	(0.163)
	iEUspd101	−0.008	0.008	−0.010	−0.006	0.003	0.003	0.007	−0.010
		(0.009)	(0.009)	(0.009)	(0.010)	(0.010)	(0.010)	(0.010)	(0.011)
	mEUCPI	-0.022^{**}	-0.015^{*}	−0.006	-0.024^{***}	-0.026^{***}	−0.005	−0.006	-0.029^{***}
		(0.008)	(0.008)	(0.008)	(0.009)	(0.009)	(0.010)	(0.009)	(0.010)
B	mEUpYOY	−0.001	0.001	−0.001	0.001	−0.0003	−0.001	−0.001	−0.001
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
	vST50BCt	-0.235^{***}	-0.185^{**}	-0.285^{***}	-0.291^{***}	-0.228^{**}	-0.313^{***}	-0.219^{**}	-0.400^{***}
		(0.087)	(0.081)	(0.087)	(0.091)	(0.094)	(0.099)	(0.093)	(0.105)
	Observations	222	222	222	222	222	222	222	222
	R^2	0.845	0.873	0.882	0.851	0.853	0.875	0.933	0.922
	Adjusted R^2	0.836	0.865	0.875	0.843	0.845	0.868	0.929	0.918
	Res. Std. Error	0.073	0.068	0.073	0.077	0.079	0.084	0.079	0.089
	F Statistic	95.025^{***}	119.452^{***}	130.134^{***}	99.782^{***}	101.363^{***}	122.188^{***}	240.982^{***}	206.077^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

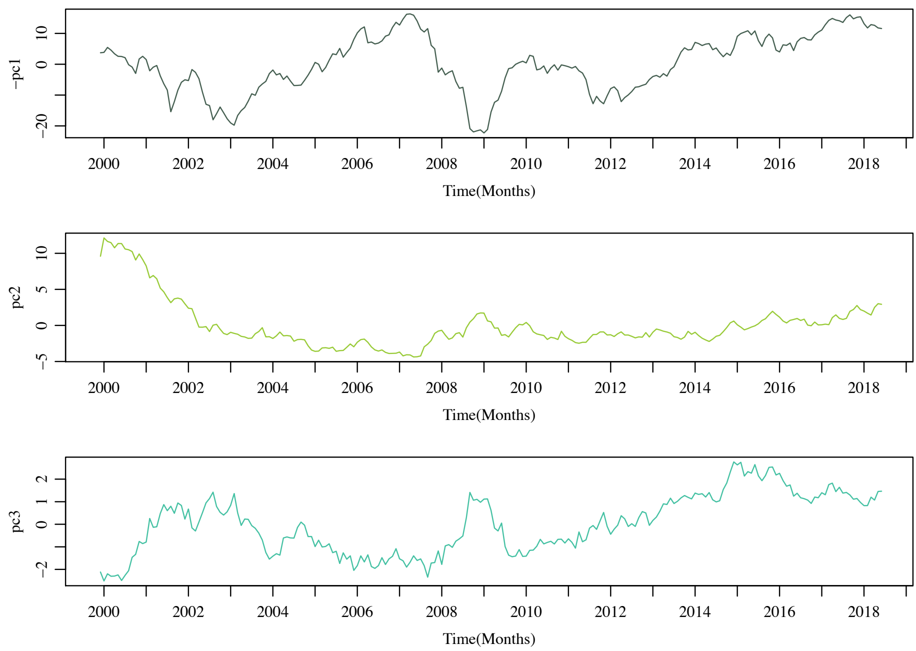


Figure 5.2: Historical development of *-pc1*, *pc2* and *pc3*

Figure 5.3 depicts the effect coefficients of *pc1*, *pc2* and *pc3* on each of the quantiles. It is evident that the (negative) first component is a trend component capturing proportional shifts in the complete distribution.[27] The second component may be interpreted as a contraction (tilt[24]) component as each of the distributional tails are affected oppositely. An increase in this component leads to an expansion of the distribution. *pc3* represents more complex skew-dynamics (convexity[24]), and an increase entails a negative drift of both tails and a positive drift of the center - skewing the distribution negatively. *pc2* has one fixed point[28] in the center of the distribution and *pc3* has two fixed points.

Because the *pc1*, *pc2* and *pc3* are linear combinations of quantiles that are assumed to be cointegrated with the underlying variables, a regular cointegrated factor analysis is conducted using independent variables from *Ap*. The resulting coefficients and R^2 values are displayed in Table 5.2 and resemble *Ap* coefficients for *pc1*, likely to capture the overall trend. The other regressions are less significant, have far weaker R^2 values and the interpretation is not straightforward, because the cointegration may be weak and nonlinear cross effects between the different principal coefficients may affect causality. Other sets of exogenous variables may describe the principal components more precisely.[29] Nonetheless, we note that *pc2*, representing dispersion, increases with volatility and

[27]Note that a dispersion proportional to the standard deviation of the quantiles is encompassed in this trend component because the quantiles are normalized in advance of the decomposition.

[28]By a fixed point, we mean a point in the distribution that is inert to movements of a certain factor.

[29]New subset regressions were conducted on each of the principal components. See Table 18 in Appendix C.X for favored coefficients. The USD/EUR exchange rate, 10-year interest rate and inflation are identified as key variables for *pc2*, *pc3* and *pc4*, but we will leave this subject as an area for further research. A preliminary hypothesis explaining the USD/EUR-rate effect is: Importing firms are valued at higher multiples than exporting firms. When EUR is devalued, importing corporation's cash flows diminish, leading to multiple distribution

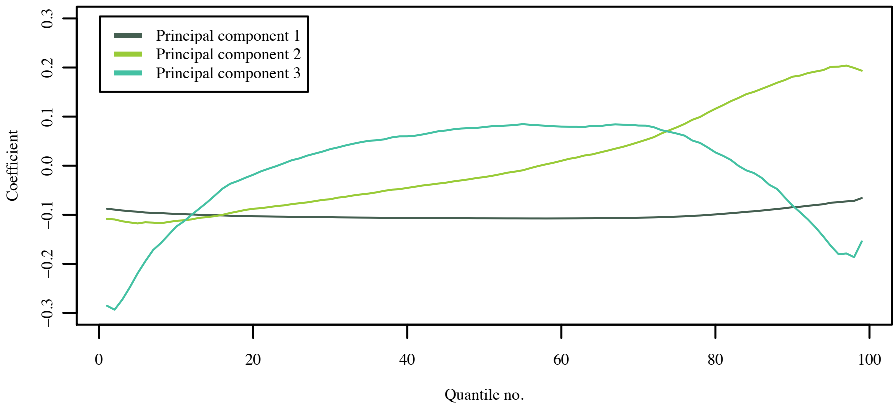


Figure 5.3: Principal component coefficients for 99 quantiles

decreases with the spread, while *pc3* is affected primarily by inflation. These results are strikingly consistent with the coefficients in Figure 5.1.

In summary, three quantile estimation models have been created: *qOLS* - OLS directly on quantiles (Equation 5.9), *qVAR* - VAR on quantile differences (Equation 5.12) and *qPCA* - PCA-based OLS on quantiles. *qOLS* and *qVAR* are a priori consistent with *Ap*, while *qPCA* has a negligible deviation, rendering all models operative as extensions of *Ap*. Figure 5.4 displays plots of each of the in-sample predictions against actual quantiles.[30] It is evident from the *qOLS* versus *qPCA* plot that *pc1*, *pc2* and *pc3* capture almost all variance in the distribution as the in-sample predictions are inseparable. In Chapter 4, *Rp* had a similar accuracy to *Ap* in in-sample modeling of the median, but *qVAR* is stronger than *qOLS* and *qPCA* at higher quantile predictions (see Table 5.3). This indicates that the propagation of distributional dynamics by building quantiles from differences implemented in *qVAR* captures some profound aspects that *qOLS* and *qPCA* fail to incorporate. Even though interdependencies between the quantiles in *qVAR* have a stabilizing effect on the quantiles close to the median, the massive parametric flexibility may cause overfitting issues. *qOLS* and *qPCA* are less prone to overfitting, because individual quantiles only depend on linear relationships, but underfitting is significant, as seen for the 90% quantile in-sample. *qPCA* represents a transparent model for the drivers of the distribution moments with far fewer degrees of freedom than *qOLS*, and infeasibility issues in predictions can be monitored by inspection of principal component coefficients, possibly involving smoothing operations.

5.2 Parametric Distribution Regression

The quantile perspective investigated in the preceding chapter quickly becomes complicated when accounting for various assumptions and constraints related to the fundamental properties of distributions. Leveraging on the discussion from this chapter's introduction, we have developed a method of regressing a dynamic parametric distribution on a set of underlying variables. Assume that the observed multiples $\omega \in \Omega_t$ are instances of the random variable M_t. The pdf of M_t, $f_{\boldsymbol{\rho}_t}$, is assumed

contraction.

[30]Note that the *qVAR* distribution was built by fixing initial quantiles $y_{t=1,q}$ and building each cross sectional distribution from the *Ap* median fitted values, according to the method explained in footnote 23.

Table 5.2: OLS model coefficient on *pc1*, *pc2* and *pc3* using *Ap* underlying variables

	Dependent variable:		
	Principal Component		
	pc1	*pc2*	*pc3*
iEUspd101	3.640***	−1.030***	−0.151
	(0.438)	(0.336)	(0.116)
mEUCPI	4.097***	−0.299	−0.688***
	(0.302)	(0.232)	(0.080)
mEUpYOY	−0.674***	0.083*	−0.028*
	(0.064)	(0.049)	(0.017)
vST50BCt	64.943***	15.886***	2.026*
	(4.421)	(3.391)	(1.166)
Constant	−113.159***	−23.284***	−1.773
	(6.840)	(5.246)	(1.804)
R^2	0.800	0.114	0.294
Adjusted R^2	0.797	0.098	0.281
Residual Std. Error	4.163	3.193	1.098
F Statistic	218.402***	7.019***	22.723***
Note:			*p<0.1; **p<0.05; ***p<0.01

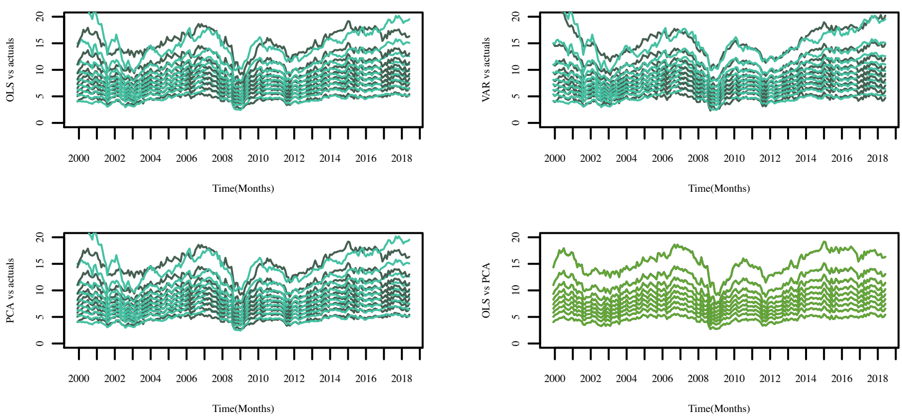


Figure 5.4: In-sample fitted quantiles vs. actual quantiles for *qOLS*, *qVAR* and *qPCA* (actuals in turquoise). Bottom right: in-sample fitted values from *qOLS* vs *qPCA*.

Table 5.3: Root Mean Square Deviations measuring complete in-sample accuracy between given models for each quantile

	10%	20%	30%	40%	50%	60%	70%	80%	90%
qOLS vs actuals	0.199	0.218	0.246	0.281	0.305	0.356	0.519	1.249	4.561
qVAR vs actuals	0.287	0.294	0.300	0.304	0.302	0.306	0.348	0.506	1.365
qPCA vs actuals	0.199	0.218	0.246	0.281	0.305	0.356	0.520	1.252	4.563
qOLS vs *qPCA*	0.000	0.000	0.001	0.000	0.000	0.000	0.001	0.002	0.002

to be a known distribution parameterized by the vector $\boldsymbol{\rho}_t$ which is dependent on the underlying macroeconomic variables through a linear relationship:[31]

$$\boldsymbol{\rho}_t = R\boldsymbol{x_t} \tag{5.13}$$

where R is a matrix of coefficients, each row mapping $\boldsymbol{x}_t$ to ρ_i. The parameters of R will be estimated to maximize resemblance between $f_{\boldsymbol{\rho}_t}$ and development of the empirical distribution of multiples Ω_t.

Resemblance may be measured by constructing a loss function $\Lambda(R) = A(D_{1:T})$, where $D_t = D(f_t, \Omega_t)$ represents a distance metric measuring the distance between individual distributions and $A(D_{1:T})$ is an aggregation function computing the total distance between two series of distributions based on the individual distances. From this, $\hat{R} = \text{argmin}_R \Lambda$ gives the parameters resulting in the best fit according to the loss function. The adequate estimation technique depends on the choice of A and D.[32] We propose a likelihood formulation of the optimization problem:[33]

$$L(R;\boldsymbol{x}_{1:T}) = \prod_{t=1}^{T_{max}} P(\Omega_t|f_{\boldsymbol{\rho}_t}) = \prod_{t=1}^{T_{max}} \prod_{\omega\in\Omega_t} P(\omega|f_{\boldsymbol{\rho}_t}) \tag{5.14}$$

$$\mathscr{L}(R;\boldsymbol{x}_{1:T}) = \sum_{t=1}^{T_{max}} \sum_{\omega\in\Omega_t} \log P(\omega|f_{\boldsymbol{\rho}_t}) \tag{5.15}$$

From visual inspection of historical data we observe that some fundamental characteristics are identified with a lognormal distribution: bell-shape, positive skewness and positive domain. Because the lognormal distribution is conventional and is associated with many naturally occurring sampling processes in business (Gibrat 1931), it is instructive to assume that M_t is lognormally distributed and parameterized in the following way: $\log M_t \sim \mathcal{N}(\boldsymbol{r}_\mu^T \boldsymbol{x}_t, e^{\boldsymbol{r}_\sigma^T \boldsymbol{x}_t})$,[34] where $R = (\boldsymbol{r}_\mu, \boldsymbol{r}_\sigma)^T$. The Likelihood function becomes:

$$\mathscr{L}(R;\boldsymbol{x}_{1:T}) = \sum_{t=1}^{T_{max}} \sum_{\omega\in\Omega_t} \log\left[\frac{1}{\omega e^{\boldsymbol{r}_\sigma^T \boldsymbol{x}_t}\sqrt{2\pi}} \exp\left\{-\frac{(\log\omega - \boldsymbol{r}_\mu^T \boldsymbol{x}_t)^2}{2e^{2\boldsymbol{r}_\sigma^T \boldsymbol{x}_t}}\right\}\right] \tag{5.16}$$

[31] The linearity constraint reduces the model flexibility, but is assumed to capture most of the variance and reduces risk of overfitting. Some distribution parameters have domain constraints, and in order to preserve the linearity condition, we will assume a logarithmic transformation. E.g., see definitions of σ for the lognormal distribution proposed later in the chapter.

[32] See Appendix C.XI for a proposal of estimation for squared distance D and logarithmic sum A.

[33] Assuming that the priors $P(R)$ are equal, and hence: $P(R|x) = \alpha P(x|R)P(R) \propto P(x|R)$, and that each observation $\omega \in \Omega$ is a random and independent draw from the underlying distribution $f_{\boldsymbol{\rho}_t}$.

[34] Here, $\boldsymbol{\rho}_t$ is defined by $\boldsymbol{\rho}_t = (\rho_\mu, \rho_\sigma)^T = R\boldsymbol{x_t} = (\boldsymbol{r}_\mu^T \boldsymbol{x}_t, \boldsymbol{r}_\sigma^T \boldsymbol{x}_t)^T$ and the lognormal distribution is defined by $\log Y_t \sim N(\rho_\mu, e^{\rho_\sigma})$. This ensures that σ is positive in the normal distribution: $\sigma = e^{\boldsymbol{r}_\sigma^T \boldsymbol{x}_t}$.

Table 5.4: Lognormal distribution parameters ($\hat{R}$) obtained by maximization of likelihood function in Equation 5.16 and descriptive statistics from resulting parameters

			$\hat{R}$				In sample statistics		
$\boldsymbol{x}_t$	Constant	Term Spread	Inflation	Growth	Volatility		Median	Max	Min
$\hat{\boldsymbol{r}}_\mu$	4.078	-0.054	-0.065	0.013	-1.154	$\hat{\mu}_{1:T}$	2.072	2.386	1.654
$\hat{\boldsymbol{r}}_\sigma$	-1.528	-0.017	-0.010	-0.002	0.580	$\hat{\sigma}_{1:T}$	0.584	0.734	0.457

This function is a multivariate, nonlinear, non-concave[35] function with possibilities of local optima. In order to find $\hat{R} = \text{argmax}_R \mathscr{L}(R;x)$, the gradient was computed:

$$\frac{\partial \mathscr{L}}{\partial \boldsymbol{r}_\sigma} = \sum_{t=1}^{T} \boldsymbol{x}_t \sum_{\omega \in \Omega_t} ((\log\omega - \boldsymbol{r}_\mu^T \boldsymbol{x}_t)^2 e^{-2\boldsymbol{r}_\sigma^T \boldsymbol{x}_t} - 1)$$

$$\frac{\partial \mathscr{L}}{\partial \boldsymbol{r}_\mu} = \sum_{t=1}^{T} \boldsymbol{x}_t \sum_{\omega \in \Omega_t} (\log\omega - \boldsymbol{r}_\mu^T \boldsymbol{x}_t) e^{-2\boldsymbol{r}_\sigma^T \boldsymbol{x}_t}$$

Several gradient descent and Newton's method based algorithms were explored, along with non-derivative search algorithms, as proposed by Nelder & Mead (1965), Corana et al. (1987) and Hooke & Jeeves (1961). Despite some issues with time complexity, local minimas and convergence, we achieved satisfying results using the Newton-based method and Hooke & Jeeves.[36] The value of $\hat{\mathscr{L}}$ does not provide much intuition, but $\hat{R}$ is presented in Table 5.4. The signs of the $\hat{\boldsymbol{r}}_\mu$ parameters are consistent with the impact in Ap, and the median $\hat{\mu}$ value is $2.1 = \log 7.9$, similar to the historical median multiple. Volatility has a relatively weak contraction effect compared to other variables, consistent with the findings in the preceding section. Using $\hat{R}$, the median $\hat{\mu}_t$ and standard deviation $\hat{\sigma}_t^2$ are calculated for each in-sample time step t in the following manner:

$$\hat{\mu}_t = e^{\boldsymbol{r}_\mu^T \boldsymbol{x}_t} \tag{5.17}$$

$$\hat{\sigma}_t^2 = (\exp(e^{2\hat{\boldsymbol{r}}_\sigma^T \boldsymbol{x}_t}) - 1)\exp(2\hat{\mu}_t + e^{2\hat{\boldsymbol{r}}_\sigma^T \boldsymbol{x}_t}) \tag{5.18}$$

Figure 5.5 shows the in-sample development of $\hat{\sigma}_t$ and $\hat{\mu}_t$, both fairly consistent with the empirical in-sample standard deviation and median. Knowing that the empirical in-sample statistics are exogenous to $\hat{R}$, their fit gives confidence in descriptive value of the model. It is clear, however, that there is some misspecification, as the standard deviation parameter deviate substantially from the sample standard deviation on several occasions. It is natural to hypothesize that the lognormal distribution is too restrictive and unable to account for the significant leptokurticity in the sample. In the introduction of this chapter, M_t was proposed to be log Student-t distributed, as we were unsuccessful in rejecting the hypothesis that the error terms, ϵ_t from the cross-sectional Equation 5.2, are Student-t distributed. In fact, lognormality on each cross-section Ω_t is confidently rejected with a JB-test, while we are unable to reject log Student-t distribution using a Kolmogorov-Smirnov

[35] Set $\boldsymbol{r}_\sigma = (r_\sigma, \boldsymbol{0})$ and $\boldsymbol{r}_\mu = (r_\mu, \boldsymbol{0})$, and compute $\frac{\partial^2 \mathscr{L}}{\partial r_\sigma r_\mu} = -\sum_{t=1}^{T} \sum_{\omega \in \Omega_t} 2x^2(\log\omega - r_\mu x)e^{-2r_\sigma x}$ corresponding to the off-diagonal element in the Hessian of the bivariate $\mathscr{L}_b$. An a,c,c,b-matrix is negative semi-definite if $\alpha^2 a + 2\alpha\beta c + \beta^2 b \leq 0$ for some vector $(\alpha\beta)$. r_μ may be set such that $\alpha\beta c$ is arbitrarily big, making the Hessian non-negative semi-definite. A non-negative semi-definite Hessian implies a non-concave function (Kreyszig 2010).

[36] The Newton-based method used is the R *nlm* function implementation of Dennis (1983) in the *stats* package.

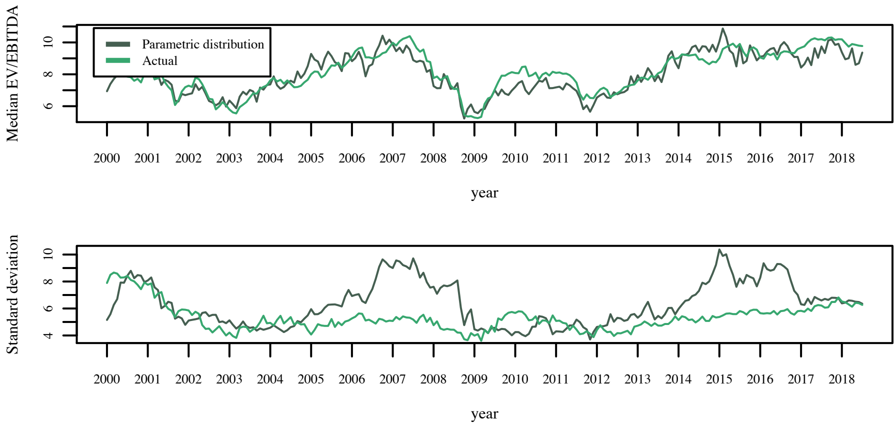


Figure 5.5: In-sample estimated standard deviation and median from lognormal distribution model against empirical sample median and sample standard deviation

test for the same samples. Consequently, we assume that $\log M_t \sim \mathcal{T}(\boldsymbol{r}_\mu \boldsymbol{x}_t, e^{\boldsymbol{r}_\sigma^T \boldsymbol{x}_t}, e^{\boldsymbol{r}_\nu^T \boldsymbol{x}_t})$, where $R = (\boldsymbol{r}_\mu, \boldsymbol{r}_\sigma, \boldsymbol{r}_\nu)^T$. We enforce non-negativity for the dispersion coefficient $e^{\boldsymbol{r}_\sigma^T \boldsymbol{x}_t}$ and degrees of freedom $e^{\boldsymbol{r}_\nu^T \boldsymbol{x}_t}$ in the familiar way. The new likelihood function becomes:

$$\mathcal{L}(R;\boldsymbol{x}_{1:T}) = \sum_{t=1}^{T_{max}} \sum_{\omega\in\Omega_t} \log\left[\frac{\Gamma(\frac{e^{\boldsymbol{r}_\nu^T\boldsymbol{x}_t}+1}{2})}{\Gamma(\frac{e^{\boldsymbol{r}_\nu^T\boldsymbol{x}_t}}{2})\omega e^{\boldsymbol{r}_\sigma^T\boldsymbol{x}_t}\sqrt{e^{\boldsymbol{r}_\nu^T\boldsymbol{x}_t}\pi}} \left(\frac{(\log\omega - \boldsymbol{r}_\mu^T\boldsymbol{x}_t)^2}{e^{2\boldsymbol{r}_\sigma^T\boldsymbol{x}_t}e^{\boldsymbol{r}_\nu^T\boldsymbol{x}_t}}\right)^{-\frac{e^{\boldsymbol{r}_\nu^T\boldsymbol{x}_t}+1}{2}}\right] \tag{5.19}$$

The maximization problem for this likelihood function is intimidating but has many of the same characteristics as the likelihood maximization for the lognormal distribution. We use the same Newton-based algorithm and find the optimal parameters $\hat{R}$ seen in Table 5.5. Profound interpretation of $\hat{R}$ has to be done with caution because the actual distributional moments relate to the parameters in nonlinear ways, but we note that the $\hat{\boldsymbol{r}_\mu}$ have promising signs and magnitudes and that $\hat{\nu}_{1:T} << \infty$ indicating excess kurtosis. The moments of the fitted log Student-t distributions were computed numerically by using conventional sample moment estimators on 10,000 samples. Figure 5.6 present the historical in-sample moments against the fitted log Student-t median, standard deviation, skewness and kurtosis.[37] The log Student-t distribution performs significantly better than the lognormal distribution in capturing the median and standard deviation and is able to capture the general movements in skew and kurtosis, although with some periodical deviations, particularly during the dot-com bubble burst for standard deviation.

Figure 5.7 displays in-sample fits of the two methods for four selected cross-sections. Both distributions, based solely on one input vector $\boldsymbol{x}_t$, fit well with the observed values, but the parametric flexibility of the log Student-t distribution gives it more accuracy, particularly in the tails. At the $t = 100$ intersection, the restrictiveness of the lognormal distribution is evident as it overestimates the mid region from ~ 12x to ~ 25x, but underestimates the mode and the upper tail. The responsiveness of log Student-t distribution is illustrated in the difference between the leptokurtic distribution at $t = 1$ and the more lognormal distribution at $t = 200$. Figure 5.8 shows in-sample estimated log

[37]For consistency sampled multiples above 50x were excluded.

Table 5.5: Log Student-t distribution parameters ($\hat{R}$) obtained by maximization of likelihood function in Equation 5.19 and descriptive statistics from resulting parameters

			$\hat{R}$				In sample statistics		
$\boldsymbol{x}_t$	Constant	Term Spread	Inflation	Growth	Volatility		Median	Max	Min
$\hat{\boldsymbol{r}}_\mu$	4.086	-0.056	-0.071	0.012	-1.151	$\hat{\mu}_{1:T}$	2.106	2.354	1.699
$\hat{\boldsymbol{r}}_\sigma$	1.056	-0.034	-0.018	0.000	0.150	$\hat{\sigma}_{1:T}$	0.408	0.440	0.390
$\hat{\boldsymbol{r}}_\nu$	4.582	-0.047	-0.029	0.005	-1.816	$\hat{\nu}_{1:T}$	5.128	6.806	3.410

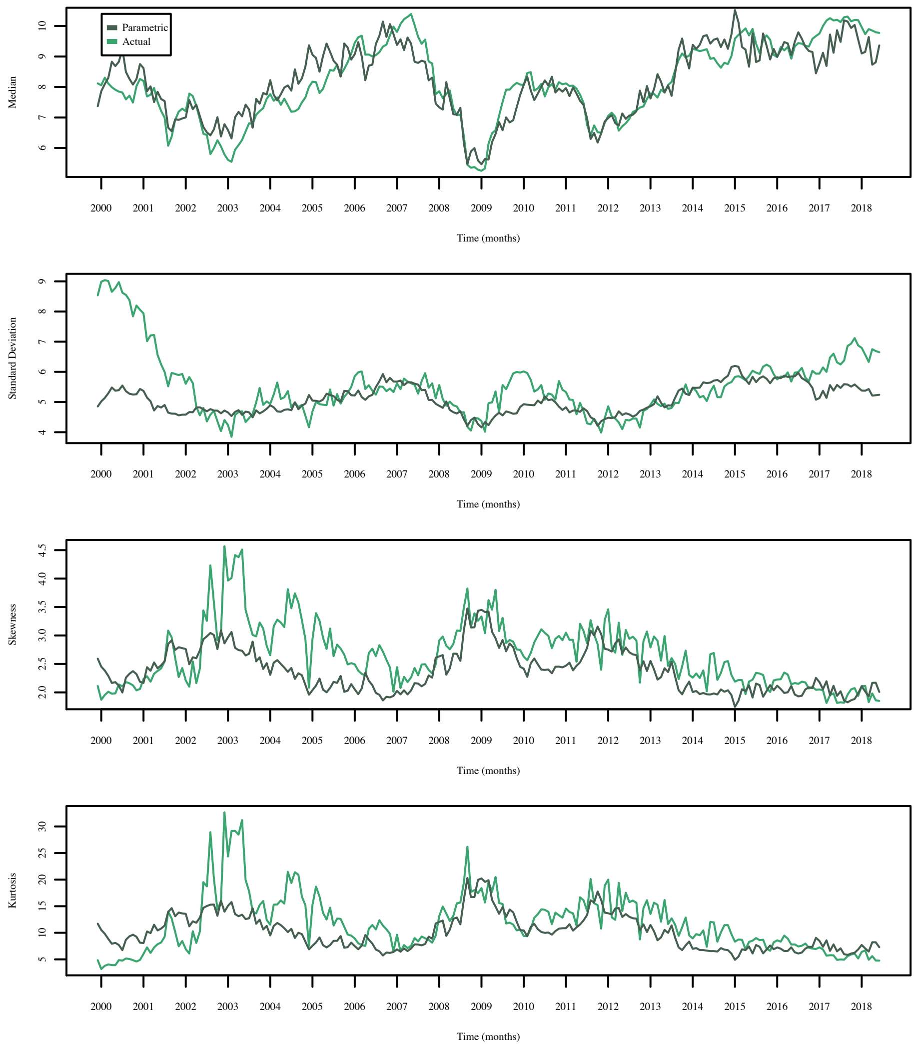


Figure 5.6: In-sample estimated median, standard deviation, skewness and kurtosis from log Student-t distribution model against empirical sample values

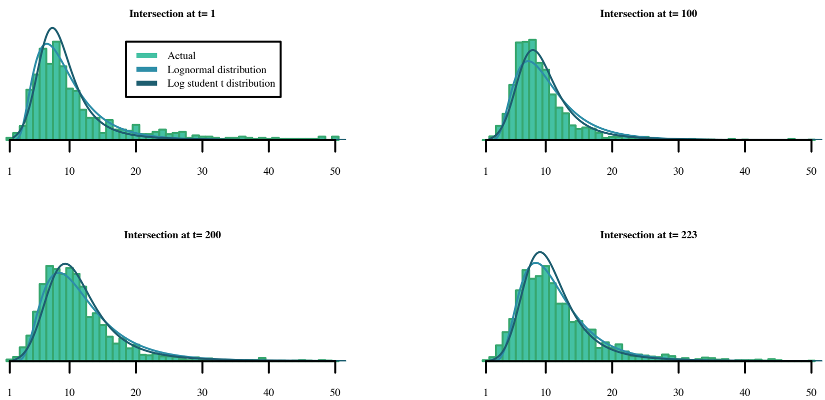


Figure 5.7: In-sample estimated parametric distributions and empirical distribution

Student-t distribution over time, illustrating the adaptiveness of the method. A distribution with additional parameters, such as the skewed generalized t distribution (Theodossiou 1998) or generalized hyperbolic distribution (introduced by Barndorff-Nielsen (1977)), could potentially yield even better fits, but the general optimization problem is $O(e^n)$ in number of parameters, and overfitting could become a significant issue. Accordingly, we consider the generalized log-t distribution to be satisfactory for our purpose, and it has a profound foundation in the quotient distribution of EV and EBITDA.

5.3 Frequency Regression

The parametric distribution regression is an elegant approach to estimation of H, but it is inherently restricted by the number of parameters and is dependent on computationally demanding optimization problems. Hence, we have developed a flexible frequency-based method for approximating H.

In this model, the pricing data range is mapped into B non-overlapping and exhaustive intervals (buckets),[38] and the pdf of M_t is assumed to be described by frequencies $\nu_{bt} \in (0, 1)$ for the buckets b at time t, such that the probability density function of the multiple ω is:

$$P(M_t = m) = \sum_{b=1}^{B} \nu_{bt} \mathbb{1}_{l_b \leq m < u_b} \tag{5.20}$$

where l_b and u_b denote the lower and upper limits of the bucket. ν_{bt} is assumed to be dependent on the underlying variables $\boldsymbol{x}_t$ through a logistic function and an i.i.d. residual process with diagonal covariance matrix ϵ_{bt}. γ_t is a normalization constant:

$$\nu_{bt} = \frac{\gamma_t}{1 + e^{\alpha_b + \boldsymbol{\beta}_b \cdot \boldsymbol{x}_t + \epsilon_{bt}}}, \sum_{b=1}^{B} \nu_{bt} = 1 \tag{5.21}$$

[38] In the implementation below, the intervals are equal and trailing, but there are many potential approaches bucket allocation, e.g., buckets based on a fixed number of observations or an OLS F-test requirement. See Appendix E.I for discussion.

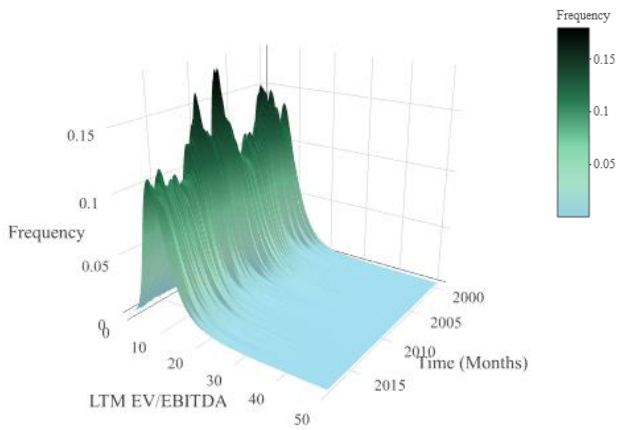


Figure 5.8: In-sample estimated log Student-t distributions over time

Now, for $B(|\boldsymbol{\beta}_t|+1)$ parameters a likelihood function is defined based on the complete training data set:

$$L(\boldsymbol{\beta}_{1:B}, \alpha_{1:B}; \boldsymbol{x}_{1:T}) = \prod_{t=1}^{T} \prod_{\omega \in \Omega_t} P(M_t = \omega) \tag{5.22}$$

Letting Ω_{bt} be the set of observations in bucket b at time t, the log likelihood function becomes:[39]

$$\mathscr{L}(\boldsymbol{\beta}_{1:B}, \alpha_{1:B}; \boldsymbol{x}_{1:T}) = \sum_{t=1}^{T} |\Omega_t| \log \gamma_t - \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log(1 + e^{\alpha + \boldsymbol{\beta}_b \cdot \boldsymbol{x}_t}) \tag{5.23}$$

$(\hat{\boldsymbol{\beta}}_{1:B}, \hat{\alpha}_{1:B}) = \text{argmax}_{(\boldsymbol{\beta}_{1:B}, \alpha_{1:B})} \mathscr{L}$ are the optimal coefficients for this model given the sample $\Omega_{1:T}$. The optimization problem has 250 parameters for Ap, and we have not been successful in developing an algorithm to compute it. The core of this part of the modeling is the heuristic approaches, which has proven very successful in giving practically viable distribution estimates. A natural approximation of v_{bt} is by the empirical frequencies $\bar{v}_{bt} = \frac{|\Omega_{tb}|}{|\Omega_t|}$, and if ϵ_{bt} is assumed to be i.i.d. with diagonal covariance matrix, the following logistic regression model may be obtained:[40]

$$\bar{v}_{bt} = \frac{\gamma_t}{1 + e^{\boldsymbol{\alpha_b} + \boldsymbol{\beta_b} \cdot \boldsymbol{x_t} + \epsilon_{bt}}} \tag{5.24}$$

For estimation purposes, a perturbation, $\delta_b = \frac{1}{2} \min_t \{\hat{v}_{bt}\}$, is added to each empirical frequency because the finite-sample bias renders some frequencies zero, giving infeasible results. The value of δ_b is set large enough not to distort the regression significantly, and small enough to represent a low-value instance. 19.5% of the frequencies observed are empty, mostly $b > 30$-buckets.[41] Equation 5.24 is reformulated and estimated using OLS:

[39] See Appendix C.XII for derivation of Equation 5.23 using Equation 5.20, 5.21 and 5.22.

[40] The assumption that ϵ_{bt} is i.i.d. and has diagonal covariance matrix is almost certain not to hold, but we are seeking practical heuristics, and will ignore the issue in parts of the discussion.

[41] See overview of magnitude of problem in Appendix C.XIII.

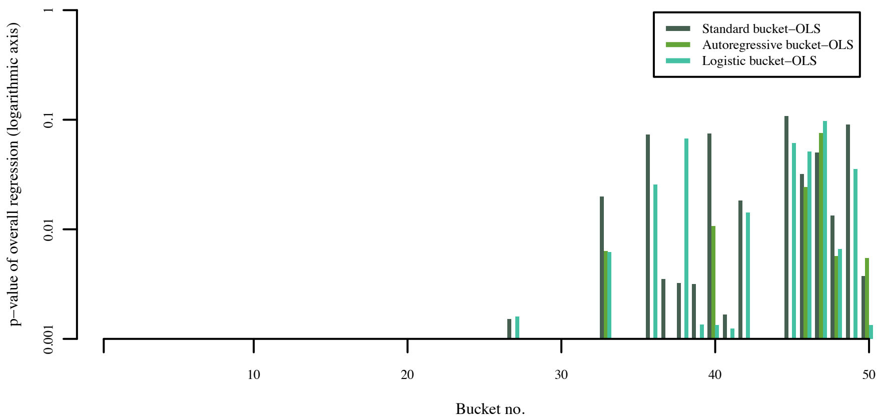


Figure 5.9: F-statistic p-value from the frequency approach regressions for each bucket

$$\log(\frac{1}{\bar{v}_{bt} + \delta_b} - 1) = \alpha_b + \boldsymbol{\beta_b} \cdot \boldsymbol{x_t} + \epsilon_{bt} \tag{5.25}$$

Because the perturbation makes the logistic regression slightly biased, we have developed a linear model, *liFreq* (Equation 5.26), that tolerates 0 frequencies, but has an invalid domain ($\notin [0, 1]$). The domain constraint turns out to be violated only for a diminishing fraction of the frequencies, mainly high-multiple buckets with few observations.

$$\bar{v}_{bt} = \alpha_b + \boldsymbol{\xi_b} \cdot \boldsymbol{x_t} + \epsilon_{bt} \tag{5.26}$$

Because the frequencies are likely to be autocorrelated, we propose an autoregressive extension of the preceding model (*auFreq*):

$$\bar{v}_{bt} = \alpha_i + \psi v_{b,t-1} + \boldsymbol{\xi_b} \cdot \boldsymbol{x_t} + \epsilon_{bt} \tag{5.27}$$

Using the fitted values from the model $\hat{v}_{bt}$, a complete estimation of the distribution is obtained through normalization $\tilde{v}_{bt} = \hat{v}_{bt}(\sum_{b=1}^{B} \hat{v}_{bt})^{-1}$, letting:

$$\hat{H}_{\boldsymbol{x}_t}(m) = \sum_{b=1}^{B} \tilde{v}_{bt} \mathbb{1}_{l_b \leq m < u_b} \tag{5.28}$$

The three approximations were estimated using OLS, although with significant issues in compliance: discreteness, censoring, autocorrelation, and others, so the output must be interpreted with care. Figure 5.9 displays reported F-test significance for the different buckets for the three models described, and unsurprisingly, the higher multiple buckets are the ones most pressing issues, as they have smaller sample sizes and are affected by the discrete nature of the sampling and the non-linearity issues. 1.0% of in-sample predictions in *liFreq* are invalid due to domain issues, and 1.5% of the *auFreq*-values, which gives confidence in the assumption that the frequencies are sufficiently linear within its domain to be estimated adequately using a linear model. Henceforth, *liFreq* will be the primary frequency based approach, based on its simplicity, low frequency of domain issues, ability to provide distributional estimates based solely on $\boldsymbol{x}_t$, and coefficient stability, which will be elaborated shortly.

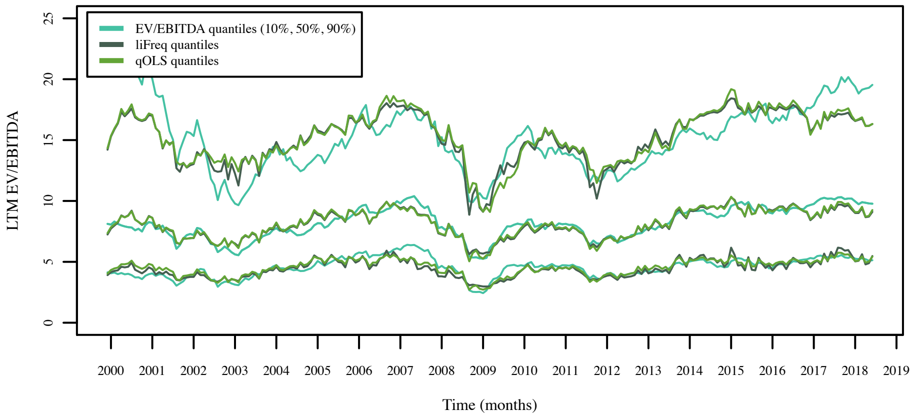


Figure 5.10: 10%, 50% and 90% quantiles against in-sample nonparametric and in-sample OLS

Figure 5.10 displays quantiles computed from the in-sample distribution predictions by *liFreq*, against actual quantiles and *qOLS* quantiles.[42] It is clear that the quantiles obtained from the frequency model have close resemblance to the *qOLS* quantiles, although the quantiles are unseen to *liFreq*. This indicates that some underlying relationship is captured in both models and gives confidence in the overall estimate of the distribution from the frequency model. Figure 5.11 depicts selected in-sample distributions of *liFreq* against empirical distributions, illustrating the flexibility and accuracy of *liFreq*. The distributions estimated are less smooth than the parametric distributions, but are able to capture irregularities, such as the flatness of $t = 200$. Nevertheless, the overall trends of the frequency approach and parametric approach are similar, see Figure 5.12.

The coefficients reveal that the *liFreq* regressions are capable of capturing some fundamental properties of the distribution. The coefficient function $\hat{\xi}_i(b)$ for exogenous variable i is expected to be continuous and smooth and represents the sensitivity of all parts of the distribution to changes in the exogenous variables. Taking the expectation as $|\Omega_t| \to \infty$ of the coefficient integral from 0 to ∞ and applying the Fundamental Theorem of Calculus:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{\epsilon_t}}\left[\int_0^\infty \hat{\xi}_i(m)dm\right] &= \int_0^\infty \mathbb{E}_{\boldsymbol{\epsilon_t}}[\hat{\xi}_i(m)]dm = \int_0^\infty \frac{\mathbb{E}_{\boldsymbol{\epsilon_t}}[d\bar{v}(m)_t]}{dx_i}dm \\
&= \int_0^\infty \frac{dv(m)_t}{dx_i}dm = \frac{d}{dx_i}\int_0^\infty v(m)_t dm = \frac{d}{dx_i}[1] \\
&= 0
\end{aligned} \tag{5.29}$$

It is evident that the coefficient integral will converge towards 0 as $|\Omega_t|$ increases.[43] Compliance with this property is an indication of stability in the coefficients, and turns out to be superior for *liFreq*, as seen in Figure 5.13 versus corresponding figures for the logistic and the lagged models in Appendix C.XIV.

The fixed point intersection of each variable marks the bucket that is unaffected by movements in the exogenous variable, and the two sides are affected oppositely by a change in the underlying.

[42] The q'th quantile was defined by $\hat{y}_{q,t} = m | \int_0^m \hat{H}_{\boldsymbol{x}_t}(m')dm' = q$, using $\hat{H}_{\boldsymbol{x}_t}(m)$ from Equation 5.28.

[43] For this calculation, the coefficients from Equation 5.26 were used. If coefficients from Equations 5.27 or 5.24 are used, corresponding results are obtained.

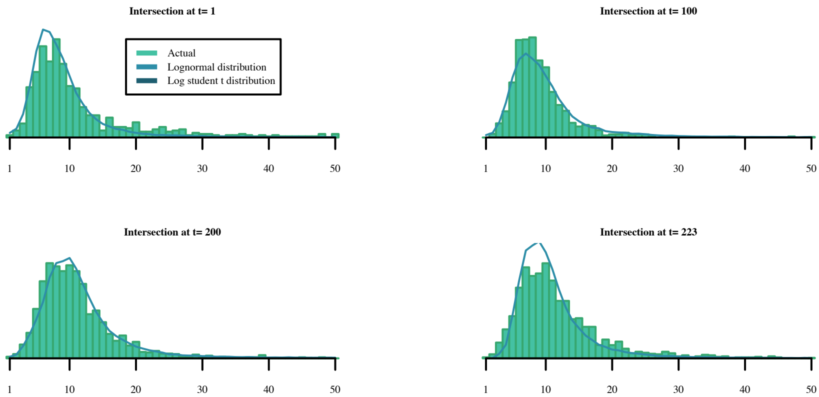


Figure 5.11: In-sample estimated *liFreq* distributions and empirical distributions

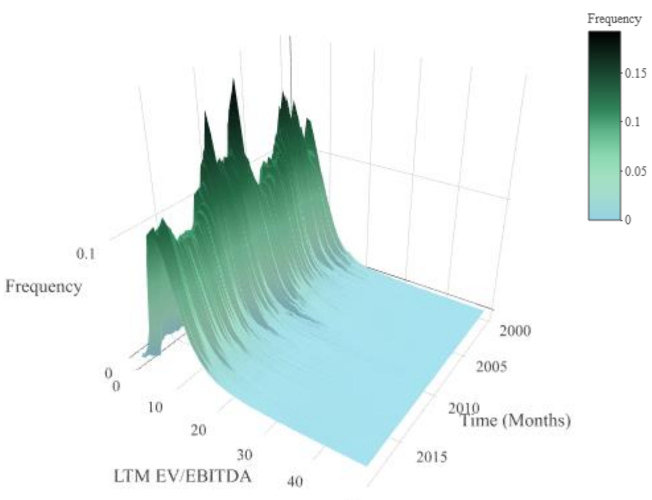


Figure 5.12: In-sample frequency model distribution of EV/EBITDA over time

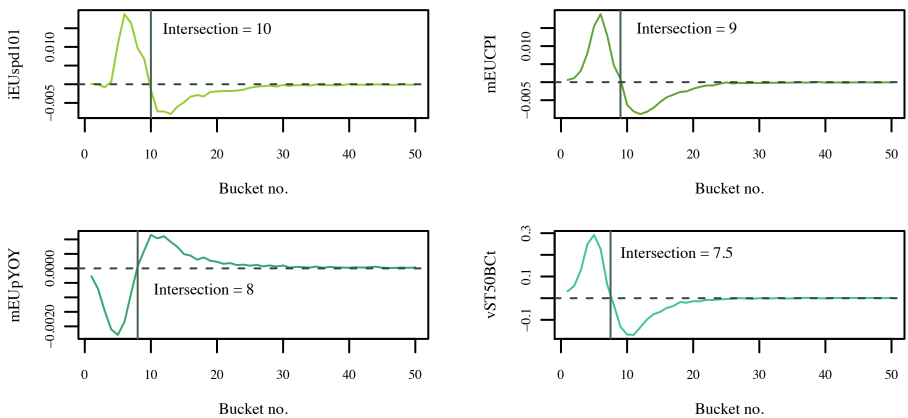


Figure 5.13: Coefficients of four exogenous *Ap* variables from 50 regressions using Equation 5.26

Moreover, if a variable changes in a way that increases the frequency of the lower buckets, the buckets of higher multiples than the x-intercept will see a decrease, and the flux through the distribution at the intercept value will be negative. This indicates that the variable affects the multiple negatively. The difference in the intercept points and shapes between the variables coefficients gives information about the behavior of the multiple, and the hypotheses of variable effects from Section 5.1 and 5.2 are supported and elaborated: while the frequency of 10x valued firms is unchanged by term structure spread movements, the same part of the distribution is the one most severely affected by volatility hikes.

5.4 Summary

A wide range of methods to model the dynamics of the distribution of multiples have been developed (i.e., approximations of H from Equation 5.1). The quantile methods (*qOLS*, *qVAR*, and *qPCA*) use quantiles or constructs from quantiles as the dependent variable, while the parametric distribution method use observed multiples as realizations of underlying stochastic variables. The frequency method (*liFreq*) is a very general method with few assumptions on the nature of the distribution dynamics and provides elaborations on the effects of individual exogenous variables. The results are highly consistent with the index analyses and across the different distribution modeling methods. All variables affect the whole distribution, but term structure spread is more important for high-valued companies, volatility is more important for medium- and low valued companies and inflation is important for the center of the distribution. Industrial production growth has a relatively even effect on the whole distribution of multiples.[44] The significance of the explanatory variables proposed in *Ap* is not as strong for the tails of the distribution, and it may be beneficial to utilize other variables in their description.

[44] As a supplement to the discussion, a simplistic analysis of how the exogenous variables impact low versus high valued companies is provided in Appendix C.XV. The results of the analysis are largely consistent with our findings here.

Chapter 6

Forecasting

After analyzing and understanding underlying macroeconomic drivers of multiples historically by examining indices and the complete distribution, we wish to establish a view on the prospects going forward. However, any forecast undertaking is subject to some notable pitfalls, and even the overall viability of scientific forecasting is debated (van Vught 1987). Hume's problem of induction may be used dogmatically to object any prediction - deterministic or probabilistic. Practically, however, forecasts serve as decision support and discussion material, and most mortals tolerate the assumption that the laws governing the past will govern the future. Of course, fully uncovering these laws is impossible, but the endeavor of science is to propose theories describing aspects of the laws. For econometric forecast models, empirical support is key, and models with causal backing in unrefuted economic theories are strengthened versus a stand-alone econometric model. Occams' razor is also core for developing and choosing plausible and transparent models (Popper & Bartley 1982). Nevertheless, Hume's objection remains, and an epistemological disclaimer is included in the sentiment of the models: We have to be *modest*.

We start by ignoring the macroeconomic environment and forecast y_t solely based on its history. In Section 6.2 the focus is shifted to the underlying environment, and we use publicly available forecasts and scenarios of the key variables along with our models to assess consequences for valuation multiples. In the last part, Section 6.3, we use vector models and Monte Carlo simulations to quantify uncertainty in the development of the underlying variables and propagate this into uncertainty concerning valuation multiples development.

6.1 Endogenous Forecasting

Some information that may be useful for forecasting of a time series can be apparent in the series' historical development: autocorrelation, trends, mean reverting tendencies and stationarity, and any forecasting based on underlying variables introduce an additional layer of uncertainty in the regression model (Ashley 1983). It is instructive to use models capturing these endogenous properties to gain an initial view on what development prospects we might expect. The models using no external information will be referred to as naïve models. Two trivial benchmark models can be constructed immediately. The Naive Martingale (*NM*): The forecast is the last observed value (Ville 1939). The Naive Gauss (*NG*): The forecast is the average of all previous forecasts. A natural extension of *NM* is the *AR(n)* model, in which forecast preserve the Markov property (Russell & Norvig 2003), but depend on parameters reflecting the complete process:[1]

$$y_t = \alpha + \gamma_1 y_{t-1} + ... + \gamma_n y_{t-n} + \epsilon_t \tag{6.1}$$

MA(m) models also preserve the Markov property, but residuals are used to smooth out a fixed part of the process - a natural extension of the *NG* model:

$$y_t = \alpha + \epsilon_t + a_1\eta_{t-1} + ... + a_m\eta_{t-m} \tag{6.2}$$

Combining these two models, we obtain the *ARMA(n,m)*-model:

$$y_t = \alpha + \gamma_1 y_{t-1} + ... + \gamma_n y_{t-n} + \eta_t + a_1\eta_{t-1} + ... + a_m\eta_{t-m} \tag{6.3}$$

and by using d_t instead of y_t, an *ARIMA(n,1,m)* is obtianed.

The model parameters were estimated by maximizing the associated likelihood function using the algorithm proposed by Nelder & Mead (1965) implemented in the arima-method from *stats* in R. A range of n and m values for the models were tested in an out-of-sample prediction test for six horizons,[2] but no lag performed significantly better than $m, n = 1$, see Table 6.1. *NM* and models forecasting something similar to the current do well on short horizons, while *NG* and models with a mean reverting tendency do better on long horizons. The *AR(n)* and *ARMA(n,m)* seem to adopt features from both strategies. No model is superior, but *ARMA(1,1)* and *AR(1)* perform well on all horizons and are pruned by Occam's razor.

Table 6.1: Root Mean Square Errors for out-of-sample prediction test for a selection of models on six horizons

	Forecast horizon					
	1 month	2 months	6 months	1 year	2 years	5 years
AR(1)	0.053	0.119	0.394	0.839	1.678	2.598
AR(2)	0.054	0.125	0.429	0.922	1.699	2.258
AR(5)	0.055	0.129	0.497	1.122	1.897	2.303
MA(1)	0.525	1.758	1.827	1.917	2.037	2.260
MA(2)	0.272	1.090	1.817	1.908	2.031	2.254
MA(5)	0.109	0.386	1.801	1.895	2.024	2.249
ARMA(1,1)	0.053	0.122	0.414	0.888	1.671	2.270
ARMA(2,2)	0.055	0.127	0.469	0.972	1.576	2.341
ARIMA(1,1)	0.053	0.121	0.394	0.804	2.061	3.526
ARIMA(2,2)	0.053	0.121	0.394	0.806	2.038	3.510
Naïve Martingale	0.052	0.115	0.375	0.810	1.927	3.514
Naïve Gauss	1.749	1.767	1.835	1.924	2.042	2.264

[1] The *ARIMA(n,d,m)* models and their assumptions will only briefly be presented here. See Tsay (2010) or Alexander (2008*a*) for elaboration.

[2] The data was separated into test- and training data according to the following indexing: $D_{\text{test}} = y_{(T-h-N+i):T}$ and $D_{\text{training}} = y_{1:(T-h-N+i-1)}$ for $T = 223$, horizons $h \in \{1, 2, 6, 12, 24, 60\}$, sample size $N = 50$ and $i \in [1, N]$. This test is biased towards testing on later parts of the sample, based on early parts, but random sampling would alter the autocorrelation, and correlated sampling processes would be biased towards prediction models resembling the generative model.

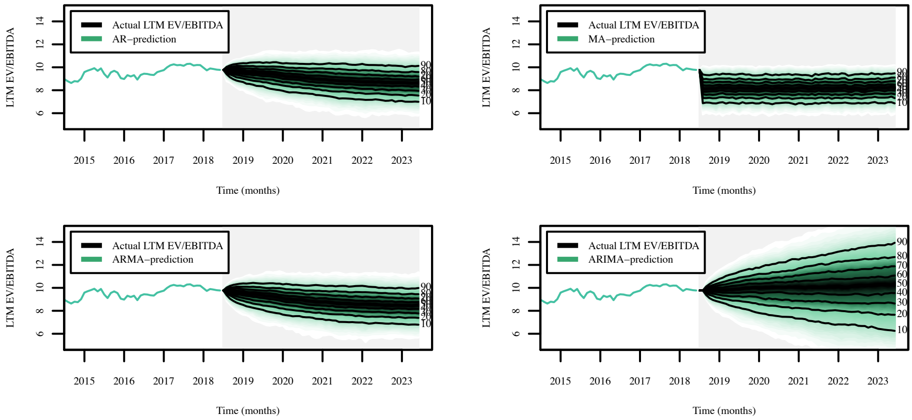


Figure 6.1: Monte Carlo simulations for four different models using the pricing index solely

Figure 6.1 displays the results from a Monte Carlo simulation of *AR(1)*, *MA(1)*, *ARMA(1,1)* and *ARIMA(1,1)*.[3] The first three models predict a mean reversion of the multiples over the next five years, while *ARIMA(1,1)* predicts a slight increase from current valuation levels, continuing the increasing trend observed towards current levels. The results are entangled with the model assumptions, and the question of which model to believe in must be answered with caution. Nevertheless, the predictions provide an initial understanding of the kinds of development that might be expected for y_t, and looking at the prediction tests in Table 6.1 for selection guidance, a future multiple decline arises as a supposition.

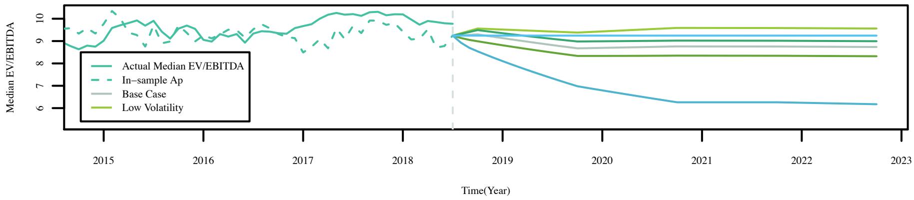
6.2 Scenario-based Forecasting

In the preceding analysis, the macroeconomic environment was ignored, although the models from Chapter 4, particularly *Ap* (see Figure 4.2), were shown to be suitable for long-term predictions when the development of the macroeconomic variables is known. This development is of course not known, but may be forecasted. In Chapter 3, exogenous variables were chosen based on criteria of predictability, and we hold these variables to be easier to predict than the multiple indices by themselves, but more importantly: the variables are meaningful for investors and represent statements they may agree or disagree with. The forecasts inherently introduce new uncertainties and complications but provide transparency and intuition into the multiple development forecast.

Modeling macroeconomic environment dynamics may be done using vector autoregressions, system dynamics, equilibrium theories, or a range of other methods. Instead of implementing this from scratch, we rely on scenarios based on estimates from a range of models created by analysts and institutions. The base case scenario is based on median consensus estimates for the macroeconomic variables obtainable from Factset.[4] No volatility forecasts were available, and the base case was set to be at a conservative increase from the current remarkably low levels. The accuracy of consensus

[3]The forecasts are obtained by 1,000 Monte Carlo simulations drawing residuals randomly with substitution from the model residuals. For *ARMA(1,1)*: $y_{t,s} = \alpha + \gamma_1 y_{t-1,s} + u_{st} + a_1 u_{s,t-1}$, where y_{ts} is simulation s at time t, and u_{ts} a randomly drawn residual. The residuals are not alarmingly autocorrelated, except for *MA(1)*, see Appendix D.I.

[4]See Section 3.3 for elaboration on how the base case scenario is constructed.



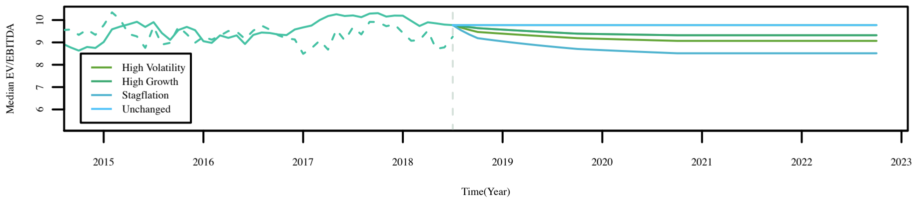


Figure 6.2: Multiple development scenarios using *Ap* (top) and *Dp* (botttom)

estimates is often questionable (Dreman & Berry 1995), and the estimates are likely to have less internal consistency than individual forecasts, but are convenient, transparent and fairly consistent with the beliefs in the industry. The base case scenario is used to derive other scenarios[5] with varying volatility and growth:

- **No-change Scenario:** All variables remain at the current level.
- **Base Case Scenario:** Volatility increases from the current level to the median level the last five years. Other variables are based on consensus estimates.
- **Low Volatility Scenario:** Volatility remains at the current level. All other variables are based on consensus estimates.
- **Increasing Volatility Scenario:** Volatility increases to the overall median last twenty years, which is still fairly conservative. All other variables remain as in the base case scenario.
- **High Growth Scenario:** Industrial production growth increases to the 90% quantile historically, i.e., 5.4 % growth. All other variables remain as in the base case scenario.
- **Stagflation Scenario:** This is a worst-case scenario, in which growth halts, inflation, volatility and term structure spread increase dramatically.

Figure 6.2 displays the resulting y_t from projections of the scenarios using the path independent predictable models, *Ap* and *Dp*. None of the scenarios result in an increase of multiples from current level because the prevailing macroeconomic conditions are strong, in particular the low volatility. Predictions of *Ap* result in larger dispersion than *Dp*, and while *Ap* predictions reflect the multiple justified by the macroeconomic variables, *Dp* predictions reflect short term dynamics in the macroeconomic variables. Evidently, *Ap* is accurate long-term and *Dp* is accurate short-term, and their performance intersection is at ~ 6 months.[6] *Ap* may be adjusted to account for the current environment by employing the autocorrelated residuals forecasting, Equation 4.10 from Section 4.4, but in order to capture both the current environment and the short-term dynamics modeled by *Dp*, we

[5]See Appendix D.II for numerical values of the scenarios.

[6]I.e., the predicative performance of *Ap* and *Dp* are the same on a six month horizon. See Figure 4.2 in Chapter 4.

Table 6.2: Test RMSE results from in-sample prediction for *Ap*, *Dp* and Combined prediction model

Model	1 month	6 months	1 year	2 years	5 years
Ap-model	0.552	0.553	0.541	0.549	0.543
Dp-model	0.215	0.644	1.025	1.387	1.727
Combined model	0.213	0.470	0.526	0.529	0.544

propose a Combined forecast model:[7]

$$y_h^c = y_h^a(1 - e^{-c(h-i)}) + y_h^d e^{-c(h-i)} \tag{6.4}$$

where y_h^c is the Combined model prediction in horizon h based on *Ap* (y_h^a) and *Dp* (y_h^d), with parameters i and c. Now $i = 0$ such that $y_0 = y_0^d$, and $c = \frac{\log 2}{6}$ such that $y_6^c = \frac{1}{2}y_6^a + \frac{1}{2}y_6^d$.[8] This forecast model takes the current macroeconomic environment into account, models short-term fluctuations and converges towards the level justified by *Ap* long-term. In fact, reversion from the current level to *Ap* levels mirrors the solution to an Ornstein-Uhlenbeck process[9] with $\mu_t = y_t^a$ given by the *Ap* prediction. Table 6.2 displays accuracy numbers from the usual in-sample prediction test.[10] The Combined model performs as good as the best of *Ap* and *Dp* on all horizons, except for $h = 6$, in which information from *Ad* and *Dp* are synthesized to create superior accuracy. Figure 6.3 shows the development forecasted by *Ap*, *Dp* and the Combined model applied to the no-change scenario. The prediction of the Combined model is an exponential decay from the current level to the level currently justified by macroeconomic conditions. Figure 6.4 shows predictions using the Combined model for all scenarios. The low volatility scenario is the only scenario remaining approximately at the current multiple level, while all others yield significant downturns in median multiple. Even the unchanged scenario predict a decrease in median multiple because it reverts to the current *Ap* prediction. The base case prediction is a significant multiple contraction, stabilizing at $\sim 8.7x$, down from the current $9.8x$ multiple. The high volatility scenario predicts a downturn to $\sim 8.3x$, while the bearish stagflation scenario gives a $\sim 6.2x$ multiple in 2021.

In order to understand the full multiple valuation impact of the scenarios, they are applied to the distribution prediction models developed in Chapter 5. Figure 6.5 displays the predicted multiple distribution at the beginning of 2021 for each scenario using the log Student-t distribution and the frequency method, respectively. All scenarios except the low volatility scenario entail a shift of the distribution towards lower multiples, but also non-trivial information about dynamics are revealed, such as dispersion and fatness of tails. As discovered in Chapter 5, the high multiple tail is less affected by increases in volatility than other parts of the distribution. Hence, in the low volatility scenario, parts of the high multiple tail (40x-50x) have lower densities than for the high growth

[7] An exponentially decaying weighting is used to get a smooth transition between the two models. Other functions, such as linear and Sigmoid functions, were considered but yielded inadequate prediction test results. Also, the exponential decay mirrors Equation 4.10 and is inspired by an Ornstein-Uhlenbeck process. See footnote 9.

[8] The parameters and the weighting functions may be subject to more sophisticated methods of estimation, such as likelihood maximation for the whole system, but it is probable that this would eliminate explanatory power from exogenous variables, as seen in Section 4.4. See Appendix E.I for a note on further research of this subject.

[9] See Uhlenbeck & Ornstein (1930) for elaboration on process. The Ornstein-Uhlenbeck process is defined by $dy_t = c(\mu - y_t)dt + \sigma dW_t$, and the solution is $y_t = y_0 e^{-ct} + \mu(1 - e^{-ct}) + \sigma \int_0^t e^{-c(t-s)} dW_s$, where the expected value of the Wiener Process dW_s is 0.

[10] The same test was used here as in Figure 4.2, i.e., given the current and future macroeconomic variables, and complete training data, predict the h horizon value.

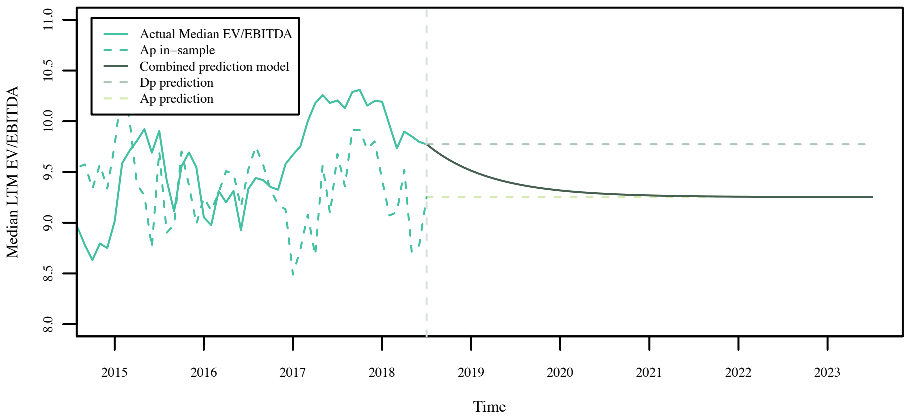


Figure 6.3: Median EV/EBITDA multiple development forecasted by *Ap*, *Dp* and the Combined model with unchanged macroeconomic environment

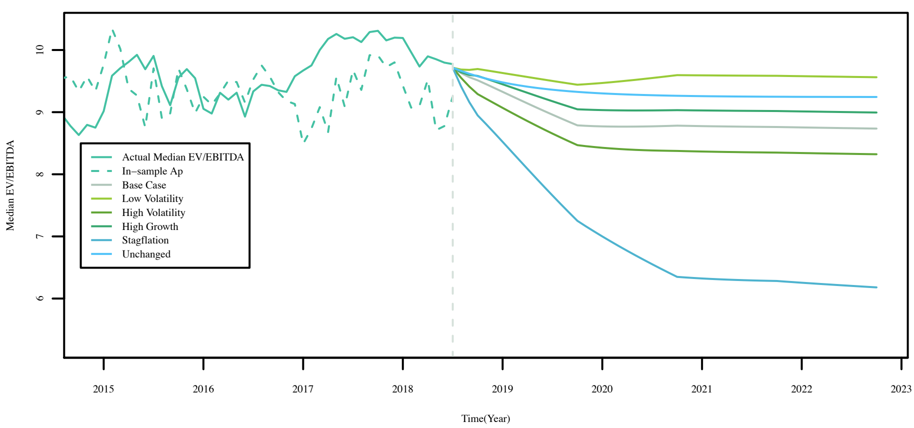


Figure 6.4: Multiple development scenarios using Combined model

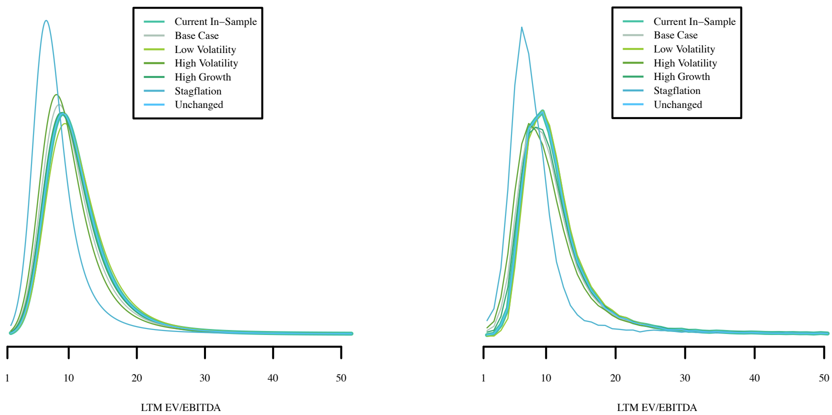


Figure 6.5: Multiple distribution development forecasted by the log Student-t parametric distribution model and the frequency distribution model for the six specific scenarios in beginning of 2021

scenario, although the low volatility scenario implies a significantly stronger median multiple. The multiple distribution predicted in the base case scenario at the beginning of 2021 is relatively consistent across models, as seen in Figure 6.6, but the log Student-t distribution is somewhat shifted to the right, and the frequency model has a dent in the top. The full dimensionality of the base case distributional predictions using the frequency model is presented in Figure 6.7, and the cross-distributional impact of the hypothesized multiple decline is evident.

6.3 Vector Model Forecasting

Although scenario modeling is transparent and easy to interpret, a severe disadvantage is that nothing can be said about uncertainty in the predictions. The future can never be predicted exactly, and probabilities can never describe the future completely, as discussed in the introduction to this chapter. Nevertheless, if some fundamental assumptions on the underlying process are made, a model can be created representing one *view* on the future, conditional on the assumptions. Instead of postulating scenarios, we will in this section postulate models of the behavior of the exogenous variables, and use Monte Carlo simulations and the models from Chapter 4 and Chapter 5 to establish a view on the development of multiple indices and distributions.

If the exogenous variables were independent, they could have been modeled with individual ARMA-models, and if they were serially independent, they could have been modeled as vector white noise drawing from an estimate of their joint distribution,[11] but in our case, the variables are both serially and cross-correlated. A VARIMA model[12] is versatile and handles serial correlation, interdependencies, and nonstationarity. However, because the exogenous variables are cointegrated a VARMA model is more appropriate.[13] Despite the arguments presented by Athanasopoulos &

[11] E.g., by assuming multinormality, and to use a Cholesky decomposed covariance matrix estimate along with the mean-vector to obtain simulations through drawing of standard normal variables.

[12] The terminology is assumed to be known. See Tsay (2010) for an explanation.

[13] See Appendix D.III for Johansen test results indicating cointegration. Toda & Yamamoto (1995) argue that VAR models are appropriate for cointegrated variables. Furthermore, the problem of Section 6.1 is repeated with predictions from a VARIMA-model: Differentiated models yield trivial flat predictions.

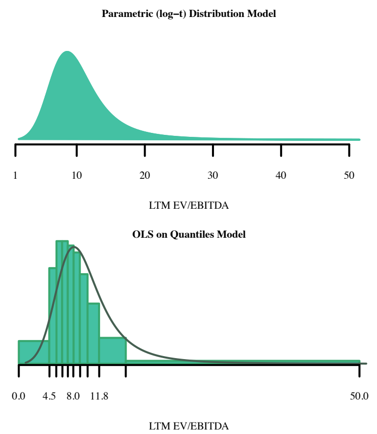


Figure 6.6: Multiple distribution prediction in beginning of 2021 in the base case scenario for models developed in Chapter 5. Log Student-t distribution are plotted in all charts for comparison (black line)

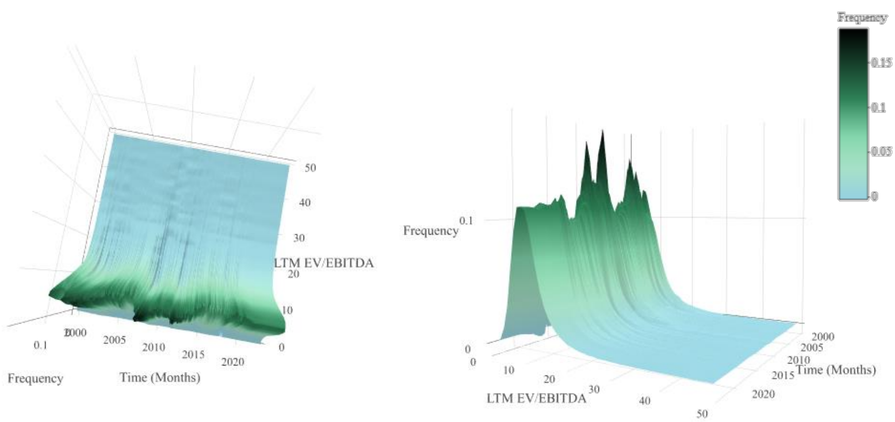


Figure 6.7: In-sample multiple distribution over time for the frequency model including base case scenario predictions

	Spread	Inflation	Growth	Volatility	EV/EBITDA
Current value	1.10%	2.10%	1.24%	12.6%	9.77x
Historical median	1.27%	2.00%	1.80%	24.0%	8.04x
Historical mean	1.24%	1.74%	1.00 %	22.1%	8.14x
Steady state solution	1.18%	1.72%	0.93%	21.5%	8.22x

Table 6.3: Steady state solution without fixed variables for VAR-*Ap* model

Vahid (2008), we ignore the MA terms as these complicate estimation and simulation adding little explanatory power. A Structural VAR(1) model captures both current and 1-lag interactions of the variables (Tsay 2010):

$$\boldsymbol{x}_t = \boldsymbol{\alpha} + B^0\boldsymbol{x}_t + B^1\boldsymbol{x}_{t-1} + \boldsymbol{\epsilon}_t \tag{6.5}$$

where $\boldsymbol{x}_t$ are the exogenous variables, $\boldsymbol{\alpha}$ is a vector of constants, $\boldsymbol{\epsilon}_t$ is an i.i.d. error process vector with $\forall(i \neq j)\text{Cov}(\epsilon_i,\epsilon_j) = 0$, and B^i are the autoregressive coefficient matrices. B^0 is such that elements $b_{ii} = 0$. The SVAR model may be reformulated into a normal VAR model:

$$\begin{aligned}\boldsymbol{x}_t &= (I - B^0)^{-1}\boldsymbol{\alpha} + (I - B^0)^{-1}B^1\boldsymbol{x}_{t-1} + (I - B^0)^{-1}\boldsymbol{\epsilon}_t \\ &= \boldsymbol{a} + V\boldsymbol{x}_{t-1} + \boldsymbol{e}_t\end{aligned} \tag{6.6}$$

The reformulation renders identification of the original coefficients impossible, and the covariance matrix of the error process $\boldsymbol{e}_t$ becomes nonzero, but neither of these effects is problematic because the forecast simulations will be unaffected. The coefficients can be estimated using OLS, and the results for the underlying variables in *Ap* are attached in Table 22 in Appendix D.IV. The main predictor of each of the variables is naturally the previous realization of itself, but there are also significant cross-variable effects. The residuals from the VAR model are largely serially uncorrelated,[14] allowing for independent random drawings of residual vectors in the Monte Carlo simulation.

In the long-run, the exogenous variables in the VAR will converge to a steady state $\boldsymbol{x}$ such that $\boldsymbol{x}_t = \boldsymbol{x}_{t-1} = \boldsymbol{x}$ when $t \to \infty$. If $\det(I - \hat{V}) \neq 0$, we can take the expectation value of the set of equations 6.6 to obtain:

$$\boldsymbol{x} = (I - \hat{V})^{-1}\hat{\boldsymbol{\alpha}} \tag{6.7}$$

The steady state solution for the VAR model based on *Ap* is displayed in Table 6.3. The current value of the term structure spread at 1.1% is close to the steady state, while inflation is somewhat lower at 2.1% and industrial production growth is higher at 1.24 %. The variable furthest away from steady state is volatility, currently among the 2%-quantile lowest values, at 12.6%. From Section 4.1.2, volatility is known to have large impact on multiple levels and thus, a volatility driven multiple mean reversion may be hypothesized. The steady state solution is a relatively optimistic scenario compared to median and mean scenarios.

The Monte Carlo simulation for each model is conducted by providing the initial value $\boldsymbol{x}_T$ and then propagating the simulation vectors by drawing cross-correlated, but not serially correlated residuals, based on the historical distribution.[15] For a simulation s at horizon h, the vector of underlying variables is:

$$\boldsymbol{x}_h^s = \hat{\boldsymbol{\alpha}} + \hat{V}\boldsymbol{x}_{h-1}^s + \boldsymbol{u}_h^s \tag{6.8}$$

[14] See ACFs for the VAR model residuals in Figure 29 in Appendix D.IV.

[15] Two methods were tested: 1) Drawing directly from the observed vectors of $\hat{\boldsymbol{e}}_t$. 2) Fitting a multivariate kernel to $\hat{\boldsymbol{e}}_t$, (Panaretos & Konis 2012) and drawing from this. No significant differences between the methods were discovered.

where $\hat{\boldsymbol{\alpha}}$ and $\hat{V}$ are the estimators of coefficients from Equation 6.6 and $\boldsymbol{u}_t^s$ is a randomly drawn residual vector. Figure 6.8 displays the historical development and the simulated probability fans for underlying variables in *Ap* based on 1,000 simulations. As expected from a VAR model, there is a clear mean reversion in all underlying variables. Furthermore, the variables are affected by the development of each other, creating feedback motion resembling a damped harmonic oscillator, particularly for industrial production growth. One consequence of this is that the median of the growth variable simulations is expected to hit 0 in 2021, before reverting towards the steady state at $\sim 1\%$. In general, the overall trend of the variables is similar to the trends in the base case scenario, except that the latter is more bullish on growth and volatility. The model could have been calibrated to fit better with the base case scenario, but this would affect the integrity and simplicity of the model. The point of mean reversion for the growth variable is 0.93%, close to the historical mean of 1.00% and far away from the historical median at 1.80%, indicating that the model is influenced by the outliers of the 2007/2008 financial crisis.

The individual simulations $\boldsymbol{x}_h^s$ are projected through to y_t^s using the relationships from Chapter 4 and drawing regression model residuals from the historical distribution. For *Ap*, the following projection was conducted:

$$y_{T+1:H}^s = \hat{\alpha} + \hat{\boldsymbol{\beta}} \cdot \boldsymbol{x}_{T+1:H}^s + \epsilon_{T+1:H}^s \tag{6.9}$$

where $\hat{\alpha}$ and $\hat{\boldsymbol{\beta}}$ are estimated coefficients and ϵ_t^s is a residual from Equation 4.1. As discussed in Section 4.4, the residuals in *Ap* are autocorrelated, and this was managed by modelling the residual process explicitly using Equation 4.9. For each s, a vector of random residuals $\hat{u}_{T+1:H}^s$ is drawn from the residuals of Equation 4.9, and this residual vector is used to propagate correlated error processes from the initial value $\hat{\epsilon}_T$, resulting in $\hat{\epsilon}_{T+1:H}^s$. Corresponding Monte Carlo simulations were conducted for the *Dp* and *Rp* models.[16] The projection of individual simulations before aggregation is done to preserve the full dynamic of the models, allowing correct estimation of the time-varying distribution of developments. One essential difference between this way of forecasting compared to scenario forecasting is that we preserve the probability distributions and effectively convolute the VAR process residuals with the regression model residuals in order to create the final forecasts, allowing us to quantify uncertainty and the complete distribution of outcomes.

The probability fan in Figure 6.9 displays the quantiles of the projected development using *Ap* and the VAR. Accounting for autocorrelation in the residuals resolves the issue of the deviation between the actual and the *Ap* estimate because the autocorrelation between the final observed residual and the consecutive simulated residuals is taken into account explicitly in the initial value of the residual process. The fan median yields a mean reversion resembling the base case scenario but is slightly more pessimistic. The upside from the current multiple level is associated with the 90% quantile of developments, while the 10% worst-case prediction is a sub 7.0x multiple by early 2020. The outputs must be interpreted with caution, but give an indication of the distribution of outcomes given the underlying VAR model and *Ap*.

The Combined model introduced in Section 6.2 performed well in the prediction tests, and projecting the Monte Carlo simulation of the underlying variables using Equation 6.4, the results in Figure 6.10 are obtained.[17] The fan resembles the *Ap* fan but is less radical in initial variance because *Dp* yields smaller deviations and governs the process initially. Furthermore, the rate of decrease of the median is smaller because *Dp* predicts an integrated process, but the two distributions converge

[16] *Dp*: $y_t^s s = y_{t-1}^s e^{\beta_d \Delta \boldsymbol{x}_t^s + \epsilon_{t,s}}$. *Rp*: $y_t^s = \hat{\alpha}_r + \hat{\gamma} y_{t-1}^s + \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_t^s + \epsilon_t^s$. None of these models have the same issues with autocorrelation in the residuals as *Ap*.

[17] The simulations are projected into *Ap* and *Dp*, drawing correlated residuals for *Ap*:$y_h^s = \hat{\alpha} + \hat{\boldsymbol{\beta}} \cdot \boldsymbol{x}_t^s + u_t^s$, where u_t^s is generated in the explained way, and *Dp*: $y_t^s = y_t^s e^{\hat{\beta}_d (\log x_t^s - \log x_{t-1}^s) + u_t^s}$, where the residuals are drawn from the ones observed historically. These are combined using Equation 6.4 to form individual Combined model simulations. The resulting error process is a convolution of three error processes, VAR residuals, *Dp* residuals and *Ap* residuals.

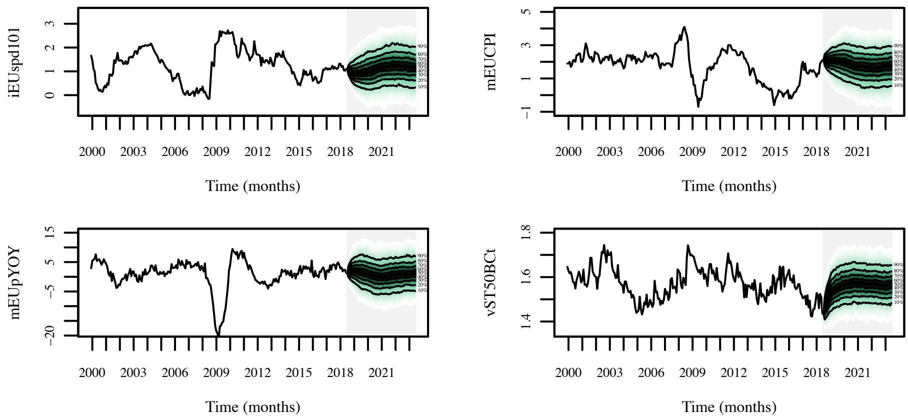


Figure 6.8: VAR model forecasting of underlying variables in *Ap*

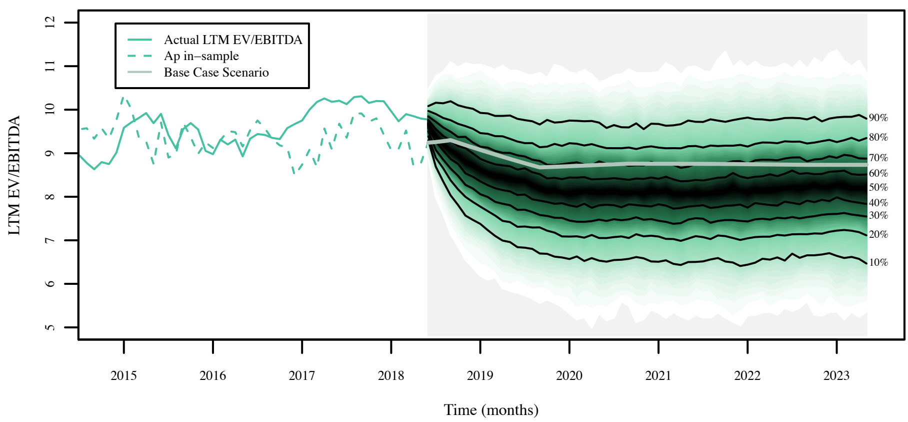


Figure 6.9: Probability fan for development of Median EV/EBITDA using VAR model for exogenous variables and *Ap* to project.

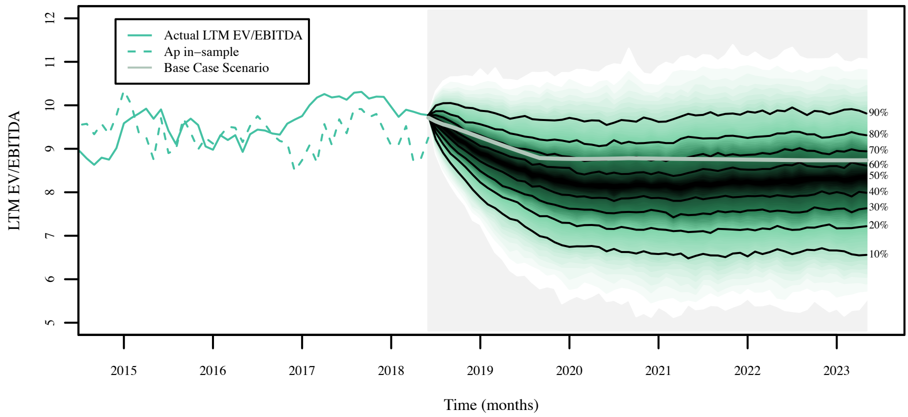


Figure 6.10: Probability fan for development of Median EV/EBITDA using VAR model for underlying variables and Combined *Ap* and *Dp*-models for projection.

long-term. The forecasts produced by the Combined model are proven to be accurate in-sample, and the predicted multiple decline is slightly less pessimistic but still severe.

The underlying simulations also provide an opportunity to form a view on uncertainty in development of distinct quantiles of the multiple distribution. In Figure 6.11 probability fans for development of 10% and 90% multiple quantiles using VAR model for exogenous variables and *qOLS* for projection are depicted. As for *Ap*, autocorrelations in residuals are mitigated by adopting an AR(1) model. The uncertainty for high valued companies is larger than for low valued companies, but not in relative terms. In Chapter 5, we emphasized that the current median multiple is lower than in 2007 but higher than in 2000, while the 90% quantile opposite: higher than 2007 but lower than in 2009. The probability fans do not forecast the development of expensive companies to be as dramatic as in these crises, possibly because OLS mitigates the impact of extreme events, necessitating the use of more complete distributional models in the forecasting. Nevertheless, the forecasted downturn of the quantiles is severe and give an indication of the shift that might be expected.

Each model from Chapter 5 is an instance of the function $F(h,s,m)$, mapping from horizon and simulation to a multiple distribution.[18] It is natural to enumerate the distributions according to a notion of *optimism* and *pessimism*, but the enumeration method is dubious. Unlike numbers, distributions do not have an inherent transitive ordering. A naïve way to produce output is to divide the Monte Carlo simulations into quantiles and propagate this ordering through to the distributions, but this does not account for cross-relationships between the series of underlying variables, and the scenarios will be unrealistic. The μ parameter of the log Student-t distribution from Section 5.2 is a location parameter and may be used to construct a transitive ordering of the simulation distributions. Since the variance in the other parameters is large, it is natural to perform smoothing of the parameters, e.g., when naming a quantile of the simulation distribution of the multiple distributions. Figure 6.12 depicts the 10%, 50% and 90% simulation distribution quantiles of the log Student-t distribution prediction at the beginning of 2021.[19] We can observe that current in-sample distribution

[18]In order to fully propagate the error process of the distribution, an autocorrelated continuous distributional residual process, $\hat{\varepsilon}(m)_h^s$ has to be defined. This is elaborated in Appendix E.I.

[19]The smoothing operation for the q% quantile parameters is calculated as the mean of parameter values between $(q-2.5)$% and $(q+2.5)$%.

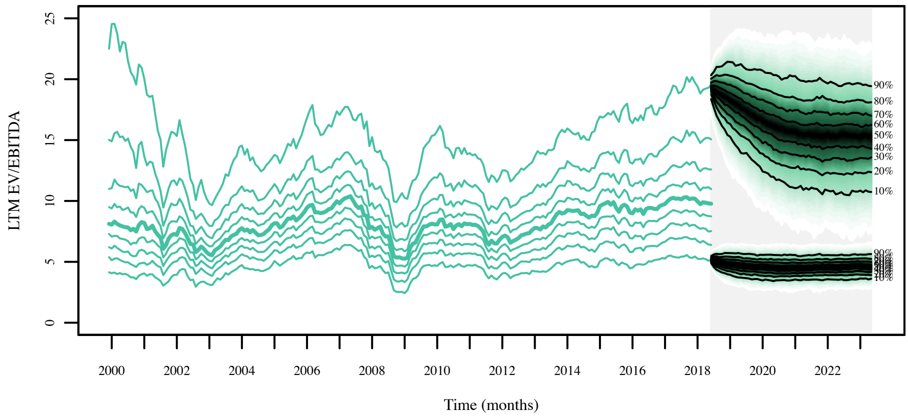


Figure 6.11: Probability fans for the development of the 10% and 90% quantiles of EV/EBITDA using VAR model for the exogenous variables and *qOLS* for projection.

of multiples is close to the predicted two year 90% quantile distribution, which is consistent with our forecasted probability fan using Monte Carlo simulations. Based on the same ordering of μ at the same point of time, Figure 6.13 illustrates the full range of distributional outcomes for all $1{,}000$ simulations, and cross-sectional medians are consistent with the range of median EV/EBITDAs at the same instance in Figure 6.9.

As well as being instances of $F(h,s,m)$, the models are instances of the function $F(\boldsymbol{x}_t, m)$, and this dimensionality allows for analyses of sensitivity of the distribution towards individual exogenous variables. In particular, volatility is an important factor, and Figure 6.14 displays the distributional sensitivity towards volatility varying from minimum to maximum level historically. All other exogenous variables are kept at the current level. Volatility has significant impact on the full distribution manifested by the shift in Figure 6.14. The resulting median multiple ranges from 6.6x and 9.5x illustrating the potentially severe effects of volatility increases.

The mean reversion produced by the models is not surprising, knowing that the current European median multiple is within the top 10% quantile of values since 2000. The only time in our sample to match this level is the months before the financial crisis in 2007. Proposed models struggle to justify the multiple levels in the current macroeconomic conditions, but as emphasized by Hume we cannot know that underlying relationships going forward will resemble the past. Nevertheless, we modestly conclude that a future multiple decline is likely.

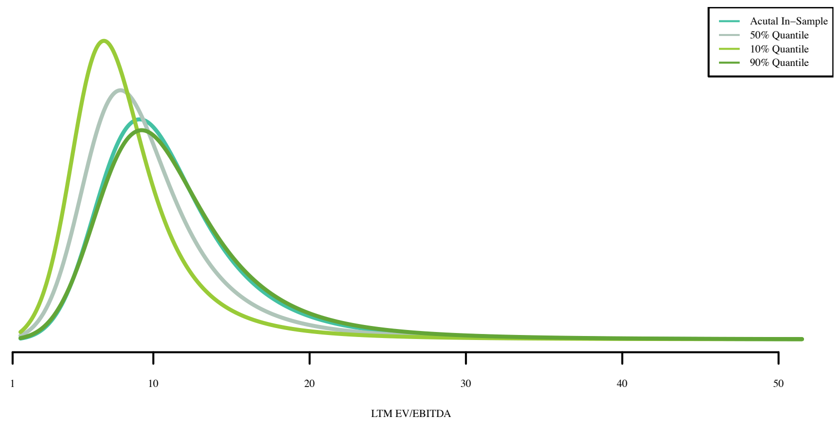


Figure 6.12: Log Student-t distribution model's 2021 prediction of distribution of multiples for quantiles of parameters derived from 1,000 Monte Carlo simulations

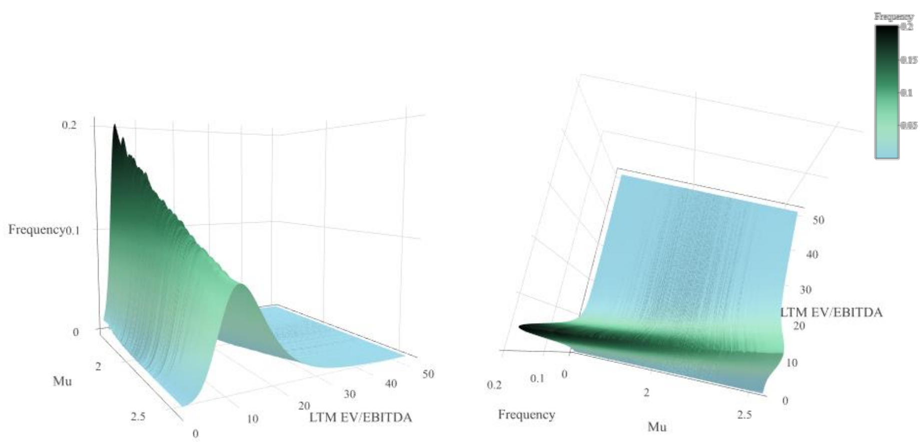


Figure 6.13: Multiple distribution in 2021 for all 1,000 Monte Carlo simulations of underlying variables using log Student-t distribution model

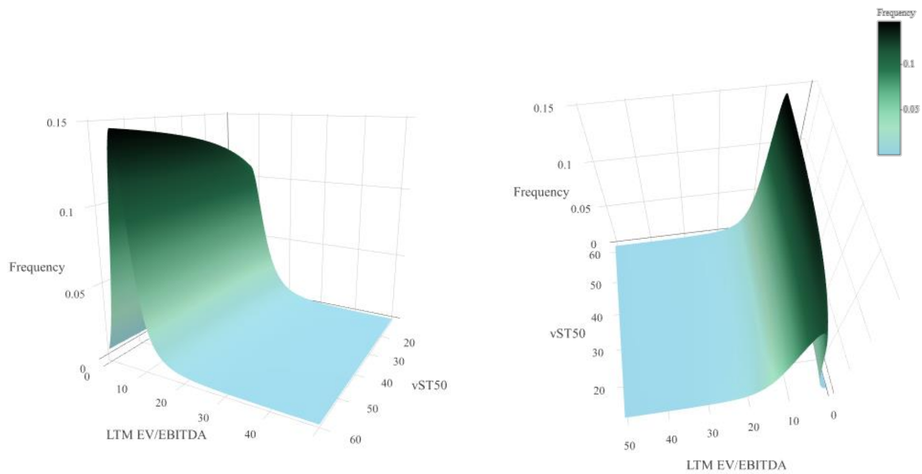


Figure 6.14: Distribution of multiples when varying volatility and keeping the other exogenous variables at current level

Chapter 7

Conclusion

Value, ethical or material, is a fundamental concept in ethics and philosophy, and definition and observance of the material part is a cornerstone in economics. Most econometric research on firm valuation focus on analysis of stock price return variables, while we have employed a Copernican Turn to understand aggregate firm valuation levels relative to profits:

$$m_c = \frac{V_c}{\pi_c} \tag{2.7}$$

The cointegrated *Ap* model (Equation 4.1) was used to identify implied volatility, inflation, term structure spread and industrial production growth as key drivers of the European median listed firm LTM EV/EBITDA multiple, and the response was as expected from economic theory introduced in Chapter 2.

$$y_t = \alpha + \boldsymbol{\beta} \cdot \boldsymbol{x_t} + \epsilon_t \tag{4.1}$$

Building on *Ap*, distributional dynamics of multiples were modeled from the quantile, parametric and frequency perspectives, obtaining consistent results regarding the distributional dynamics in response to changes of underlying variables: implied volatility disproportionally impacts low multiples, while term structure spread affect high multiples. Inflation primarily affects the center of the distribution, and industrial production growth has a balanced impact. The models produce a complete cross-sectional distribution based on four exogenous variable data points at one instance t.

Forecasting the distribution and the aggregates is essentially an attempt to predict changes in perceptions of future conditions. Triangulated by endogenous forecasting models, scenario analysis and simulations of underlying variables, we are unable to justify prevailing multiple levels in the current macroeconomic environment. This gives us reason to believe that we are at the verge of a turning point in the business cycle.

Our research builds a platform from which further research that expands domain, accounts for new assumptions or aims to improve accuracy may depart, see Appendix E.I for a non-exhaustive list of ideas.

Other datasets and granularities may be used to enhance model performance. Our research was exclusively concerned with rational factors, and a natural extension is to include irrational, behavioral indicators in the exogenous variables, e.g., obtained through semantic analyses of social media data. By segmenting the data into industries, geographies and other classes, hypotheses on class-specific drivers and multiples[1] can be tested. A firm size independent β multiple was proposed in Chapter 5, and it would be interesting to investigate implications of its implementation and to do corresponding analyses for other multiple types.

There is a vast range of econometric models that may capture additional features of the multiple index processes, leading to more accurate descriptions of causality and predictions. Heteroskedasticity is significant and may be dealt with using a GARCH-model. Autocorrelation in *Ap* is dealt with using the exponential ARMA-family of models, but Long memory models may be more adequate in explaining the slow decay in the ACF from Figure 4.9 (Mandelbrot & Van Ness 1968). Further advancements can likewise be made to the multiple distribution models. Solving $\hat{\boldsymbol{\beta}}_{1:B} = \text{argmax}_{\boldsymbol{\beta}_{1:B}} \mathscr{L}$ in Equation 5.23 would yield the mathematically optimal parameters for the frequency model, but the issue of overfitting may become pressing. These issues may be depressed by imposed constraints on frequencies, e.g., by dividing buckets into monotonous sections, or correspondingly for differentiated buckets. In principle, the inclusion of B and individual bucket ranges as parameters of $\mathscr{L}$ may yield an optimal solution to the bucket allocation problem. Increased parametric flexibility may also be a natural next step for the parametric method, e.g., by fitting a five-parametric generalized *skewed* t distribution, using a spline composition strategy or by modeling the tails using extreme value analysis (Fisher & Tippett 1928). In Appendix E.I, an approach inspired by Quantum mechanics is also proposed.

Cylicality and timing effects are not modeled in a sophisticated way in our models, but Hamilton (1989) introduces Markov Switching Models (MSM) for describing regime nonstationarity, and we propose describing the macroeconomic factors with a two-state MSM with switching probabilities dependent on the absolute level of the macroeconomic variables (Diebold et al. 1993):[2]

$$i \in \{1,2\}, \Delta \boldsymbol{x}_t = \boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_{ti} \tag{7.1}$$

$$P(s_t = 1) = \frac{P(s_{t-1} = 1)}{1 + e^{\boldsymbol{\beta}_1 \boldsymbol{x}_{t-1}}} + \frac{e^{\boldsymbol{\beta}_2 \boldsymbol{x}_{t-1}} P(s_{t-1} = 2)}{1 + e^{\boldsymbol{\beta}_2 \boldsymbol{x}_{t-1}}} \tag{7.2}$$

This model outputs macroeconomic environment dependent probabilities of a regime change which can be used to model the probability of a financial crisis or recession, but we were unable to estimate the parameters using MLE.[3] The underlying variable modelling may be extended further by modelling causal relationships explicitly through a system dynamic model including nonlinear feedback loops, decision variables, irrational responses and policy effects. Inclusion of panel data and panel analysis in the system dynamic model allows contemporaneous modelling of individual firm development, industry-specific dynamics and interaction patterns. Panel analyses by itself may be useful for idiosyncratic risk modelling, and for obtaining a bottom-up perspective on the distribution of corporations (Matyas & Sevestre 1996).[4]

Combining the above-mentioned model extensions, a sophisticated and comprehensive, but complex and potentially unintelligible, model will be obtained. Overfitting, data quality, and interpretability become problematic, and Hume's problem of induction remains: We do not know that the future will resemble the past. Our approach to modeling has been pragmatic, exploratory,

[1]Industry-specific multiples include EV/Reserves for upstream O&G, EV/# of users for technology and Price/Book Value for the financial industry.

[2]$\boldsymbol{x}_t$ are exogenous variables, $\boldsymbol{\epsilon}_i$ is an i.i.d. error process and $\boldsymbol{\alpha}_i$, $\boldsymbol{\beta}_i$, $P(s_0 = 1)$, the joint residual distribution parameters (assuming multinormality) Σ and $\boldsymbol{\mu}$ are model parameters.

[3]Explicit maximization of $\mathscr{L} = \sum_{t=1}^{T} \log \sum_{i=1}^{2} f_{\boldsymbol{\epsilon}_{ti}} (\log \frac{\boldsymbol{x}_t}{\boldsymbol{x}_{t-1}} - \boldsymbol{\alpha}_i)(\frac{P(s_{t-1}=1)}{1+e^{\boldsymbol{\beta}_1 \boldsymbol{x}_t}} + \frac{e^{\boldsymbol{\beta}_2 \boldsymbol{x}_t} P(s_{t-1}=2)}{1+e^{\boldsymbol{\beta}_2 \boldsymbol{x}_t}})$ is $O(2^T)$ and intractable, so we developed an EM-based algorithm for estimation but were unable to get satisfactory results (Dempster et al. 1977). See Appendix E.I

[4]Panel analysis is concerned with the regression $y_{ct} = \alpha_c + \boldsymbol{\beta}_c \cdot \boldsymbol{x}_t + \epsilon_{ct}$, and in particular the description of ϵ_{ct}. See Appendix E.I.

transparent and focused on practical consequences rather than theoretical deviations. Occams' razor has been used diligently to uncover consequences of the Copernican turn from Chapter 2, and the forecasts are unambiguous: the current multiple level of 9.8x is unsustainable. Our mean reverting vector model predicts a 90% confidence range of 6.4x to 9.6x by early 2020. A decline is likely to be catalyzed by increased volatility, as the current index is close to an all-time low at 12.6%. We consider this a paradox given the prevailing high level of global geopolitical risk.[5] We can do nothing but present a troubling question: *Will geopolitical risk trigger investor fear and push the world into a prolonged period of lower multiples?*

[5]CBOE Volatility Index (VIX) and Global Economic Policy Uncertainty Index have followed each other closely historically. However, in recent years VIX has been decreasing, while global economic policy uncertainty has moved in the opposite direction. See Appendix E.II for indexed plot and calculation of correlation coefficients.

Bibliography

Al-Anaswah, N. & Wilfling, B. (2011), 'Identification of speculative bubbles using state-space models with markov-switching', *Journal of Banking Finance* **35**(5), 1073 – 1086.

Alexander, C. (2008*a*), *Market Risk Analysis*, Vol. II, John Wiley Sons, Ltd.

Alexander, C. (2008*b*), *Market Risk Analysis*, Vol. I, John Wiley Sons, Ltd.

Alvarez-Cuadrado, F. & El-Attar Vilalta, M. (2012), 'Income inequality and saving', *Oxford Bulletin of Economics and Statistics* .

Ang, A., Piazzesi, M. & Wei, M. (2004), 'What does the yield curve tell us about gdp growth?', *National Bureau of Economic Research* (10672).

Ashley, R. (1983), 'On the usefulness of macroeconomic forecasts as inputs to forecasting models', *Journal of forecasting.* **2**(3), 211–223.

Athanasopoulos, G. & Vahid, F. (2008), 'Varma versus var for macroeconomic forecasting', *Journal of Business & Economic Statistics* **26**(2), 237–252.

Baker, M. & Wurgler, J. (2007), Investor sentiment in the stock market, Working Paper 13189, National Bureau of Economic Research.

Baker, S. R., Bloom, N. & Davis, S. J. (2016), 'Measuring economic policy uncertainty', *The Quarterly Journal of Economics* **131**(4), 1593–1636.

Baltagi, B. (2008), *Econometric analysis of panel data*, John Wiley & Sons.

Barberis, N., Huang, M. & Santos, T. (1999), Prospect theory and asset prices, Working Paper 7220, National Bureau of Economic Research.

Barberis, N. & Thaler, R. (2002), A survey of behavioral finance, Working Paper 9222, National Bureau of Economic Research.

Barndorff-Nielsen, O. (1977), 'Exponentially decreasing distributions for the logarithm of particle size', *Proceedings of the Royal Society of London* **353**(1674), 401–419.

Basu, S. (1977), 'Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis', *The journal of Finance* **32**(3), 663–682.

Baumohl, B. (2012), *The Secrets of Economic Indicators: Hidden Clues to Future Economic Trends and Investment Opportunities*, Pearson Education.

Beaver, W. & Morse, D. (1978), 'What determines price-earnings ratios?', *Financial Analysts Journal* **34**(4), 65–76.

Bernanke, B. (2005), The global saving glut and the us current account deficit, Technical report, Board of Governors of the Federal Reserve System (US).

Bernanke, B. S. & Mishkin, F. S. (1997), 'Inflation targeting: A new framework for monetary policy', *Journal of Economic Perspectives* **11**(2), 97–116.

Beveridge, S. & Nelson, C. R. (1981), 'A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the business cycle', *Journal of Monetary Economics* **7**(2), 151 – 174.

Black, F. (1976), 'The dividend puzzle', *Journal of portfolio management* **2**(2), 5–8.

Black, F. & Scholes, M. (1973), 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**(3), 637–654.

Blumer, A., Ehrenfeucht, A., Haussler, D. & Warmuth, M. K. (1987), 'Occam's razor', *Information Processing Letters* **24**(6), 377 – 380.

Bourguignon, F. (1981), 'Pareto superiority of unegalitarian equilibria in stiglitz'model of wealth distribution with convex saving function', *Econometrica: Journal of the Econometric Society* pp. 1469–1475.

Box, G. E. P. & Cox, D. R. (1964), 'An analysis of transformations', *Journal of the Royal Statistical Society. Series B (Methodological)* **26**(2), 211–252.

Campbell, J. Y. (1987), 'Stock returns and the term structure', *Journal of Financial Economics* **18**(2), 373 – 399.

Campbell, J. Y. & Mankiw, N. G. (1987), Permanent and transitory components in macroeconomic fluctuations, Working Paper 2169, National Bureau of Economic Research.

Campbell, J. Y. & Shiller, R. J. (1988), 'Stock prices, earnings, and expected dividends', *The Journal of Finance* **43**(3), 661–676.

Chandler (1962), 'Strategy and structure: Chapters in the history of the industrial enterprise. by alfred d. chandler, jr.', *Journal of American History* **49**(3), 536–537.

Chen, S.-S. (2009), 'Predicting the bear stock market: Macroeconomic variables as leading indicators', *Journal of Banking Finance* **33**(2), 211 – 223.

Cheng, A. C. (1995), 'The uk stock market and economic factors: A new approach', *Journal of business finance accounting.* **22**(1), 129–142.

Chinn, M. D. & Ito, H. (2007), 'Current account balances, financial development and institutions: Assaying the world "saving glut"', *Journal of international money and Finance* **26**(4), 546–569.

Cochrane, D. & Orcutt, G. H. (1949), 'Application of least squares regression to relationships containing auto-correlated error terms', *Journal of the American statistical association* **44**(245), 32–61.

Corana, A., Marchesi, M., Martini, C. & Ridella, S. (1987), 'Minimizing Multimodal Functions of Continuous Variables with the "Simulated Annealing" algorithm', *ACM Trans. Math. Softw.* **13**(3), 262–280.

Cormen, T. H., Leiserson, C. E., Rivest, R. L. & Stein, C. (2009), *Introduction to algorithms*, MIT press.

Davis, P. J. (1959), 'Leonhard euler's integral: A historical profile of the gamma function: In memoriam: Milton abramowitz', *The American Mathematical Monthly* **66**(10), 849–869.

Dempster, A. P., Laird, N. M. & Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via the em algorithm', *Journal of the royal statistical society. Series B (methodological)* pp. 1–38.

Dennis, J. (1983), 'Rb schnabel numerical methods for unconstrained optimization and nonlinear equations prentice-hall', *New Jersey* .

Dickey, D. A. & Fuller, W. A. (1981), 'Likelihood ratio statistics for autoregressive time series with a unit root', *Econometrica* **49**(4), 1057–1072.

Diebold, F., Lee, J.-H. & Weinbach, G. C. (1993), Regime switching with time-varying transition probabilities, Working Papers 93-12, Federal Reserve Bank of Philadelphia.

Ding, Z., Granger, C. W. & Engle, R. F. (1993), 'A long memory property of stock market returns and a new model', *Journal of Empirical Finance* **1**(1), 83 – 106.

Dixit, A. K. & Pindyck, R. S. (1994), *Investment under uncertainty*, Princeton university press.

Dreman, D. N. & Berry, M. A. (1995), 'Analyst forecasting errors and their implications for security analysis', *Financial Analysts Journal* **51**(3), 30–41.

Edgeworth, F. Y. (1884), 'The philosophy of chance', *Mind* **9**(34), 223–235.

Edwards, R. B. (2010), *The Essentials of Formal Axiology*, Upa.

Efron, B. (1979), 'Bootstrap methods: Another look at the jackknife', *The Annals of Statistics* **7**(1), 1–26.

Efron, B. & Tibshirani, R. J. (1993), *An Introduction to the Bootstrap*, number 57 *in* 'Monographs on Statistics and Applied Probability', Chapman & Hall/CRC, Boca Raton, Florida, USA.

Enders, W. & Granger, C. W. J. (1998), 'Unit-root tests and asymmetric adjustment with an example using the term structure of interest rates', *Journal of Business & Economic Statistics* **16**(3), 304–311.

Engle, R. F. & Granger, C. W. J. (1987), 'Co-integration and error correction: Representation, estimation, and testing', *Econometrica* **55**(2), 251–276.

Fama, E. F. (1970), 'Efficient capital markets: A review of theory and empirical work', *The Journal of Finance* **25**(2), 383–417.

Fama, E. F. & French, K. R. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**(1), 3 – 56.

Fama, E. F. & French, K. R. (2015), 'A five-factor asset pricing model', *Journal of Financial Economics* **116**(1), 1 – 22.

Farrar, D. E. & Glauber, R. R. (1967), 'Multicollinearity in regression analysis: The problem revisited', *The Review of Economics and Statistics* **49**(1), 92–107.

Fernandez, P. et al. (2001), 'Valuation using multiples. how do analysts reach their conclusions?', *IESE Business School* pp. 1–13.

Fisher, I. (1919), 'Stabilizing the dollar', *The American Economic Review* **9**(1), 156–160.

Fisher, I. (1930), 'The theory of interest', *New York* **43**.

Fisher, R. A. & Tippett, L. H. C. (1928), Limiting forms of the frequency distribution of the largest or smallest member of a sample, *in* 'Mathematical Proceedings of the Cambridge Philosophical Society', Vol. 24, Cambridge University Press, pp. 180–190.

Forrester, J. W. (1997), 'Industrial dynamics', *The Journal of the Operational Research Society* **48**(10), 1037–1041.

Frazzini, A. & Lamont, O. A. (2005), Dumb money: Mutual fund flows and the cross-section of stock returns, Working Paper 11526, National Bureau of Economic Research.

Furnival, G. M. & Wilson, R. W. (1974), 'Regressions by leaps and bounds', *Technometrics* **16**(4), 499–511.

Gauß, C. F. (1821), 'Abhandlungen betreffend die wahrscheinlichkeitsrechnung als grundlage des gaußchen fehlerfortpflanzungsgesetzes'.

Gibrat, R. (1931), *Les inégalités économiques: applications: aux inégalitês des richesses, à la concentration des entreprises, aux populations des villes, aux statistiques des familles, etc: d'une loi nouvelle: la loi de l'effet proportionnel*, Librairie du Recueil Sirey.

Gordon, M. (1962), *The Investment, Financing, and Valuation of the Corporation*, Irwin series in economics, R.D. Irwin.

Granger, C. & Newbold, P. (1974), 'Spurious regressions in econometrics', *Journal of Econometrics* **2**(2), 111–120.

Gultekin, B. (1983), 'Stock market returns and inflation: Evidence from other countries', *The Journal of Finance* .

Hamilton, J. D. (1989), 'A new approach to the economic analysis of nonstationary time series and the business cycle', *Econometrica* **57**(2), 357–384.

Hartman, R. S. (1967), *The structure of value; foundations of scientific axiology by Robert S. Hartman. Foreword by George Kimball Plochmann. Pref. by Paul Weiss*, Southern Illinois University Press Carbondale.

Hooke, R. & Jeeves, T. A. (1961), 'Direct search solution of numerical and statistical problems', *J. ACM* **8**(2), 212–229.

Hosking, J. R. (1980), 'The multivariate portmanteau statistic', *Journal of the American Statistical Association* **75**(371), 602–608.

Huerta, G. & West, M. (1999), 'Priors and component structures in autoregressive time series models', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**(4), 881–899.

Hume, D. (1779), *An Enquiry Concerning Human Understanding*, A. Millar.

Jarque, C. M. & Bera, A. K. (1980), 'Efficient tests for normality, homoscedasticity and serial independence of regression residuals', *Economics Letters* **6**(3), 255 – 259.

Jevons, W. S. (1871), 'The principles of political economy'.

Johansen, S. (1991), 'Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models', *Econometrica* **59**(6), 1551–1580.

Joyce, M., Lasaosa, A., Stevens, I., Tong, M. et al. (2011), 'The financial market impact of quantitative easing in the united kingdom', *International Journal of Central Banking* **7**(3), 113–161.

Kane, A., Marcus, A. J. & Noh, J. (1996), 'The p/e multiple and market volatility', *Financial Analysts Journal* **52**(4), 16–24.

Kant, I. (1949), 'Critique of practical reason, and other writings in moral philosophy'.

Kaplan, S. N. & Ruback, R. S. (1995), 'The valuation of cash flow forecasts: An empirical analysis', *The Journal of Finance* **50**(4), 1059–1093.

Kessel, R. A. (1971), The cyclical behavior of the term structure of interest rates, *in* 'Essays on Interest Rates, Volume 2', NBER, pp. 337–390.

Keynes, J. M. (1936), *The General Theory of Employment, Interest, and Money*.

Koenker, R. & Bassett, G. (1978), 'Regression quantiles', *Econometrica* **46**(1), 33–50.

Koenker, R. W. & D'Orey, V. (1987), 'Algorithm as 229: Computing regression quantiles', *Journal of the Royal Statistical Society. Series C (Applied Statistics)* **36**(3), 383–393.

Kohavi, R. et al. (1995), A study of cross-validation and bootstrap for accuracy estimation and model selection, *in* 'Ijcai', Vol. 14, Montreal, Canada, pp. 1137–1145.

Koller, T., Goedhart, M., Wessels, D. et al. (2010), *Valuation: measuring and managing the value of companies*, Vol. 499, John Wiley and sons.

Kreyszig, E. (2010), *Advanced engineering mathematics*, John Wiley & Sons.

Kwiatkowski, D., Phillips, P. C., Schmidt, P. & Shin, Y. (1992), 'Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?', *Journal of Econometrics* **54**(1), 159 – 178.

Kwon, C. S. & Shin, T. S. (1999), 'Cointegration and causality between macroeconomic variables and stock market returns', *Global Finance Journal* **10**(1), 71 – 81.

L Little, P. & L Little, B. (2000), 'Do perceptions of corporate social responsibility contribute to explaining differences in corporate price-earnings ratios? a research note', **3**, 137–142.

Landreth, H. & Colander, D. C. (2002), *History of economic thought*, Houghton Mifflin College Division.

Lee, B.-S. (1992), 'Causal relations among stock returns, interest rates, real activity, and inflation', *Journal of Finance* pp. 1591–1603.

Li, H. & Maddala, G. (1997), ‘Bootstrapping cointegrating regressions’, *Journal of Econometrics* **80**(2), 297 – 318.

Lintner, J. (1965), ‘The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets’, *The Review of Economics and Statistics* **47**(1), 13–37.

Lundgren, J., Rönnqvist, M. & Värbrand, P. (2010), *Optimization*, Professional Publishing.

Maio, P. & Philip, D. (2015), ‘Macro variables and the components of stock returns.’, *Journal of empirical finance.* **33**, 287–308.

Mandelbrot, B. B. & Van Ness, J. W. (1968), ‘Fractional brownian motions, fractional noises and applications’, *SIAM review* **10**(4), 422–437.

Mankiw, N. G. (2014), *Principles of macroeconomics*, Cengage Learning.

Markowitz, H. (1952), ‘Portfolio selection’, *The Journal of Finance* **7**(1), 77–91.

Marx, K. (1867), ‘Das kapital: kritik der politischen ökonomie’, *Germany: Verlag von Otto Meisner* **1885**, 1894.

Massey Jr, F. J. (1951), ‘The kolmogorov-smirnov test for goodness of fit’, *Journal of the American statistical Association* **46**(253), 68–78.

Matyas, L. & Sevestre, P. (1996), ‘The econometrics of panel data’, *A Handbook of the Theory with* .

Meaning, J. & Zhu, F. (2011), ‘The impact of recent central bank asset purchase programmes’.

Menger, C. (1871), *Grundsätze der volkswirthschaftslehre*, Vol. 1, W. Braumüller.

Miller, M. H. & Modigliani, F. (1961), ‘Dividend policy, growth, and the valuation of shares’, *The Journal of Business* **34**(4), 411–433.

Mitchell, T. M. (1997), *Machine Learning*, 1 edn, McGraw-Hill, Inc., New York, NY, USA.

Modigliani, F. & Cohn, R. A. (1979), ‘Inflation, rational valuation and the market’, *Financial Analysts Journal* **35**(2), 24–44.

Nelder, J. A. & Mead, R. (1965), ‘A simplex method for function minimization’, *Computer Journal* **7**, 308–313.

Nelson, C. R. & Plosser, C. R. (1982), ‘Trends and random walks in macroeconmic time series: some evidence and implications’, *Journal of monetary economics* **10**(2), 139–162.

Nicholson, S. F. (1960), ‘Price-earnings ratios’, *Financial Analysts Journal* **16**(4), 43–45.

Nietzsche, F. W. .-. (1910), *The will to power : an attempted transvaluation of all values*, Gordon Press, New York.

Panaretos, V. M. & Konis, K. (2012), ‘Nonparametric construction of multivariate kernels’, *Journal of the American Statistical Association* **107**(499), 1085–1095.

Pang, B., Lee, L. et al. (2008), ‘Opinion mining and sentiment analysis’, *Foundations and Trends® in Information Retrieval* **2**(1–2), 1–135.

Pearson, K. (1901), 'Liii. on lines and planes of closest fit to systems of points in space', *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **2**(11), 559–572.

Perron, P. (1989), 'The great crash, the oil price shock, and the unit root hypothesis', *Econometrica* **57**(6), 1361–1401.

Piketty, T. (2015), 'About capital in the twenty-first century', *American Economic Review* **105**(5), 48–53.

Plato (2008), *The Republic*, Oxford University Press.

Popper, K. & Bartley, W. (1982), *The Open Universe: An Argument for Indeterminism*, Karl R. Popper, Hutchinson.

Ramcharran, H. (2002), 'An empirical analysis of the determinants of the p/e ratio in emerging markets', *Emerging Markets Review* **3**(2), 165 – 178.

Reilly, R. R. & Wecker, W. E. (1973), 'On the weighted average cost of capital', *The Journal of Financial and Quantitative Analysis* **8**(1), 123–126.

Ritter, J. R. (2005), 'Economic growth and equity returns', *Pacific-Basin Finance Journal* **13**(5), 489–503.

Ritter, J. R. & Warr, R. S. (2002), 'The decline of inflation and the bull market of 1982-1999', *Journal of Financial and Quantitative Analysis* **37**(1), 29–61.

Roman, S. (2007), *Advanced Linear Algebra*, Graduate Texts in Mathematics, Springer New York.

Rostow, W. W. (1959), 'The stages of economic growth', *The economic history review.* **12**(1), 1–16.

Rubin, I. I. (1978), 'Abstract labour and value in marx's system', *Capital & Class* **2**(2), 107–109.

Russell, S. J. & Norvig, P. (2003), *Artificial Intelligence: A Modern Approach*, 2 edn, Pearson Education.

Sakia, R. (1992), 'The box-cox transformation technique: a review', *The statistician* pp. 169–178.

Schreiner, A., Spremann, K., Aboody, D., Berndt, T., Bernhard, J., Gantenbein, P., Lang, S., Liu, J., Thomas, J. & , S. (2007), 'Multiples and their valuation accuracy in european equity markets'.

Schrödinger, E. (1926), 'An undulatory theory of the mechanics of atoms and molecules', *Physical review* **28**(6), 1049.

Scott, M. & Bruce, R. (1987), 'Five stages of growth in small business', *Long range planning : the journal of the Society for Long Range Planning and of the European Planning Federation* **20**(3), 45–52.

Shamsuddin, A. F. & Hillier, J. R. (2004), 'Fundamental determinants of the australian price-earnings multiple', *Pacific-Basin Finance Journal* **12**(5), 565 – 576.

Sharpe, W. F. (1964), 'Capital asset prices: A theory of market equilibrium under conditions of risk', *The Journal of Finance* **19**(3), 425–442.

Shiller, R. (2000), *Irrational Exuberance*, New York Times bestseller, Princeton University Press.

Smith, A. (1776), *An Inquiry to the Nature and Causes of the Wealth of Nations*, W. Strahan and T. Cadell, London, UK.

Stigler, S. M. (1973), 'The asymptotic distribution of the trimmed mean', *The Annals of Statistics* **1**(3), 472–477.

Theodossiou, P. (1998), 'Financial data and the skewed generalized t distribution', *Management science* **44**(12-part-1), 1650–1661.

Thorbecke, W. (1997), 'On stock market returns and monetary policy', *The Journal of Finance* **52**(2), 635–654.

Toda, H. Y. & Yamamoto, T. (1995), 'Statistical inference in vector autoregressions with possibly integrated processes', *Journal of Econometrics* **66**(1-2), 225–250.

Townsend, J. S. (2000), *A modern approach to quantum mechanics*, University Science Books.

Tsay, R. S. (2010), *Financial Time Series*, John Wiley & Sons, Ltd.

Uhlenbeck, G. E. & Ornstein, L. S. (1930), 'On the theory of the brownian motion', *Phys. Rev.* **36**, 823–841.

van Vught, F. (1987), 'Pitfalls of forecasting: Fundamental problems for the methodology of forecasting from the philosophy of science', *Futures* **19**(2), 184 – 196.

Ville, J. (1939), *Etude critique de la notion de collectif.*

Walpole, R. E., Myers, R. H., Myers, S. L. & Ye, K. (1993), *Probability and statistics for engineers and scientists*, Vol. 5, Macmillan New York.

Walpole, R., Myers, R., Myers, S. & Ye, K. (2012), 'Probability & statistics for engineers & scientists, d. lynch, ed., boston, massachusetts: Pearson education'.

Walras, L. (1874), *Éléments d'économie politique pure, ou, Théorie de la richesse sociale*, F. Rouge.

White, C. B. (2000), 'What p/e will the u.s. stock market support?', *Financial Analysts Journal* **56**(6), 30–38.

Wilkinson, J. H. (1965), *The algebraic eigenvalue problem*, Vol. 87, Clarendon Press Oxford.

Williams, J. B. (1938), *The theory of investment value*, Vol. 36, Harvard university press Cambridge, MA.

Yan, G. C. C. (2011), 'A financial engineering approach to identify stock market bubble', *Systems Engineering Procedia* **2**, 153 – 162. Complexity System and Engineering Management.

Yeo, I. & Johnson, R. A. (2000), 'A new family of power transformations to improve normality or symmetry', *Biometrika* **87**(4), 954–959.

Zorn, T., Dudney, D. & Jirasakuldech, B. (2008), 'P/e changes: some new results', *Journal of Forecasting* **28**(4), 358–370.

Appendices

A Appendix - Chapter 3

A.I Overview of Initially Selected Macroeconomic Variables

Category	Description	Unit	Source	Assigned name
Interest Rates	Long term government bond yield (average), NSA - Euro Zone	Percent	Eurostat	iEULT
	Euro Benchmark Bond - 10 Year - Yield	Yield	Tullet Prebon Information	iEU10
	Euro Benchmark Bond - 5 Year - Yield	Yield	Tullet Prebon Information	iEU5
	Euro Benchmark Bond - 1 Year - Yield	Yield	Tullet Prebon Information	iEU1
	3-Month Euribor (% per annum, last) - Euro Zone	Percent	ECB - European Central Bank	iEURIBOR3MTH
	EURIBOR (ACT/360) - 3 Month - Yield	Yield	European Money Markets Institute	iEURIBOR3MTHY
	EURIBOR (ACT/360) - 1 Month - Yield	Yield	European Money Markets Institute	
	10 Year, Government Bond Yield, Real Return, Percent - Europe	Percent	ECB - European Central Bank	iEU10R
	Government Bond, 10 Year - Euro Zone	Percent	ECB - European Central Bank	
	10 Years, ECB, Spot Rate, AAA-Rated Government Bond Yield, Percent - Euro Zone	Percent	ECB - European Central Bank	
	MFI interest rates on euro-denominated deposits, Repos, New business (% per annum) - Euro Zone	Percent	ECB - European Central Bank	
	MFI interest rates on euro-denominated deposits, Repos, Outstanding amounts (% per annum) - Euro Zone	Percent	ECB - European Central Bank	
	Up to 1 Year, Lending Rates, Households, New Business, For House Purchase, AAR/NDER, Percent - Euro Zone	Percent	ECB - European Central Bank	
Exchange Rates (Spot)	ECB Reference Exchange Rate, UK Pound Sterling/Euro, 2:15 pm (C.E.T.), Against ECU up to December 1998 - United Kingdom	GBP/EUR	ECB - European Central Bank	fxGBPEUR
	ECB Reference Exchange Rate, US Dollar/Euro, 2:15 pm (C.E.T.), Against ECU up to December 1998 - United States	USD/EUR	ECB - European Central Bank	fxUSDEUR
	ECB Reference Exchange Rate, Japanese Yen/Euro, 2:15 pm (C.E.T.), Against ECU up to December 1998 - Japan	JPY/EUR	ECB - European Central Bank	
	ECB Reference Exchange Rate, Chinese Yuan Renminbi/Euro, 2:15 pm (C.E.T.), Against ECU up to December 1998 - China	CNY/EUR	ECB - European Central Bank	
Economic (EU)	MFI Loans, Households, Consumer Credit, Outstanding Amount (EA Changing Composition), SA, Bil EUR - Euro Zone	Billions of EUR	ECB - European Central Bank	
	Citi Inflation Surprise Index, Euro-Zone - Euro Zone	Index	Citigroup	
	Aggregated Balance Sheet of Eurosystem, Total Assets/Liabilities, Levels (EA Changing Composition), EOP, NSA, Bil EUR - Euro Zone	Billions of EUR	ECB - European Central Bank	mEUA
	Aggregated Balance Sheet of Eurosystem, Assets, Levels, External Assets (EA Changing Composition), EOP, NSA, Bil EUR - Euro Zone	Billions of EUR	ECB - European Central Bank	mEUEXTA
	Aggregated Balance Sheet of Eurosystem, Liabilities, Levels, External Liabilities (EA Changing Composition), EOP, NSA, Bil EUR - Euro Zone	Billions of EUR	ECB - European Central Bank	mEUEXTL
	Aggregated Balance Sheet of Eurosystem, Assets, Levels, Fixed Assets (EA Changing Composition), EOP, NSA, Bil EUR - Euro Zone	Billions of EUR	ECB - European Central Bank	mEUFA
	Standardized Unemployment Rate (Euro Area 17 fixed, SA, percent of civilian labor force) - Euro Zone	Percent	ECB - European Central Bank	mEUSUR
	Unemployment Rate, Euro Area 19, SA, Percent - Euro Zone	Percent	Eurostat	mEUUR
	Harmonized Consumer Prices (HICP), Overall, (Annual rate of change, NSA) - Euro Zone	Percent	ECB - European Central Bank	mEUCPI
	PPI, Industry, Euro 19, 2015=100, NSA, Index - Euro Zone	Index	Eurostat	mEUPPI
	Retail Sales, Turnover and Volume of Sales, Except of Motor Vehicles and Motorcycles, Euro Area (19 Countries), Y/Y % Change, CA, Percent - Euro Zone	Percent	Eurostat	
	Retail Sales, Turnover and Volume of Sales, Except of Motor Vehicles and Motorcycles, Euro Area 19 Countries, M/M% Change, SCA, Percent - Euro Zone	Percent	Eurostat	
	Consumer Survey, Consumer confidence indicator, Balance, Nace R2, SA, Index - Euro Zone	Index	Eurostat	sEUCC
	Business Survey, Industrial confidence indicator, Balance, Nace R2, SA, EA19 - Euro Zone	Index	Eurostat	sEUIC
	Business Survey, Economic Sentiment Indicator, Euro Area (19 Countries), SA, Index - Euro Zone	Index	Eurostat	sEUES
	Business Survey, ZEW, Economic Expectations, Balance, Percent - Euro Zone	Percent	ZEW - Center for European Economic Research	sEUEE
	Business Survey, ZEW, Short-Term Interest Rates, Balance, Percent - Euro Zone	Percent	ZEW - Center for European Economic Research	sEUSTI
	Business Survey, ZEW, Inflation Rate, Balance, Percent - Euro Zone	Percent	ZEW - Center for European Economic Research	sEUINFL
	Business Survey, Industrial Confidence Indicator, SA, Index - Euro Zone	Index	Eurostat	
	PMI Composite Sector, Output Index, SA - Euro Zone	Index	Markit Economics (NTC)	
	PMI Construction Sector, Total Activity Index, SA - Euro Zone	Index	Markit Economics (NTC)	
	PMI Manufacturing Sector, PMI Index, SA - Euro Zone	Index	Markit Economics (NTC)	
	PMI Services Sector, Business Activity Index, SA - Euro Zone	Index	Markit Economics (NTC)	
	Retail PMI, SA - Euro Zone	Index	Markit Economics (NTC)	
Economic (US)	Leading Index, 10 Year, Interest Rate Spread, Treasury Bonds Less Federal Funds, Percent - United States	Percent	Conference Board	mUSi10
	US Future Inflation Gauge, Index - United States	Index	ECRI - Economic Cycle Research Institute	
	Leading Index, Leading Credit Index - United States	Percent	Conference Board	mUSCRED
	H.8, Assets of Commercial Banks, Total assets, SA, Bil USD - United States	Billions of USD	Federal Reserve System	mUSCBA
	H.8, Assets of Large Domestically Chartered Commercial Banks, Total assets, SA, Bil USD - United States	Billions of USD	Federal Reserve System	mUSDCBA
	Monthly Treasury Statement, Federal Debt, Amount Outstanding, Total, NSA, Mil USD - United States	Millions of USD	U.S. Department of Treasury	mUSFD
	Civilian Unemployment Rate, SA, Percent - United States	Percent	Conference Board	mUSUR
	Employment, Population Ratio, SA, Percent - United States	Percent	U.S. Department of Labor	mUSER
	Retail and Food Service Sales, NAICS, (Incl. in Advance Release), SA/WDA, Mil USD - United States	Millions of USD	U.S. Census Bureau	mUSRET
	Retail Trade, Chained 2009, SA, Mil USD - United States	Millions of USD	BEA - US Bureau of Economic Analysis	mUSRETCH
	Chicago PMI, Business Barometer Index, SA, Index - United States	Index	MNI/Deutsche Borse Group	mUSPMI
	PMI Manufacturing Sector, PMI Index (Markit), SA - United States	Index	Markit Economics (NTC)	
	PMI Manufacturing Sector, PMI Index (Markit), NSA - United States	Index	Markit Economics (NTC)	
	PMI Services Sector-Business Activity Index, NSA - United States	Index	Markit Economics (NTC)	
	PMI Composite Sector-Output Index, SA, USD - United States	Index	Markit Economics (NTC)	
	PMI Services Sector-Business Activity Index, SA - United States	Index	Markit Economics (NTC)	
Industrial Production	Euro Zone - Industrial Production Mth End		FactSet Standardized Economics	mEUp
	United States - Industrial Production Mth End		FactSet Standardized Economics	mUSp
	China - Industrial Production Mth End		FactSet Standardized Economics	
Volatility	CBOE Market Volatility Index - Index Price Level	Index	FactSet Prices	vVIX
	Euro STOXX 50 Volatility - Index Price Level	Index	STOXX	vST50
EPUI	Economic Policy Uncertainty, News Based, Index - United States	Index	Economic Policy Uncertainty	sUSEPU
	Economic Policy Uncertainty, News Based, Index - China	Index	Economic Policy Uncertainty	sCHEPUI
	Economic Policy Uncertainty, News-Based, Index - Europe	Index	Economic Policy Uncertainty	sEUEPUI

Table 1: Complete list of 65 initial macroeconomic variables

A.II Correlation Matrices for Independent Variables

Correlations based on monthly values from Januar 2000 to January 2018.

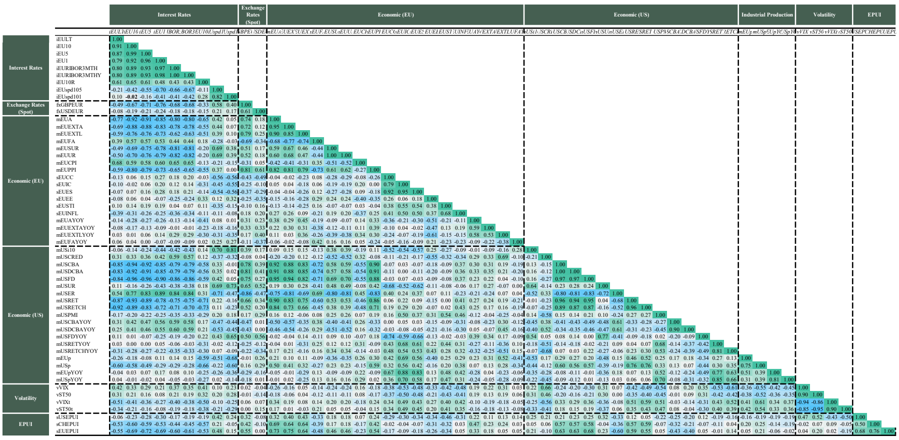


Figure 1: Correlation matrix for absolute variables

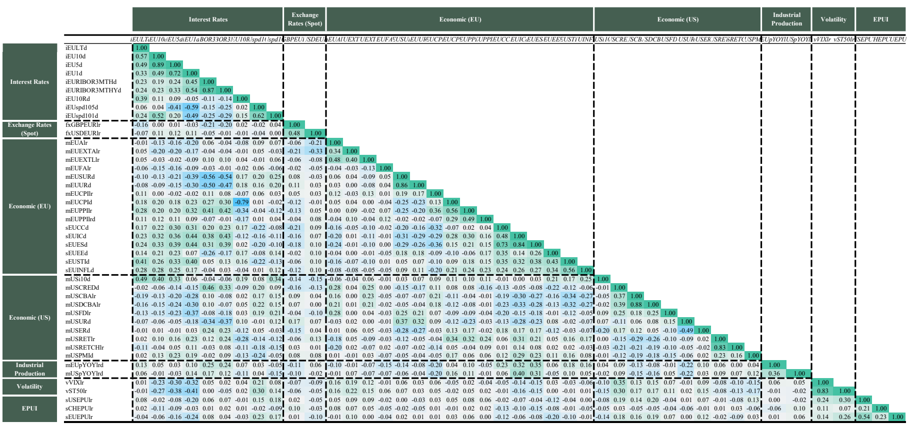


Figure 2: Correlation matrix for differentiated variables

		Interest Rates		FX	Economic (EU)									Economic (US)		Ind. Prod.	Volatility		EPUI
		iEU10	iEUspd101	fxUSDEUR	mEUFA	mEUCPI	sEUCC	sEUEE	sEUSTI	sEUINFL	mEUAYOY	mEUEXTAYOY	mEUFAYOY	mUSPMI	mUSCBAYOY	mEUpYOY	vST50	vST50t	sEUEPUI
Interest Rates	iEU10	1.00																	
	iEUspd101	-0.02	1.00																
Exchange Rates	fxUSDEUR	-0.19	0.17	1.00															
Economic (EU)	mEUFA	0.57	-0.03	-0.34	1.00														
	mEUCPI	0.59	-0.15	0.05	0.35	1.00													
	sEUCC	0.06	-0.56	-0.49	0.08	-0.08	1.00												
	sEUEE	0.06	0.32	-0.35	0.29	-0.40	0.26	1.00											
	sEUSTI	0.14	-0.15	0.16	0.16	0.03	0.38	0.38	1.00										
	sEUINFL	-0.31	-0.08	0.20	-0.21	-0.37	0.41	0.37	0.68	1.00									
	mEUAYOY	-0.28	0.01	0.23	-0.19	0.14	-0.36	-0.51	-0.21	-0.11	1.00								
	mEUEXTAYOY	-0.17	-0.16	0.33	-0.38	0.11	-0.10	-0.47	0.13	0.19	0.59	1.00							
	mEUFAYOY	0.04	0.27	-0.37	0.42	0.05	-0.05	0.21	-0.23	-0.23	-0.09	-0.22	1.00						
Economic (US)	mUSPMI	-0.20	0.18	0.29	0.08	0.07	0.16	0.31	0.54	0.46	-0.12	-0.04	-0.04	1.00					
	mUSCBAYOY	0.42	-0.44	0.01	0.38	0.26	0.00	-0.15	-0.09	-0.31	-0.08	-0.23	-0.12	-0.27	1.00				
Industrial	mEUpYOY	0.03	-0.36	-0.16	0.13	0.22	0.67	0.13	0.48	0.42	-0.28	0.04	-0.06	0.52	-0.12	1.00			
Volatility	vST50	0.21	0.28	-0.14	-0.12	0.08	-0.37	-0.21	-0.43	-0.41	0.19	0.13	0.06	-0.45	-0.01	-0.36	1.00		
	vST50t	-0.21	-0.29	0.15	0.21	-0.04	0.34	0.20	0.41	0.35	-0.16	-0.18	-0.08	0.47	0.08	0.33	-0.95	1.00	
EPUI	sEUEPUI	-0.69	0.15	0.00	-0.48	-0.23	-0.17	-0.26	-0.34	0.05	0.33	0.18	0.05	0.05	-0.43	-0.06	0.20	-0.19	1.00

Figure 3: Final correlation matrix for absolute variables

		Interest Rates		FX	Economic (EU)									
		iEU10d	iEUspd105d	fxUSDEURlr	mEUAlr	mEUEXTAlr	mEUEXTLlr	mEUFAlr	mEUURd	mEUCPId	mEUPPIlrd	sEUCCd	sEUEEd	sEUSTId
Interest Rates	iEU10d	1.00												
	iEUspd105d	0.04	1.00											
FX	fxUSDEURlr	0.11	-0.04	1.00										
Economic (EU)	mEUAlr	-0.13	0.09	-0.21	1.00									
	mEUEXTAlr	-0.20	0.05	-0.33	0.34	1.00								
	mEUEXTLlr	-0.03	-0.01	-0.08	0.48	0.40	1.00							
	mEUFAlr	-0.15	0.06	-0.05	-0.04	-0.03	-0.13	1.00						
	mEUURd	-0.09	0.16	0.03	0.03	0.00	-0.08	0.04	1.00					
	mEUCPId	0.20	0.01	-0.01	0.05	0.04	0.00	-0.04	-0.23	1.00				
	mEUPPIlrd	0.12	0.01	0.08	-0.04	0.10	-0.04	0.12	-0.02	0.29	1.00			
	sEUCCd	0.22	-0.22	0.09	-0.16	-0.05	-0.10	-0.02	-0.16	-0.07	0.04	1.00		
	sEUEEd	0.21	-0.08	0.10	-0.04	0.00	-0.01	-0.05	0.18	-0.10	0.17	0.35	1.00	
	sEUSTId	0.26	-0.22	0.10	-0.16	-0.07	-0.10	0.01	0.07	0.09	0.15	0.35	0.43	1.00
Economic (US)	mUSi10d	0.40	0.08	-0.15	-0.06	-0.04	0.06	-0.01	0.07	0.11	0.11	-0.01	0.21	0.17
	mUSCREDd	-0.06	0.20	-0.13	0.28	0.04	0.25	0.00	-0.17	0.08	-0.16	-0.13	-0.22	-0.12
	mUSCBAlr	-0.13	0.17	0.04	0.16	0.00	0.23	-0.05	-0.07	-0.11	-0.01	-0.19	-0.16	-0.34
	mUSFDlr	-0.15	0.19	-0.10	0.28	0.00	0.04	-0.03	0.21	-0.09	-0.04	-0.20	-0.01	-0.12
	mUSERd	0.01	0.05	0.04	0.01	0.06	0.05	-0.03	-0.27	0.13	-0.02	0.18	-0.12	-0.03
	mUSRETCHlr	-0.04	-0.18	0.01	-0.20	0.02	-0.07	0.02	-0.02	0.05	0.09	0.01	0.02	0.02
	mUSPMId	0.13	-0.24	0.08	0.01	-0.01	0.03	-0.07	-0.04	0.17	0.06	0.12	0.11	0.16
Industrial Production	mEUpYOYlrd	0.05	0.03	0.06	-0.10	-0.01	-0.07	-0.15	-0.08	0.04	-0.05	0.23	0.06	0.18
	mUSpYOYlrd	-0.01	0.04	-0.02	-0.01	0.07	-0.07	-0.07	-0.04	0.16	-0.01	0.06	-0.05	0.10
Volatility	vST50lr	-0.27	0.30	-0.05	0.16	0.22	0.15	0.06	0.03	-0.02	0.02	-0.01	0.00	-0.01
EPUI	sEUEPUlr	-0.06	0.23	-0.10	-0.01	0.10	0.00	-0.04	0.01	0.03	0.00	-0.12	-0.20	-0.10

		Economic (US)							Ind. Prod.		Vol.	EPUI
		mUSi10d	mUSCREDd	mUSCBAlr	mUSFDlr	mUSERd	mUSRETCHlr	mUSPMId	mEUpYOYlrd	mUSpYOYlrd	vST50lr	sEUEPUlr
Economic (US)	mUSi10d	1.00										
	mUSCREDd	-0.01	1.00									
	mUSCBAlr	-0.05	0.37	1.00								
	mUSFDlr	0.09	0.25	0.18	1.00							
	mUSERd	-0.20	0.17	0.12	-0.10	1.00						
	mUSRETCHlr	-0.03	-0.21	-0.21	-0.10	-0.02	1.00					
	mUSPMId	-0.01	-0.12	-0.19	-0.15	0.02	0.16	1.00				
Industrial Production	mEUpYOYlrd	0.04	0.09	-0.13	-0.01	0.10	0.00	0.04	1.00			
	mUSpYOYlrd	-0.02	0.09	-0.15	0.05	0.03	0.07	0.12	0.36	1.00		
Volatility	vST50lr	-0.15	0.30	0.17	0.11	0.15	-0.13	-0.17	-0.01	-0.02	1.00	
EPUI	sEUEPUlr	-0.14	0.18	0.16	0.07	0.12	-0.09	0.03	0.01	0.06	0.26	1.00

Figure 4: Final correlation matrix for differentiated variables

A.III Selection of Variables based on Subset Regressions

For *Ap*, the R^2 gain from four independent variables to five independents is diminishing. All coefficients are significant with $p < 0.001$. The variables included in the optimal models are fairly consistent, from inflation, production growth and volatility for $n = 3$, to inflation, production growth, volatility and spread for $n = 4$ and inflation, production growth, volatility, spread and long interest rates for $n = 5$. $n = 4$ does not yield optimal RSS and BIC, as seen in Figure 11 and 12 respectively, but this is evaluated against the cost of adding more variables for prediction.

For *Ad*, *Dd*, *Dp*, *Rp* and *Rd* the selection is done based on similar rationale, and the algorithm outputs are displayed in Figure 6 to 10.

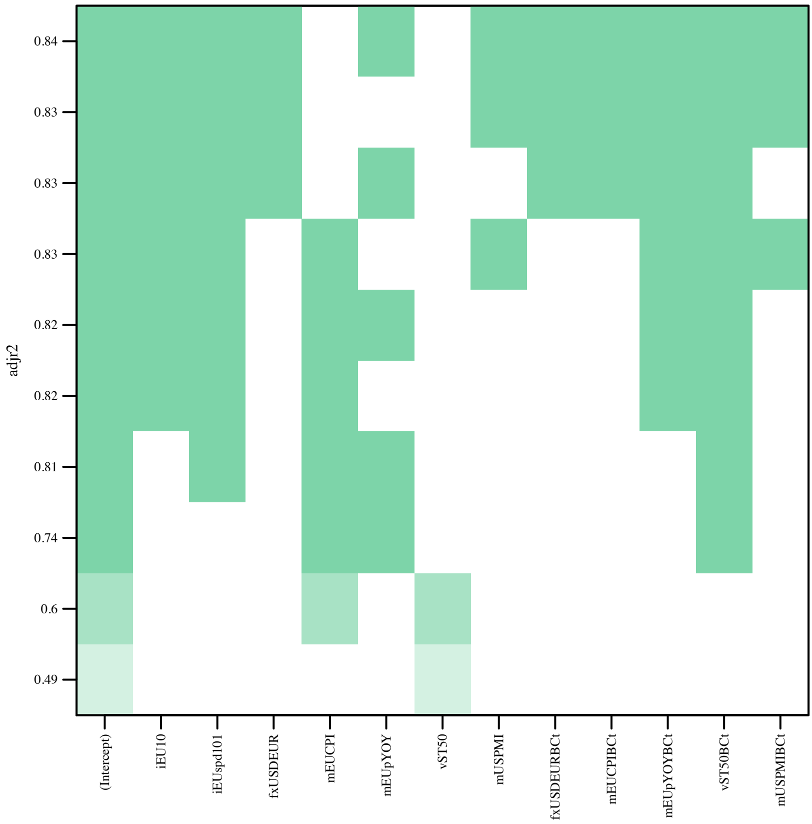


Figure 5: Adj. R^2 for *Ap* depending on chosen independent variables

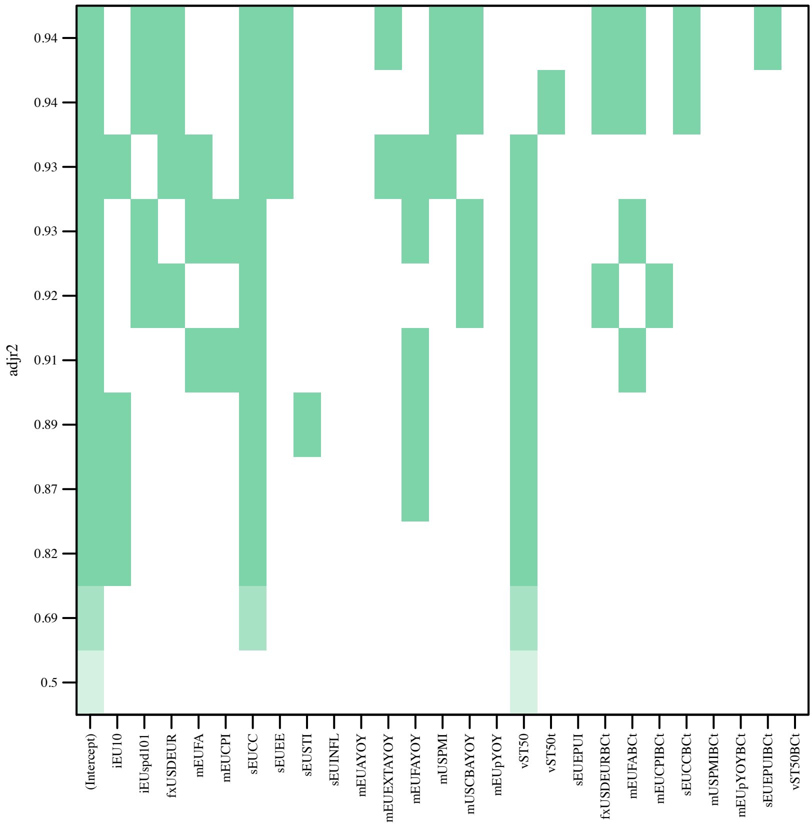


Figure 6: Adj. R^2 for *Ad* depending on chosen independent variables

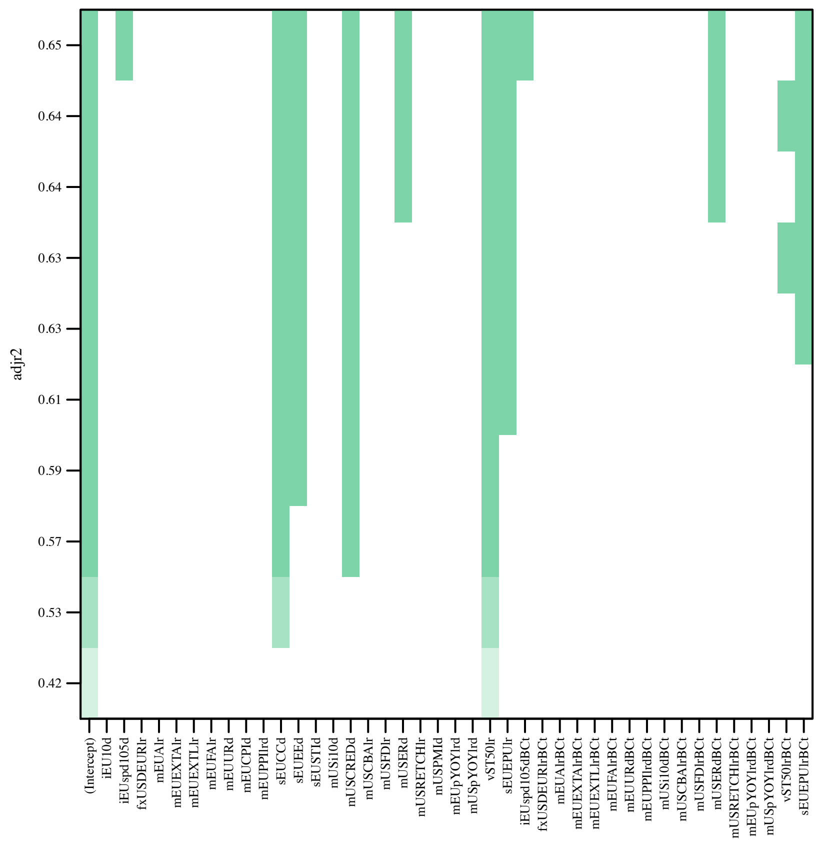


Figure 7: Adj. R^2 for *Dd* depending on chosen independent variables

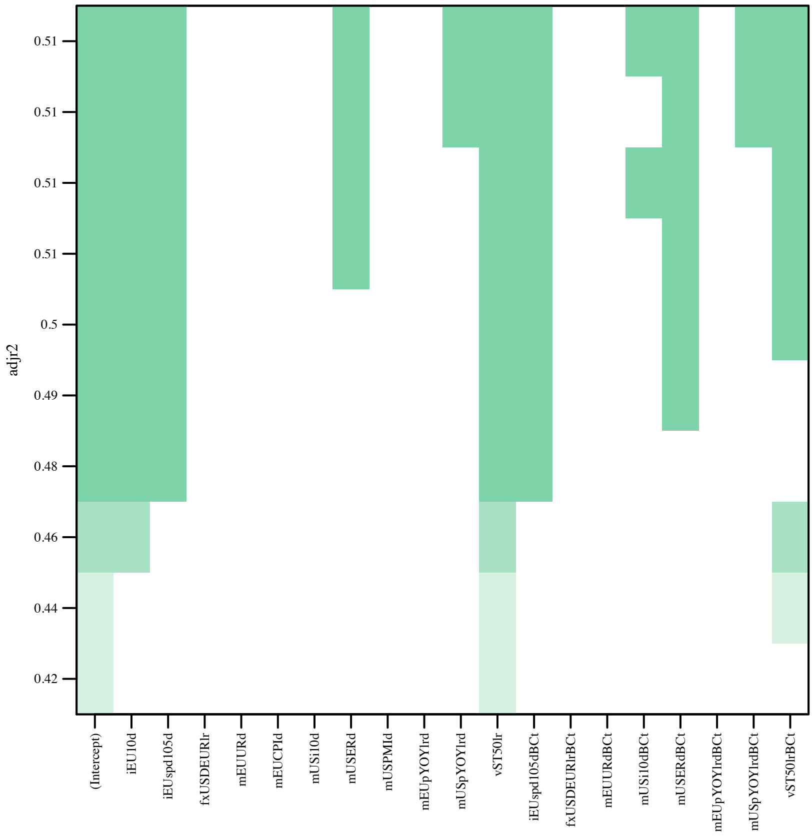


Figure 8: Adj. R^2 for *Dp* depending on chosen independent variables

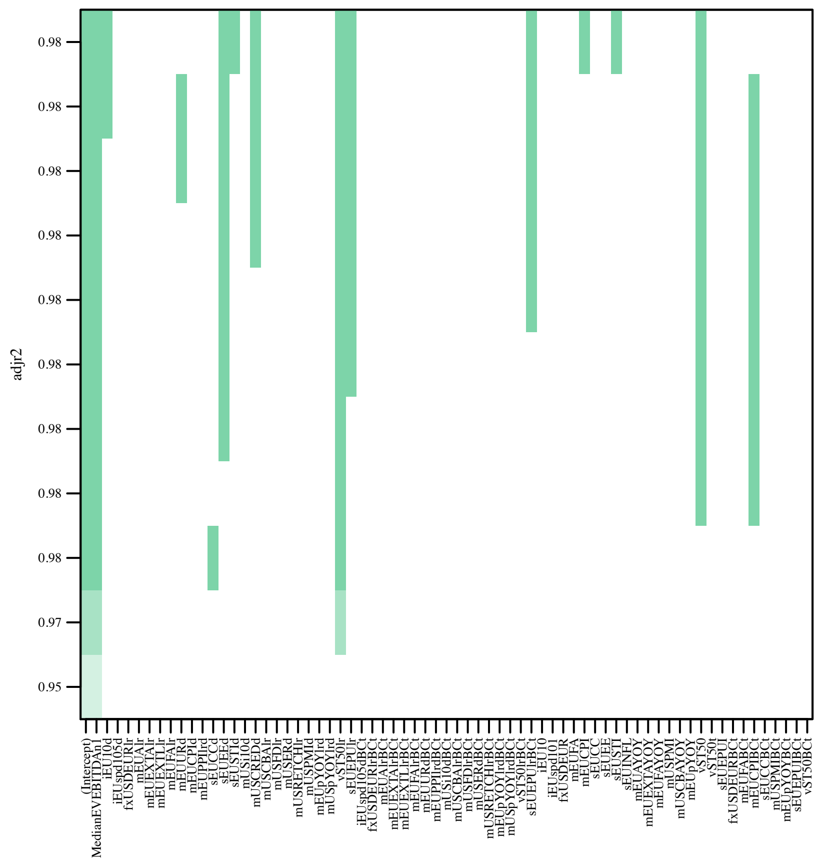


Figure 9: Adj. R^2 for *Rd* depending on chosen independent variables

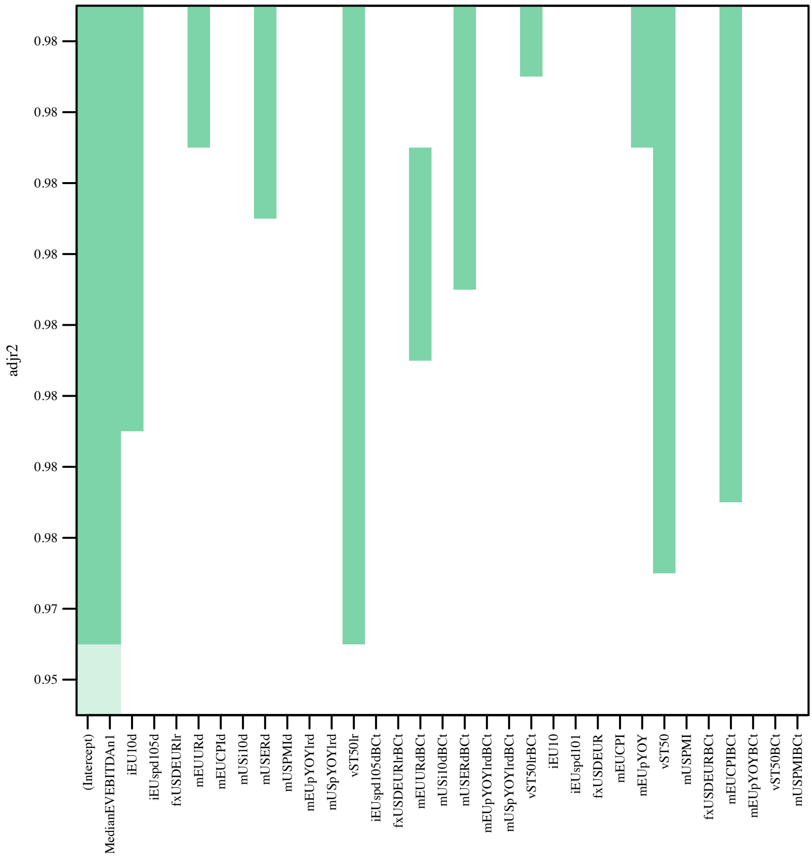


Figure 10: Adj. R^2 for Rp depending on chosen independent variables

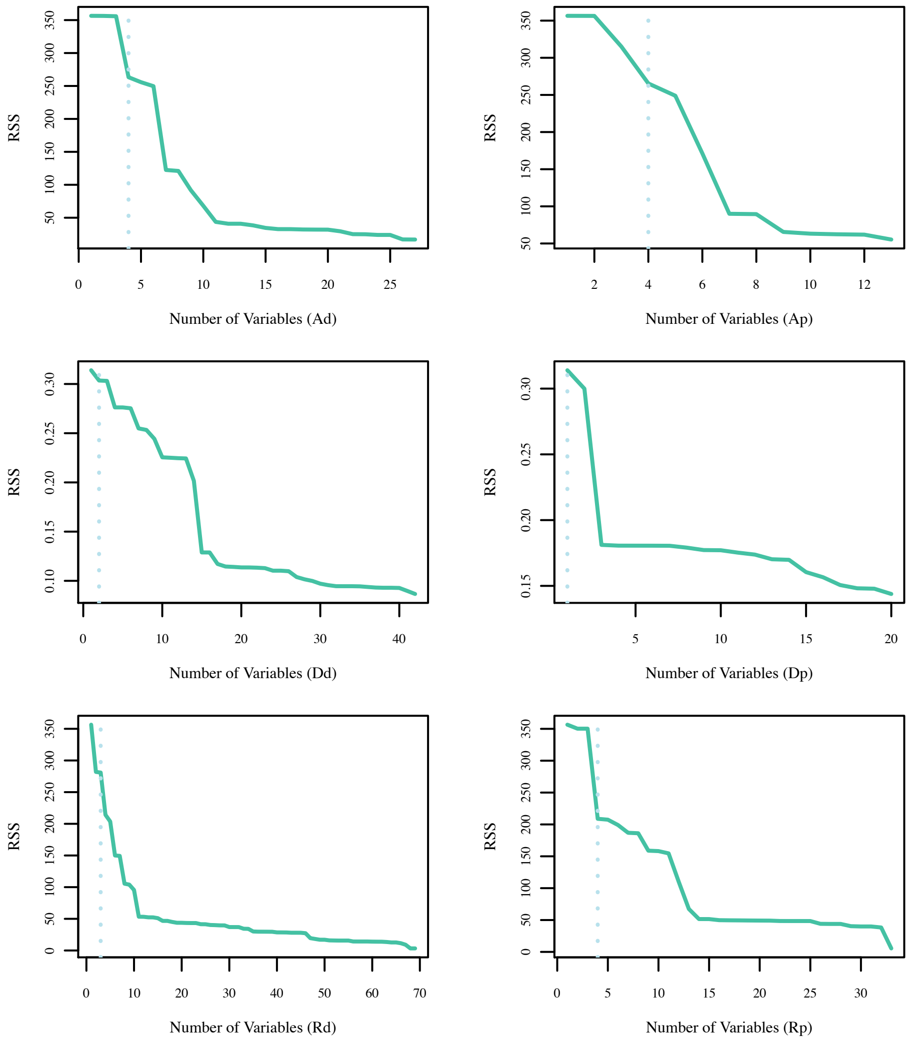


Figure 11: RSS against number of variables in the six models from the subset regressions

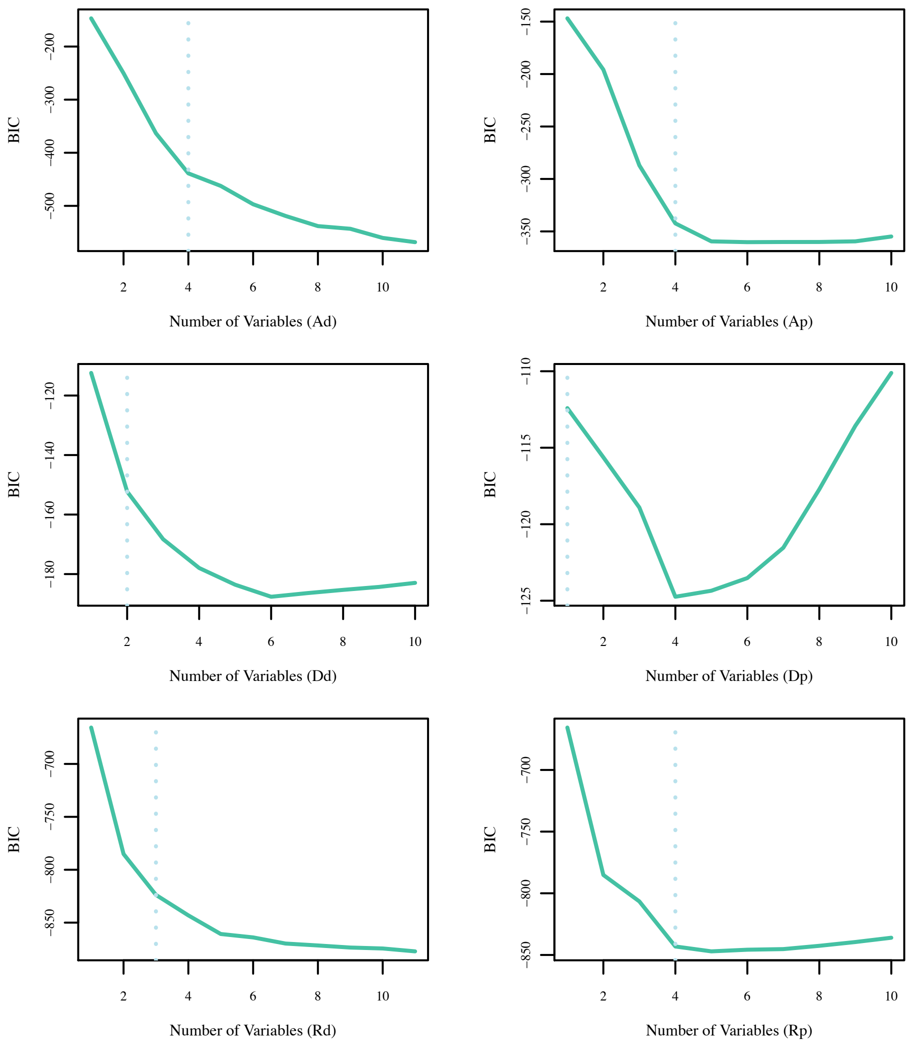


Figure 12: BIC against number of variables in the six models from the subset regressions

B Appendix - Chapter 4

B.I Out-of-Sample Testing for Models

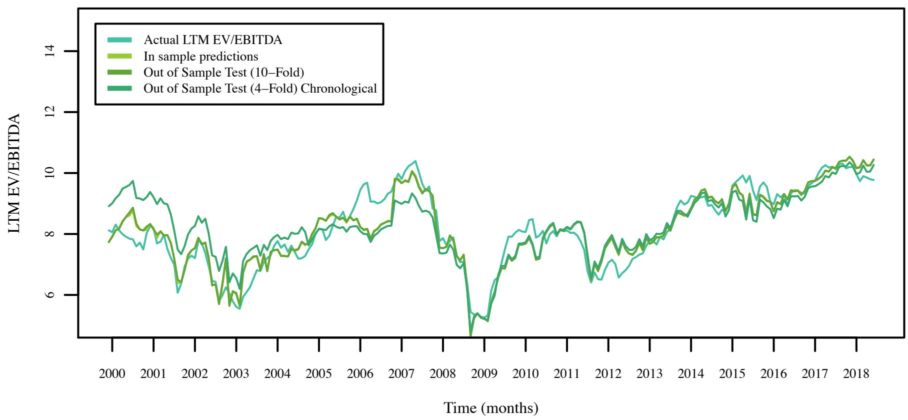


Figure 13: In- and out-of-sample predictions for *Ad*

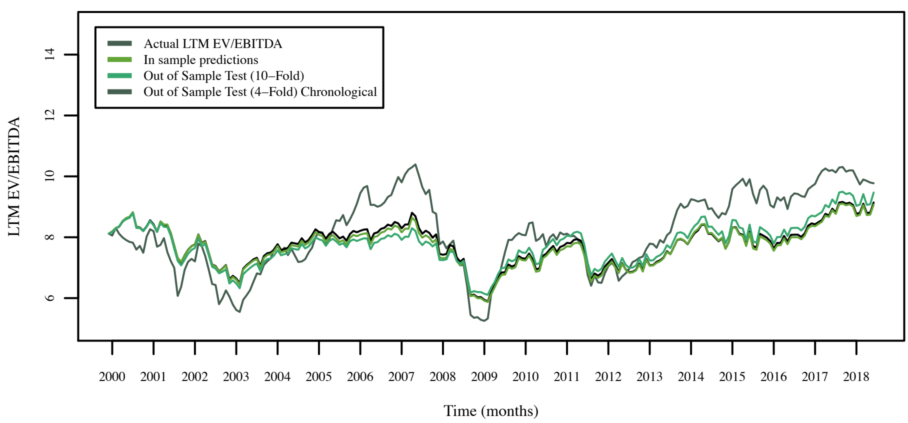


Figure 14: In- and out-of-sample predictions for *Dd*

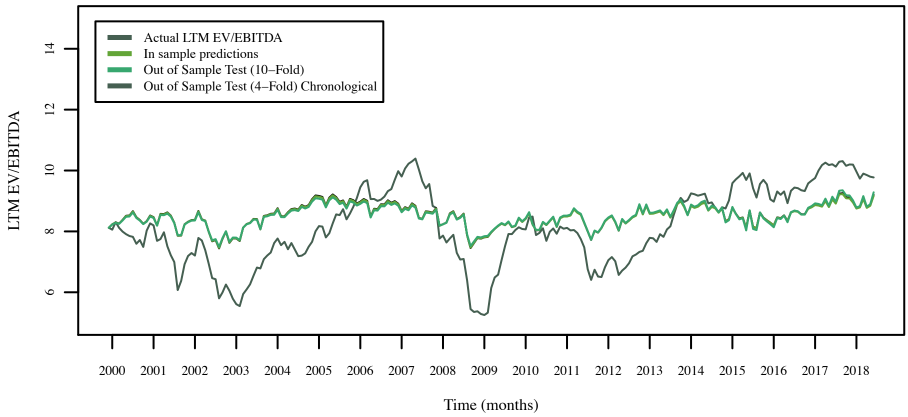


Figure 15: In- and out-of-sample predictions for *Dp*

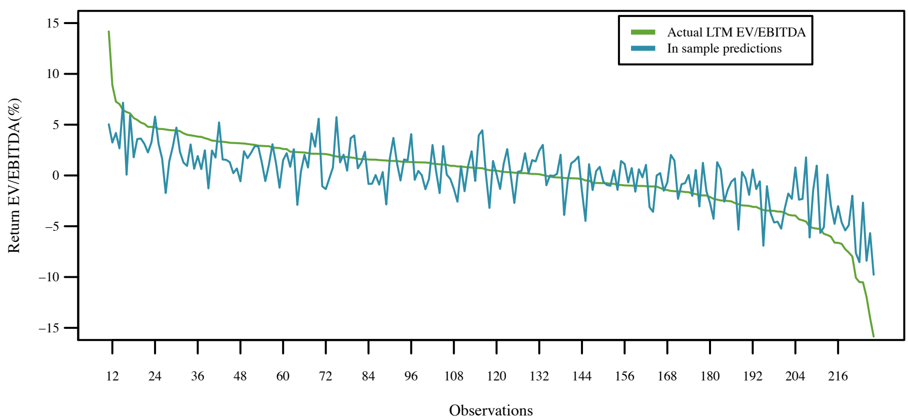


Figure 16: Sorted in-sample predictions for *Dd*

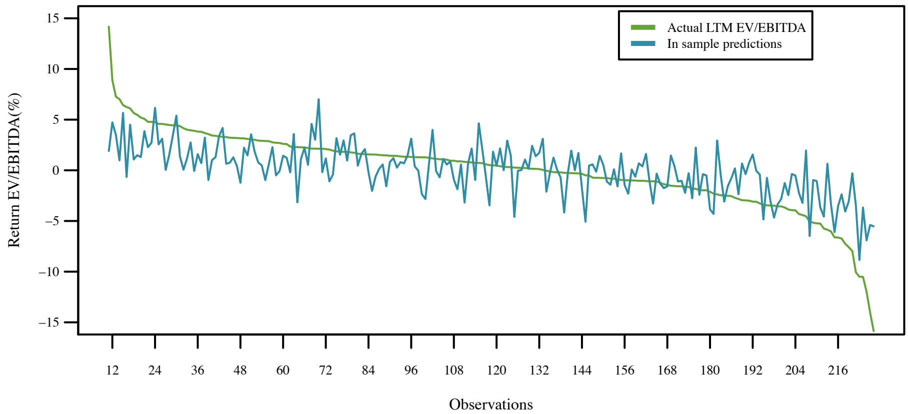


Figure 17: Sorted in-sample predictions for *Dp*

B.II Proof by Induction of Formula for *Rx* Model Predictions

Given that $\mathbb{E}(\epsilon_t) = 0$, the linear estimator of y_t in the *Rx* models is specified by the following:

$$\hat{y}_t = \hat{\alpha}_r + \hat{\gamma} y_{t-1} + \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_t \tag{3}$$

Now, the following is assumed to hold for k:

$$\hat{y}_{t+k} = \hat{\alpha}_r \frac{1-\hat{\gamma}^k}{1-\hat{\gamma}} + \hat{\gamma}^k y_t + \sum_{i=1}^{k} \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+i} \hat{\gamma}^{k-i} \tag{4}$$

Inserting (4) into (3) we obtain:

$$\begin{aligned} \hat{y}_{t+1+k} &= \hat{\alpha}_r + \hat{\gamma}\hat{\alpha}_r \frac{1-\hat{\gamma}^k}{1-\hat{\gamma}} + \hat{\gamma}^{k+1} y_t + \sum_{i=1}^{k} \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+i} \hat{\gamma}^{k+1-i} + \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+1+k} \\ &= \hat{\alpha}_r \frac{1-\hat{\gamma}^{k+1}}{1-\hat{\gamma}} + \hat{\gamma}^{k+1} y_t + \sum_{i=1}^{k+1} \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+i} \hat{\gamma}^{k+1-i} \end{aligned}$$

Hence, if (4) is true for k, it is true for $k+1$. Now, looking at (3) for $t_0 = t+1$:

$$\begin{aligned} \hat{y}_{t+1} &= \hat{\alpha}_r + \hat{\gamma} y_t + \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+1} \\ &= \hat{\alpha}_r \frac{1-\hat{\gamma}^1}{1-\hat{\gamma}} + \hat{\gamma}^1 y_t + \sum_{i=1}^{1} \hat{\boldsymbol{\beta}}_r \cdot \boldsymbol{x}_{t+i} \hat{\gamma}^{1-i} \end{aligned}$$

Which is (4) for $k = 1$. By induction, this entails that that (4) is true for all $k \in \mathbb{N}$. Q.E.D.

B.III Mean Absolute Deviation Prediction Test

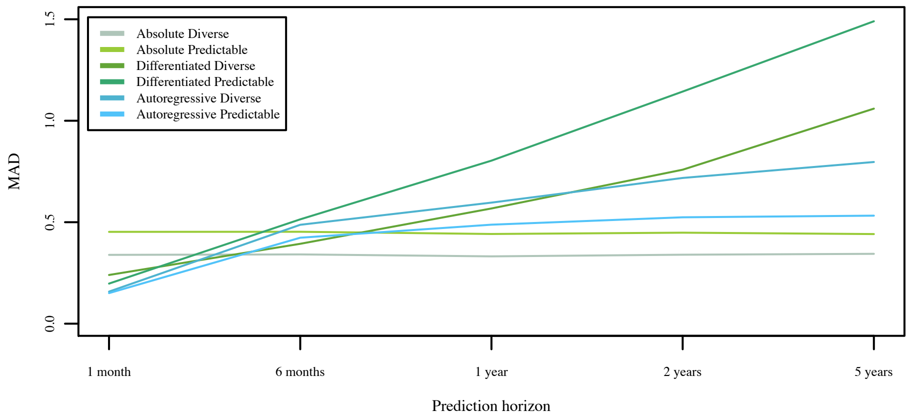


Figure 18: In-sample prediction MAD values given independent variables in horizon period

B.IV Johansen Test Results for *Ad*

Table 2: Results from Johansen test on *Ad*

	Johansen Statistic			
	Test	10%	5%	1%
r ≤ 4	0.52	6.50	8.18	11.65
r ≤ 3	9.00	15.66	17.95	23.52
r ≤ 2	25.14	28.71	31.52	37.22
r ≤ 1	48.91	45.23	48.28	55.43
r = 0	77.49	66.49	70.60	78.87
Eigenvalues				
MedianEVEBITDA	iEU10	sEUCC	vST50	mEUFAYOY
0.12	0.10	0.07	0.038	0.002

B.V Small vs. Large Company Index Analysis

The following indices are define to test if large EV and large EBITDA companies are affected less by changes in volatility:

- *Small EV Index:* Companies with EV ≤ 30% quantile at time t
- *Large EV Index:* Companies with EV ≥ 70% quantile at time t
- *Small EBITDA Index:* Companies with EBITDA ≤ 30% quantile at time t
- *Large EBITDA Index:* Companies with EBITDA ≥ 70% quantile at time t
- *Small EV & EBITDA Index:* Companies with EV ≤ 30% quantile and EBITDA ≤ 30% quantile at time t

- *Large EV & EBITDA Index:* Companies with EV ≥ 70% quantile and EBITDA ≥ 70% quantile at time t

The median of the company multiples at time t is taken in each intersection to create pricing indices.

Table 3: Regression results for small and large company indices

	Dependent variable:					
	Small EV	Large EV	Small EBITDA	Large EBITDA	Small EV & EBITDA	Large EV & EBITDA
iEUspd101	-0.142^{***}	-0.379^{***}	-0.252^{***}	-0.361^{***}	-0.145^{***}	-0.364^{***}
	(0.035)	(0.036)	(0.034)	(0.036)	(0.036)	(0.036)
mEUCPI	-0.417^{***}	-0.483^{***}	-0.454^{***}	-0.458^{***}	-0.433^{***}	-0.492^{***}
	(0.033)	(0.034)	(0.032)	(0.034)	(0.034)	(0.034)
mEUpYOY	0.239^{***}	0.397^{***}	0.352^{***}	0.402^{***}	0.242^{***}	0.397^{***}
	(0.036)	(0.037)	(0.035)	(0.037)	(0.037)	(0.037)
vST50BCt	-0.623^{***}	-0.312^{***}	-0.486^{***}	-0.341^{***}	-0.608^{***}	-0.326^{***}
	(0.035)	(0.036)	(0.034)	(0.036)	(0.036)	(0.035)
Constant	−0.000	0.000	−0.000	−0.000	−0.000	0.000
	(0.032)	(0.033)	(0.031)	(0.033)	(0.033)	(0.032)
Observations	223	223	223	223	223	223
R^2	0.776	0.764	0.791	0.767	0.768	0.769
Adjusted R^2	0.772	0.760	0.788	0.763	0.764	0.765
Residual Std. Error (df = 218)	0.478	0.490	0.461	0.487	0.486	0.485
F Statistic (df = 4; 218)	188.672^{***}	176.663^{***}	206.688^{***}	179.321^{***}	180.409^{***}	181.864^{***}

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

The coefficients for the Small EV & EBITDA index resemble coefficients of the Small EV index suggesting that small EV corresponds to small EBITDA. On the contrary, the coefficients for the Large EV & EBITDA index resemble coefficients of the Large EBITDA index and equivalently for the large indices suggesting that large EV corresponds to large EBITDA. In short, this reflects the size independency of multiples.

For changes in volatility, we observe that large companies by both EV and EBITDA are relatively less sensitive to volatility compared to small companies by both EV and EBITDA. Thus, the hypothesis that large companies are less affected by changes in volatility has been verified.

For changes in inflation, we observe that large companies are somewhat more affected compared to small companies. Growth and interest rate spread impacts larger companies more severely.

B.VI Ramsey Regression Equation Specification Error Test (RESET) Results

The R *lmtest* package was used for conducting a Ramsey RESET test for each of the linear models. Nonlinear relationships are conspicuous in *Ad* and *Dx*, but not in *Ap*.

Table 4: Ramsey RESET test results

	RESET statistic	df1	df2	p-value
Ad	5.582	2	216	0.004
Ap	0.953	2	216	0.387
Dd	5.425	2	219	0.005
Dp	2.930	2	220	0.056
Rp	0.208	2	217	0.812
Rd	2.050	2	216	0.131

B.VII Results from Taylor Regressions

A subset regression was conducted using third degree perturbation of the Taylor series. Table 5 displays R^2 and BIC values for the best regression for each # of variables. The R^2 are generally low, though with a substantial increase from $n = 1$ to $n = 2$. $n = 2$ also represents a strong BIC, compared to $n = 1$, $n = 3$ and $n = 4$. Table 6 displays the variable selections, where a = iEUspd101, b = mEUCPI, c = mEUpYOYlr and d = vSt50BCt, and concatenations represent product variables of the mentioned. The most significant explanatory variables in $n = 2$ are aa and aaa, indicating that the Taylor expansion perturbations of the term structure spread have an explanatory value. Table 7 displays the resulting coefficients for $n = 1$ and $n = 2$, and the F-test is significant at 90% and 99% respectively, keeping the potential errors of the residuals in mind. The $n = 2$ coefficients are consistent with magnitude and sign expected from a Taylor expansion of an inverse relationship.

Table 5: R^2 and BIC-values for best subset-regression for each # of variables

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
R^2	0.008	0.041	0.041	0.058	0.084	0.120	0.150	0.165	0.178	0.201
BIC	8.023	4.760	9.240	9.758	7.758	3.290	−0.119	0.184	1.118	−0.761

Table 6: Subset regression variable selection results for 1-10 variables

	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$
a						*				*
b							*			*
c										
d					*			*	*	
aa		*	*	*			*		*	*
ab								*		
ac									*	
ad			*	*		*				*
bb					*	*	*	*	*	*
bc										
bd					*	*	*	*	*	*
cc										
cd										
dd										
aaa		*	*							*
aab										
aac										
aad				*			*		*	
abb										
abc										
abd								*		
acc	*									
acd										
add				*						
bbb								*	*	*
bbc										
bbd					*	*	*	*	*	*
bcc										
bcd										
bdd					*	*	*	*	*	*
ccc										
ccd										
cdd										
ddd										

Table 7: Resulting coefficients for regressions on subset regression-selected variables for $n = 1$ and $n = 2$

	Dependent variable:	
	residuals	
	$n = 1$-regression	$n = 2$-regression
acc	0.0004*	
	(0.0003)	
aaa		0.128***
		(0.039)
aa		−0.299***
		(0.098)
Constant	−0.020	0.131*
	(0.039)	(0.074)
Observations	223	223
R^2	0.012	0.050
Adjusted R^2	0.008	0.041
Res. Std. Error	0.553	0.542
F Statistic	2.784*	5.802***
Note:		*p<0.1; **p<0.05; ***p<0.01

C Appendix - Chapter 5

C.I Bivariate Plots of log EV and log EBITDA

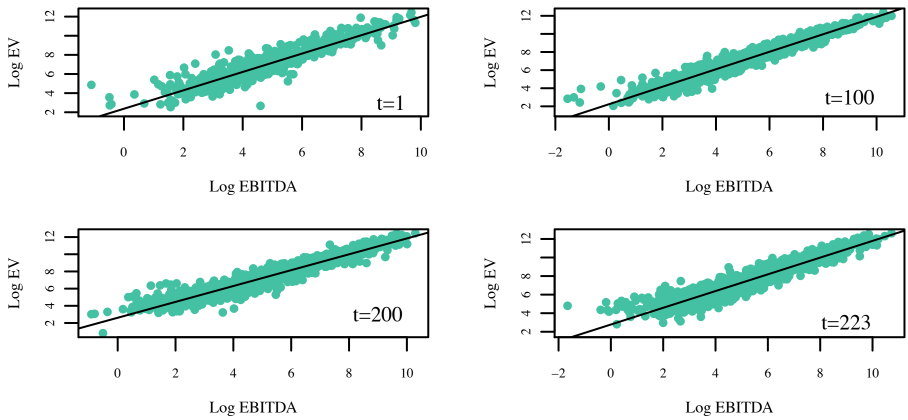


Figure 19: Bivariate plots of log EV versus log EBITDA for $t = \{1, 100, 200, 223\}$

C.II Regression Results for log EV versus log EBITDA

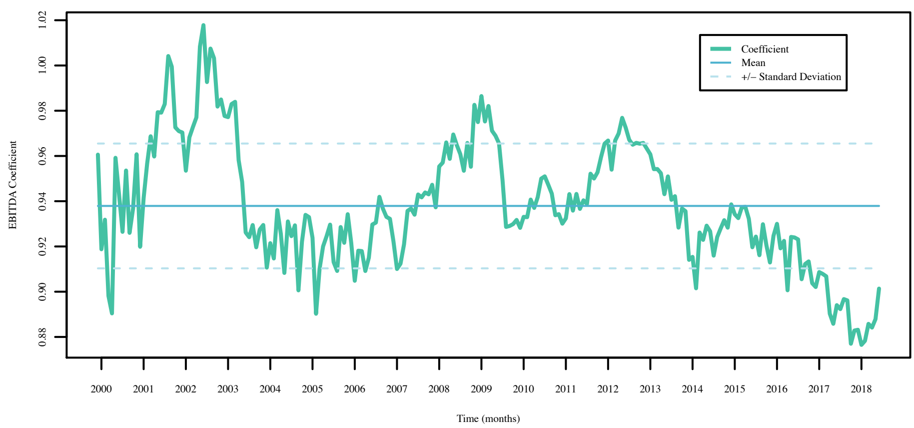


Figure 20: β_t-coefficients from the estimation of Equation 5.2. Corresponding p-values are <0.01 at each intersection

C.III Normality Test on Residuals

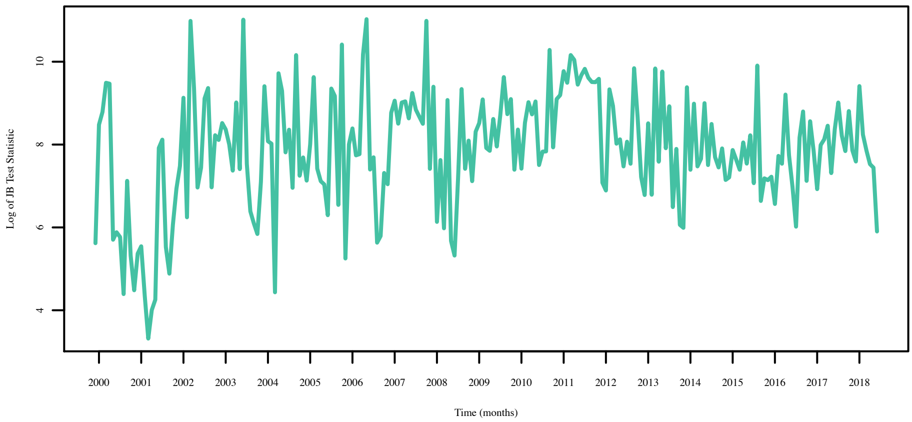


Figure 21: Log of JB test statistic at each intersection. Corresponding p-values are <0.01 for all t

C.IV KS Test on Residuals for Student-t Distribution

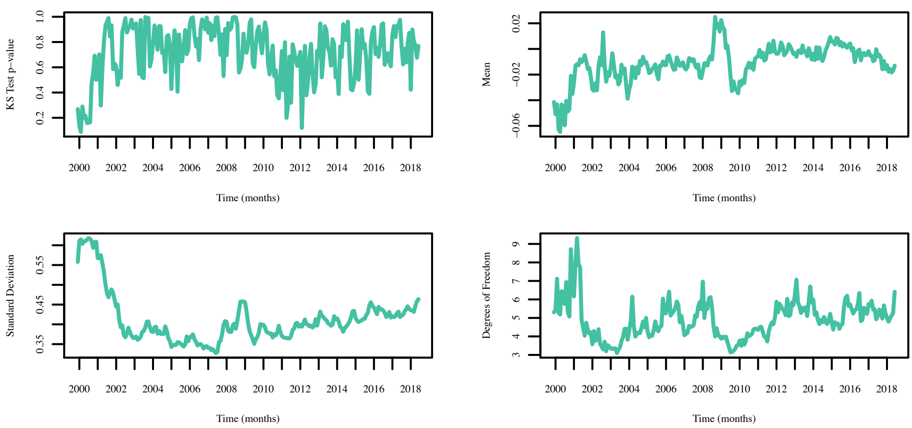


Figure 22: Plots of KS test p-value and historically fitted mean, standard deviation and degrees of freedom for student t distribution of residuals

C.V Regression EV/EBITDA vs. EV and vs. EBITDA

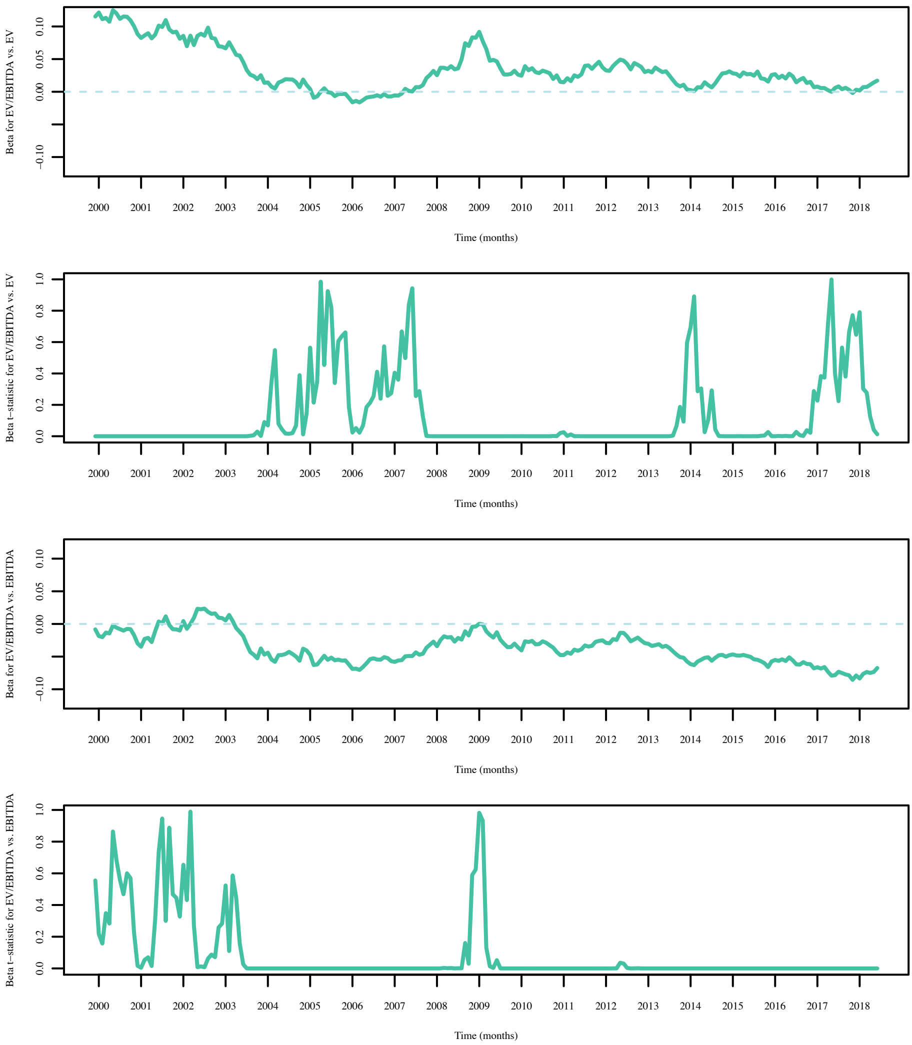


Figure 23: β_t-coefficients and corresponding p-values for regression of log EV/EBITDA vs. log EV (top two plots) and log EV/EBITDA vs. log EBITDA (bottom two plots)

C.VI Johansen Test for Regression on Quantiles

Table 8: Johanesen test statistics for compositions of individual quantiles and exogenous variables for *Ap*

	Critical value			Quantiles								
H_0	10%	5%	1%	10%	20%	30%	40%	50%	60%	70%	80%	90%
$r <= 4$	6.5	8.2	11.7	6.7	6.1	5.3	4.9	4.3	4.2	4.0	3.9	4.7
$r <= 3$	15.7	17.9	23.5	15.8	16.7	16.2	15.9	15.9	17.0	16.4	16.2	17.0
$r <= 2$	28.7	31.5	37.2	36.7	40.1	41.0	42.4	43.9	46.5	43.2	39.2	41.9
$r <= 1$	45.2	48.3	55.4	77.4	80.1	79.8	80.6	82.1	85.9	81.9	80.2	82.9
$r = 0$	66.5	70.6	78.9	123.2	123.3	121.3	125.0	125.8	132.1	129.5	126.4	138.0

C.VII Regression on Quantiles

Table 9: Normalized regression coefficients for 9 quantiles of EV/EBITDA for *Ap* model

	Dependent variable:								
	Normalized EV/EBITDA Quantiles								
	10%	20%	30%	40%	Median	60%	70%	80%	90%
iEUspd101	-0.159^{***}	-0.192^{***}	-0.213^{***}	-0.246^{***}	-0.268^{***}	-0.291^{***}	-0.313^{***}	-0.336^{***}	-0.350^{***}
	(0.041)	(0.038)	(0.036)	(0.034)	(0.032)	(0.032)	(0.034)	(0.042)	(0.053)
mEUCPI	-0.269^{***}	-0.379^{***}	-0.431^{***}	-0.455^{***}	-0.475^{***}	-0.480^{***}	-0.472^{***}	-0.424^{***}	-0.323^{***}
	(0.039)	(0.036)	(0.034)	(0.032)	(0.031)	(0.031)	(0.032)	(0.040)	(0.050)
mEUpYOY	0.304^{***}	0.310^{***}	0.319^{***}	0.324^{***}	0.335^{***}	0.347^{***}	0.355^{***}	0.377^{***}	0.365^{***}
	(0.043)	(0.039)	(0.037)	(0.035)	(0.034)	(0.033)	(0.035)	(0.043)	(0.055)
vST50BCt	-0.576^{***}	-0.551^{***}	-0.528^{***}	-0.508^{***}	-0.489^{***}	-0.464^{***}	-0.429^{***}	-0.329^{***}	-0.188^{***}
	(0.041)	(0.038)	(0.035)	(0.034)	(0.032)	(0.032)	(0.034)	(0.042)	(0.053)
Constant	−0.000	−0.000	0.000	−0.000	−0.000	−0.000	0.000	0.000	0.000
	(0.038)	(0.034)	(0.032)	(0.031)	(0.030)	(0.029)	(0.031)	(0.038)	(0.048)
Observations	223	223	223	223	223	223	223	223	223
R^2	0.689	0.742	0.769	0.790	0.809	0.812	0.788	0.680	0.490
Adjusted R^2	0.684	0.737	0.765	0.786	0.806	0.808	0.784	0.674	0.481
Residual Std. Error (df = 218)	0.562	0.513	0.485	0.463	0.441	0.438	0.465	0.571	0.721
F Statistic (df = 4; 218)	120.976^{***}	156.597^{***}	181.675^{***}	204.451^{***}	231.099^{***}	235.272^{***}	202.598^{***}	115.625^{***}	52.366^{***}

Note: Variables were normalized in advance of regression for comparability $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 10: Regression coefficients for all 6 models on all 9 quantiles. Normalized dependent variables variables for comparison across quantiles

		Quantiles								
		10%	20%	30%	40%	50%	60%	70%	80%	90%
Ad	Intercept	2.17	2.29	2.36	2.41	2.45	2.46	2.42	2.19	1.68
	iEU10	−0.08	−0.15	−0.19	−0.20	−0.21	−0.20	−0.19	−0.14	−0.02
	sEUCC	0.03	0.05	0.05	0.06	0.07	0.07	0.08	0.10	0.10
	vST50	−0.06	−0.06	−0.05	−0.05	−0.05	−0.04	−0.04	−0.03	−0.02
	mEUFAYOY	−0.03	−0.03	−0.03	−0.02	−0.02	−0.02	−0.01	−0.01	0.001
Ap	Intercept	13.85	13.54	13.15	12.78	12.43	11.90	11.13	8.80	5.42
	iEUspd101	−0.23	−0.28	−0.31	−0.35	−0.38	−0.42	−0.45	−0.48	−0.50
	mEUCPI	−0.28	−0.39	−0.45	−0.47	−0.49	−0.50	−0.49	−0.44	−0.34
	mEUpYOY	0.06	0.06	0.06	0.07	0.07	0.07	0.07	0.08	0.07
	vST50BCt	−8.35	−7.99	−7.66	−7.37	−7.10	−6.73	−6.23	−4.77	−2.72
Dd	vST50lr	−3.16	−3.37	−3.48	−3.40	−3.52	−3.50	−3.44	−3.27	−3.12
	sEUCCd	0.22	0.22	0.21	0.22	0.22	0.21	0.20	0.18	0.14
Dp	vST50lr	−3.19	−3.40	−3.50	−3.43	−3.54	−3.53	−3.47	−3.30	−3.14
Rd	Intercept	−0.01	−0.01	−0.02	−0.02	−0.02	−0.02	−0.02	−0.01	−0.01
	y_{t-1}	−0.11	−0.10	−0.09	−0.09	−0.08	−0.08	−0.09	−0.09	−0.11
	vST50lr	−3.08	−3.30	−3.42	−3.35	−3.46	−3.45	−3.40	−3.23	−3.08
	sEUCCd	0.23	0.23	0.22	0.23	0.23	0.22	0.21	0.19	0.15
Rp	Intercept	1.61	1.81	1.84	1.96	1.92	1.87	1.66	1.33	1.05
	y_{t-1}	−0.41	−0.46	−0.47	−0.50	−0.49	−0.47	−0.42	−0.32	−0.23
	vST50lr	−2.13	−2.24	−2.34	−2.21	−2.34	−2.39	−2.48	−2.55	−2.58
	vST50	−0.05	−0.05	−0.05	−0.06	−0.06	−0.05	−0.05	−0.04	−0.03
	mEUCPIBCt	−0.17	−0.20	−0.21	−0.23	−0.22	−0.23	−0.21	−0.18	−0.15

Table 11: *Ad* regression results for all quantiles

	Dependent variable:								
	EV/EBITDA Quantiles								
	10%	20%	30%	40%	Median	60%	70%	80%	90%
iEU10	−0.067***	−0.137***	−0.191***	−0.235***	−0.269***	−0.282***	−0.299***	−0.273***	−0.072
	(0.016)	(0.017)	(0.018)	(0.018)	(0.019)	(0.019)	(0.021)	(0.032)	(0.063)
sEUCC	0.027***	0.042***	0.055***	0.069***	0.083***	0.102***	0.130***	0.191***	0.309***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.005)	(0.007)	(0.015)
vST50	−0.052***	−0.051***	−0.053***	−0.055***	−0.058***	−0.059***	−0.060***	−0.056***	−0.053***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)	(0.013)
mEUFAYOY	−0.025***	−0.027***	−0.027***	−0.027***	−0.025***	−0.022***	−0.019***	−0.010**	0.003
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.009)
Constant	6.285***	7.731***	8.903***	10.087***	11.256***	12.477***	14.042***	16.263***	20.340***
	(0.082)	(0.085)	(0.090)	(0.092)	(0.094)	(0.095)	(0.104)	(0.159)	(0.318)
Observations	223	223	223	223	223	223	223	223	223
R^2	0.766	0.809	0.830	0.857	0.876	0.893	0.901	0.855	0.747
Adjusted R^2	0.761	0.805	0.827	0.854	0.874	0.891	0.899	0.853	0.742
Residual Std. Error (df = 218)	0.392	0.406	0.430	0.442	0.450	0.454	0.499	0.759	1.521
F Statistic (df = 4; 218)	178.066***	230.387***	266.906***	326.253***	385.690***	456.593***	494.318***	322.585***	160.861***
Note:									*p<0.1; **p<0.05; ***p<0.01

Table 12: *Dd* regression results for all quantiles

	Dependent variable:								
	Log return EV/EBITDA Quantiles								
	10%	20%	30%	40%	Median	60%	70%	80%	90%
vST50lr	-0.160^{***}	-0.144^{***}	-0.138^{***}	-0.127^{***}	-0.132^{***}	-0.126^{***}	-0.130^{***}	-0.127^{***}	-0.137^{***}
	(0.014)	(0.011)	(0.010)	(0.010)	(0.009)	(0.009)	(0.010)	(0.011)	(0.013)
sEUCCd	0.011^{***}	0.009^{***}	0.008^{***}	0.008^{***}	0.008^{***}	0.008^{***}	0.007^{***}	0.007^{***}	0.006^{***}
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Observations	223	223	223	223	223	223	223	223	223
R^2	0.452	0.500	0.516	0.506	0.530	0.521	0.494	0.441	0.377
Adjusted R^2	0.447	0.496	0.511	0.501	0.526	0.517	0.489	0.436	0.372
Residual Std. Error (df = 221)	0.038	0.030	0.028	0.026	0.026	0.025	0.027	0.029	0.035
F Statistic (df = 2; 221)	91.111^{***}	110.574^{***}	117.635^{***}	113.110^{***}	124.821^{***}	120.184^{***}	107.904^{***}	87.031^{***}	66.944^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 13: *Dp* regression results for all quantiles

	Dependent variable:								
	Log return EV/EBITDA Quantiles								
	10%	20%	30%	40%	Median	60%	70%	80%	90%
vST50lr	-0.162^{***}	-0.146^{***}	-0.139^{***}	-0.128^{***}	-0.133^{***}	-0.127^{***}	-0.131^{***}	-0.128^{***}	-0.138^{***}
	(0.015)	(0.012)	(0.011)	(0.011)	(0.010)	(0.010)	(0.011)	(0.011)	(0.013)
Observations	223	223	223	223	223	223	223	223	223
R^2	0.344	0.390	0.415	0.398	0.425	0.421	0.407	0.368	0.332
Adjusted R^2	0.341	0.387	0.413	0.395	0.422	0.419	0.405	0.365	0.329
Residual Std. Error (df = 222)	0.041	0.033	0.030	0.029	0.029	0.027	0.029	0.031	0.036
F Statistic (df = 1; 222)	116.212^{***}	142.072^{***}	157.709^{***}	146.729^{***}	163.910^{***}	161.514^{***}	152.582^{***}	129.026^{***}	110.285^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

Table 14: *Rd* regression results for all quantiles

	Dependent variable:								
	EV/EBITDA Quantiles								
	10%	20%	30%	40%	Median	60%	70%	80%	90%
$y_{t-1,q}$	0.972^{***}	0.978^{***}	0.981^{***}	0.983^{***}	0.983^{***}	0.983^{***}	0.981^{***}	0.980^{***}	0.975^{***}
	(0.013)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)	(0.012)	(0.012)
vST50lr	-0.634^{***}	-0.720^{***}	-0.806^{***}	-0.842^{***}	-0.982^{***}	-1.045^{***}	-1.217^{***}	-1.397^{***}	-1.894^{***}
	(0.055)	(0.056)	(0.059)	(0.064)	(0.070)	(0.077)	(0.096)	(0.125)	(0.195)
sEUCCd	0.045^{***}	0.048^{***}	0.049^{***}	0.054^{***}	0.060^{***}	0.064^{***}	0.071^{***}	0.078^{***}	0.090^{***}
	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.015)	(0.024)
Constant	0.128^{**}	0.126^{*}	0.124^{*}	0.129^{*}	0.141	0.157	0.198^{*}	0.240^{*}	0.365^{*}
	(0.059)	(0.065)	(0.070)	(0.077)	(0.086)	(0.096)	(0.118)	(0.142)	(0.187)
Observations	223	223	223	223	223	223	223	223	223
R^2	0.965	0.973	0.976	0.978	0.978	0.977	0.973	0.971	0.969
Adjusted R^2	0.965	0.973	0.976	0.977	0.977	0.977	0.972	0.970	0.968
Residual Std. Error (df = 219)	0.151	0.152	0.161	0.174	0.191	0.209	0.260	0.340	0.533
F Statistic (df = 3; 219)	$2,032.603^{***}$	$2,654.535^{***}$	$2,974.166^{***}$	$3,189.823^{***}$	$3,174.165^{***}$	$3,153.289^{***}$	$2,614.874^{***}$	$2,430.799^{***}$	$2,270.458^{***}$

Note: *p<0.1; **p<0.05; ***p<0.01

Table 15: *Rp* regression results for all quantiles

	Dependent variable:								
	EV/EBITDA Quantiles								
	10%	20%	30%	40%	Median	60%	70%	80%	90%
$y_{t-1,q}$	0.899***	0.897***	0.900***	0.900***	0.899***	0.903***	0.910***	0.935***	0.951***
	(0.019)	(0.016)	(0.015)	(0.015)	(0.015)	(0.015)	(0.016)	(0.015)	(0.013)
vST50lr	−0.450***	−0.502***	−0.569***	−0.577***	−0.688***	−0.748***	−0.913***	−1.130***	−1.613***
	(0.063)	(0.062)	(0.065)	(0.069)	(0.077)	(0.083)	(0.104)	(0.133)	(0.203)
vST50	−0.010***	−0.011***	−0.012***	−0.014***	−0.015***	−0.015***	−0.016***	−0.014***	−0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)
mEUCPIBCt	−0.033***	−0.042***	−0.047***	−0.054***	−0.060***	−0.066***	−0.071***	−0.075***	−0.094***
	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.010)	(0.012)	(0.015)	(0.023)
Constant	0.775***	0.952***	1.053***	1.195***	1.334***	1.408***	1.476***	1.310***	1.345***
	(0.121)	(0.129)	(0.139)	(0.151)	(0.171)	(0.188)	(0.225)	(0.244)	(0.284)
Observations	223	223	223	223	223	223	223	223	223
R^2	0.967	0.975	0.979	0.980	0.980	0.980	0.975	0.972	0.971
Adjusted R^2	0.966	0.975	0.978	0.980	0.979	0.979	0.974	0.972	0.970
Residual Std. Error (df = 218)	0.148	0.146	0.153	0.164	0.181	0.198	0.251	0.332	0.518
F Statistic (df = 4; 218)	1,579.876***	2,157.560***	2,484.210***	2,708.025***	2,651.599***	2,633.116***	2,116.586***	1,919.259***	1,804.542***

Note: *p<0.1; **p<0.05; ***p<0.01

Table 16: Hosking test statistics for Residuals from VAR model in Equation 5.12

Lags	Statistic	Degrees of Freedom	p-value
5	302.083	320	0.757
10	643.778	640	0.451
15	1,015.347	960	0.105
20	1,394.331	1,280	0.014
25	1,770.809	1,600	0.002
30	2,091.937	1,920	0.003

C.VIII Johansen test for Quantiles

Table 17: Results from Johansen test on the system of nine quantiles.

	Johansen Statistic			
	Test	10%	5%	1%
r ≤ 8	0.86	6.5	8.18	11.65
r ≤ 7	10.76	15.66	17.95	23.52
r ≤ 6	34.86	28.71	31.52	37.22
r ≤ 5	69.41	45.23	48.28	55.43
r ≤ 4	119.96	66.49	70.6	78.87
r ≤ 3	182.77	85.18	90.39	104.2
r ≤ 2	254.76	118.99	124.25	136.06
r ≤ 1	335.24	151.38	157.11	168.92
r = 0	424.38	186.54	192.84	204.79

C.IX Moments of EV/EBITDA Distribution

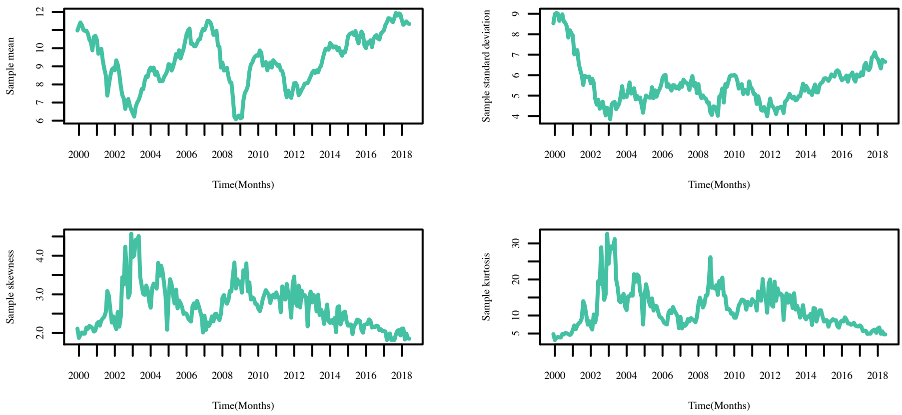


Figure 24: EV/EBITDA distribution median, standard deviation, skewness and kurtosis development

C.X Principal Component Analysis on Quantiles

Table 18: Coefficients for resulting models from subset regression on principal components

	Dependent variable:					
	$\Delta pc2$	pc2	$\Delta pc3$	pc3	$\Delta pc4$	pc4
fxUSDEURlr	-3.584^{**}					
	(1.482)					
fxUSDEUR		-13.092^{***}				
		(0.987)				
iEU10d			-0.334^{***}			
			(0.122)			
iEU10				-0.542^{***}		
				(0.037)		
mEUCPId					0.016	
					(0.110)	
mEUCPI						0.175^{***}
						(0.034)
Constant	-0.028	15.910^{***}	0.008	1.537^{***}	0.002	-0.306^{***}
	(0.035)	(1.212)	(0.023)	(0.123)	(0.028)	(0.068)
Observations	222	223	222	223	222	223
R^2	0.026	0.443	0.033	0.489	0.0001	0.107
Adjusted R^2	0.021	0.440	0.028	0.487	-0.004	0.103
Residual Std. Error	0.523	2.515	0.339	0.928	0.420	0.487
F Statistic	5.849^{**}	175.774^{***}	7.478^{***}	211.549^{***}	0.022	26.605^{***}
Note:						*p<0.1; **p<0.05; ***p<0.01

C.XI Parametric Distribution Parameter Estimation using Squared Distance and Logarithmic Sum

D and A from Section 5.2 may be defined in multiple ways. It is natural to fit a kernel $k_t(m)$ to the empirical data Ω_t, and to use a squared distance measure D and a logarithmic sum aggregation function, from which the following loss function is obtained:

$$\Lambda = \sum_{t=1}^{T_{max}} log \int_{m=0}^{\infty} (f_{Rx_t}(m) - k_t(m))^2 dm \tag{5}$$

The loss function may be minimized by setting $\nabla\Lambda = \mathbf{0}$, assuming that f_{Rx_t} has a particular parametric form and solving the resulting set of equations:

$$\frac{\partial H}{\partial \boldsymbol{r}} = \sum_{t=1}^{T} \frac{\int_{m=0}^{\infty} 2(f_{Rx_t}(m) - k_t(m))dm}{\int_{m=0}^{\infty} (f_{Rx_t}(m) - k_t(m))^2 dm} \frac{\partial f_{Rx_t}}{\partial \boldsymbol{r}} = \mathbf{0} \tag{6}$$

C.XII Derivation of Equation 5.23

We start with the likelihood function:

$$L(\boldsymbol{\beta}_{1:B}, \alpha_{1:B}; \boldsymbol{x}_{1:T}) = \prod_{t=1}^{T} \prod_{\omega \in \Omega_t} P(M_t = \omega) = \prod_{t=1}^{T} \prod_{\omega \in \Omega_t} \sum_{b=1}^{B} \nu_{bt} \mathbb{1}_{l_b \leq m < u_b} \tag{7}$$

1. Take the logarithm on both sides
2. Sum Ω_t over $\mathbb{1}_{l_b \leq \omega < u_b}$
3. Use Equation 5.20 to expand ν_{bt}
4. Separate the logarithm
5. Sum from $b = 1$ to B in first expression

$$\begin{aligned}
\mathscr{L}(\boldsymbol{\beta}_{1:B}, \alpha_{1:B}; \boldsymbol{x}_{1:T}) &= \sum_{t=1}^{T} \sum_{\omega \in \Omega_t} \log \sum_{b=1}^{B} \nu_{bt} \mathbb{1}_{l_b \leq \omega < u_b} \\
&= \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log \nu_{bt} \\
&= \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log \frac{\gamma_t}{1 + e^{\alpha_b + \boldsymbol{\beta}_b \cdot \boldsymbol{x}_t}} \\
&= \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log \gamma_t - \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log(1 + e^{\alpha_b + \boldsymbol{\beta}_b \cdot \boldsymbol{x}_t}) \\
&= \sum_{t=1}^{T} |\Omega_t| \log \gamma_t - \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log(1 + e^{\alpha_b + \boldsymbol{\beta}_b \cdot \boldsymbol{x}_t})
\end{aligned}$$

Q.E.D.

C.XIII Overview of 0-value Empirical Frequencies

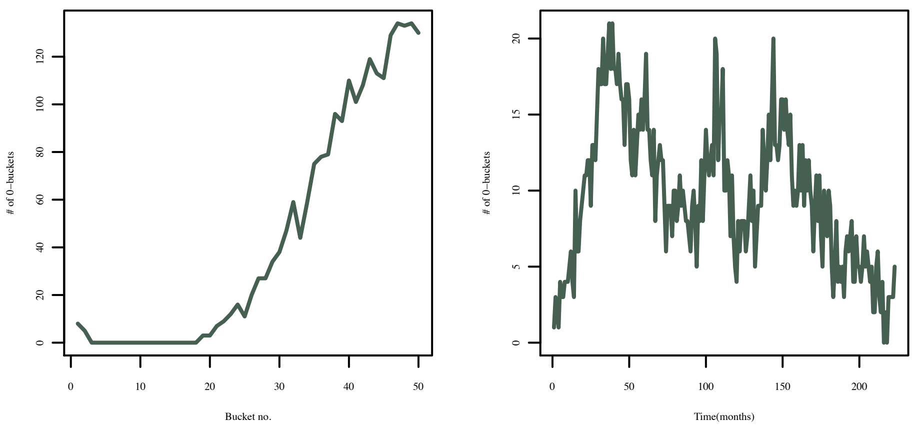


Figure 25: # of 0-buckets in-sample

C.XIV Non-Parametric Distribution Model Coefficients

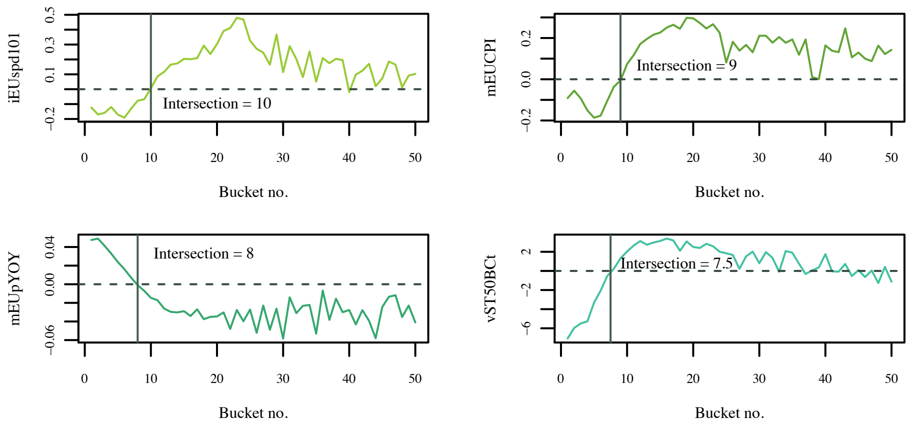


Figure 26: Coefficients for bucket model with logistic regression variables

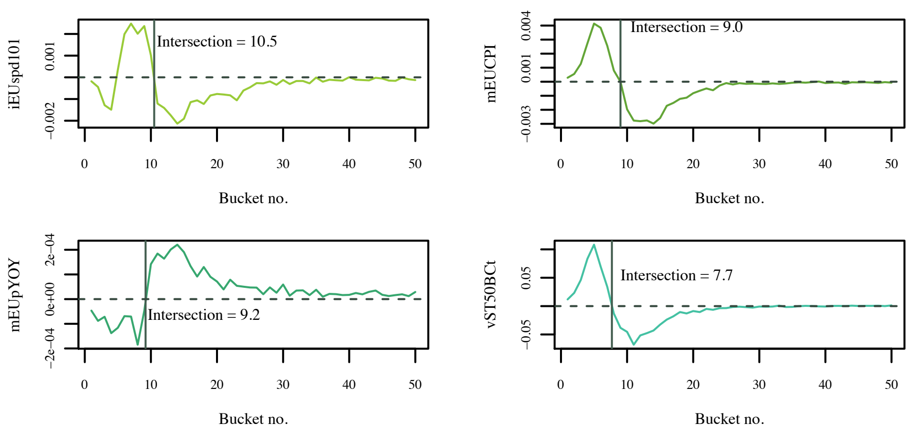


Figure 27: Coefficients for bucket model with lagged bucket variables

C.XV An Analysis on Low vs. High Value Companies

We defined the following indices in order to test how low and high valued companies are affected by the exogenous variables from *Ap*:

- *Low value company index:* Companies with median EV/EBITDA less than or equal to the median 30% EV/EBITDA quantile over the whole time period under consideration
- *High value company index:* Companies with median EV/EBITDA greater than or equal to the median 70% EV/EBITDA quantile over the whole time period under consideration

After finding these sets of companies, the median of the companies' multiples were taken to create a pricing index.

Table 19: Regression results for low and high value company multiple indices

	Dependent variable:	
	Low value	High value
iEUspd101	-0.205^{***}	-0.289^{***}
	(0.041)	(0.032)
mEUCPI	-0.256^{***}	-0.468^{***}
	(0.039)	(0.030)
mEUpYOY	0.355^{***}	0.356^{***}
	(0.043)	(0.033)
vST50BCt	-0.516^{***}	-0.469^{***}
	(0.041)	(0.032)
Constant	0.000	−0.000
	(0.038)	(0.029)
Observations	223	223
R^2	0.690	0.817
Adjusted R^2	0.684	0.814
Residual Std. Error (df = 218)	0.562	0.432
F Statistic (df = 4; 218)	121.125^{***}	243.341^{***}
Note:	*p<0.1; **p<0.05; ***p<0.01	

We observe that term structure spread is more important for high-valued companies, volatility is more important for low valued companies. Industrial production growth has similar impact on both low and high valued companies. However, inflation seems to impact higher valued companies more.

D Appendix - Chapter 6

D.I ACF Functions for Residuals in Naïve Models

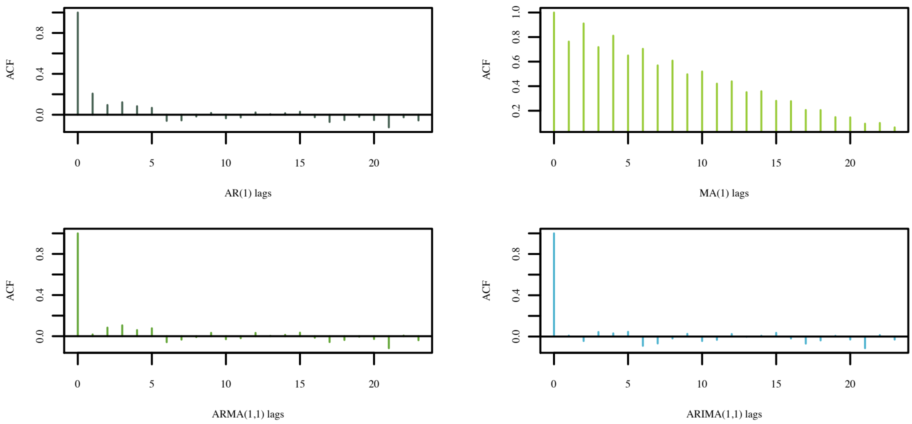


Figure 28: ACF values for the Naïve Forecasting models

D.II Exogenous Variables' Values for Specific Scenarios

Table 20: Underlying variable scenarios

	Unit	2018YE	2019YE	2020YE	2021YE	2022YE
No Change						
vST50	Index	12.60	12.60	12.60	12.60	12.60
iEUspd101	%	1.10	1.10	1.10	1.10	1.10
mEUCPI	%	2.10	2.10	2.10	2.10	2.10
mEUpYOY	%	1.24	1.24	1.24	1.24	1.24
Base Case						
vST50	Index	14.00	17.00	18.00	18.00	18.00
iEUspd101	%	1.06	1.50	1.11	1.25	1.23
mEUCPI	%	1.70	1.60	1.70	1.60	1.65
mEUpYOY	%	1.80	1.45	2.40	2.40	2.40
Low Volatility						
vST50	Index	12.60	12.60	12.60	12.60	12.60
iEUspd101	%	1.06	1.50	1.11	1.25	1.23
mEUCPI	%	1.70	1.60	1.70	1.60	1.65
mEUpYOY	%	1.80	1.45	2.40	2.40	2.40
Increasing Volatility						
vST50	Index	16.00	20.00	22.10	22.10	22.10
iEUspd101	%	1.06	1.50	1.11	1.25	1.23
mEUCPI	%	1.70	1.60	1.70	1.60	1.65
mEUpYOY	%	1.80	1.45	2.40	2.40	2.40
High Growth						
vST50	Index	14.00	17.00	18.00	18.00	18.00
iEUspd101	%	1.06	1.50	1.11	1.25	1.23
mEUCPI	%	1.70	1.60	1.70	1.60	1.65
mEUpYOY	%	4.00	5.00	5.40	5.40	5.40
Stagflation						
vST50	Index	20.00	30.00	35.50	35.50	35.50
iEUspd101	%	1.06	1.80	2.00	2.00	2.00
mEUCPI	%	1.90	2.50	3.00	3.00	3.00
mEUpYOY	%	3.50	2.50	2.00	2.00	1.00

D.III Johansen Test on Exogenous Variables of *Ap*

Table 21: Results from Johansen test on exogenous variables of *Ap*

	Johansen Statistic			
	Test	10%	5%	1%
$r \leq 3$	6.556	6.500	8.180	11.650
$r \leq 2$	20.598	15.660	17.950	23.520
$r \leq 1$	55.764	28.710	31.520	37.220
$r = 0$	93.233	45.230	48.280	55.430

D.IV VAR Model for Exogenous Variables in *Ap*

Table 22: VAR model for the independent variables in *Ap*

	Dependent variable:			
	y			
	iEUspd101	mEUCPI	mEUpYOY	vST50BCt
Lagged iEUspd101	0.930***	0.018	0.385**	−0.002
	(0.019)	(0.026)	(0.164)	(0.004)
Lagged mEUCPI	0.010	0.950***	−0.238**	0.003
	(0.013)	(0.018)	(0.113)	(0.003)
Lagged mEUpYOY	−0.011***	0.015***	0.960***	−0.001
	(0.003)	(0.004)	(0.024)	(0.001)
Lagged vST50BCt	0.349*	−0.005	−4.298**	0.845***
	(0.195)	(0.271)	(1.704)	(0.040)
Constant	−0.474	0.057	6.763**	0.242***
	(0.302)	(0.420)	(2.641)	(0.061)
Observations	215	215	215	215
R^2	0.938	0.937	0.905	0.726
Adjusted R^2	0.937	0.936	0.903	0.721
Residual Std. Error (df = 210)	0.178	0.247	1.555	0.036
F Statistic (df = 4; 210)	800.605***	781.271***	499.760***	139.117***
Note:		*p<0.1; **p<0.05; ***p<0.01		

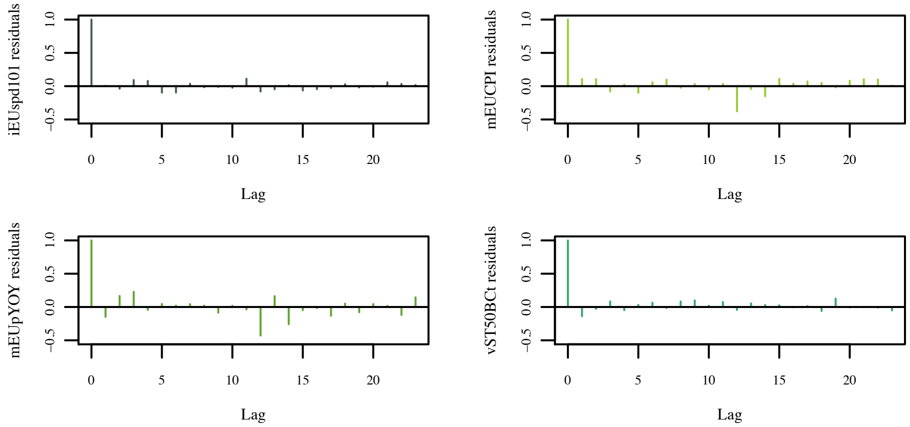


Figure 29: ACFs for VAR model residuals. Similar plots were constructed preliminarily for some residuals on other residuals, without uncovering any significant issues.

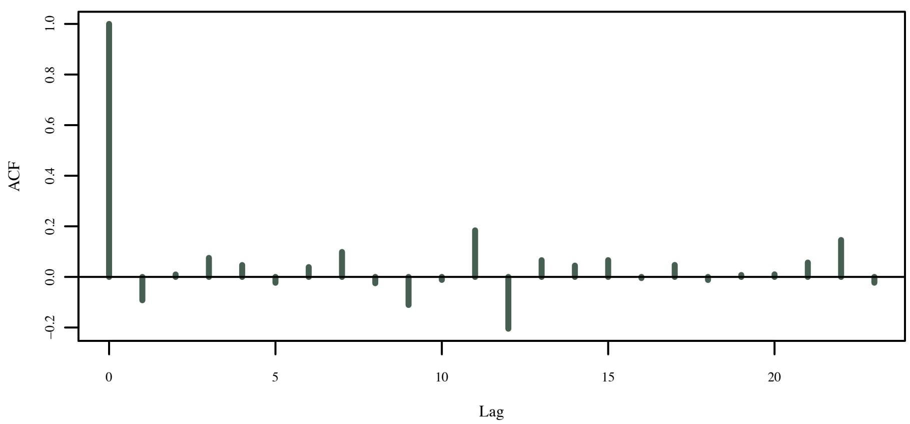


Figure 30: ACF for residual autocorrelation correction AR-model residuals. Some 12-month annual correlation.

E Appendix - Chapter 7

E.I Non-exhaustive List of Further Research Proposals

Markov Swiching Models

Heteroskedasticity and regime switching are likely to be present in the processes we have looked at, and a natural extension of our research would be to create a Markov Switching Model as introduced by Hamilton (1989). We propose a time-varying MSM (see (Diebold et al. 1993)) for the underlying macroeconomic variables in which there are two states, bear and bull market[6], and the probability of switching is dependent on the absolute level. This model will output the probability of entering a bear market at any point in time. The change is governed by the following set of equations:

$$i \in \{1,2\}, \Delta \boldsymbol{x}_t = \boldsymbol{\alpha}_i + \boldsymbol{\epsilon}_{ti} \tag{8}$$

In which $\boldsymbol{x}_t$ are the underlying variables, $\boldsymbol{\alpha}_i$ is the drift constant and $\boldsymbol{\epsilon}_{ti}$ is a i.d.d. distributed error vector. The transition probabilities are thus specified by:

$$P(s_t = 1) = \frac{P(s_{t-1} = 1)}{1 + e^{\boldsymbol{\beta_1} \boldsymbol{x}_{t-1}}} + \frac{e^{\boldsymbol{\beta_2} \boldsymbol{x}_{t-1}} P(s_{t-1} = 2)}{1 + e^{\boldsymbol{\beta_2} \boldsymbol{x}_{t-1}}} \tag{9}$$

An expression for the model parameters, i.e., the initial probability ($P(s_0 = 1)$), the parameters of $f_{\boldsymbol{\epsilon}_{ti}}$, the joint residual distribution ($\boldsymbol{\theta}$), the drift parameters($\boldsymbol{\alpha}(i)$), and the transition probability parameters ($\boldsymbol{\beta_1}, \boldsymbol{\beta_2}$) may be obtained by formulating the MLE expression:

$$\begin{aligned}\mathscr{L}(P(s_0 = 1), \boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}) = \sum_{t=1}^{T} \log \sum_{i=1}^{2} f_{\boldsymbol{\epsilon}_{ti}} (\log \frac{\boldsymbol{x}_t}{\boldsymbol{x}_{t-1}} - \boldsymbol{\alpha}_i) \\ (\frac{1}{1 + e^{\boldsymbol{\beta}_1 \boldsymbol{x}_t}} P(s_{t-1} = 1) + (1 - \frac{1}{1 + e^{\boldsymbol{\beta}_2 \boldsymbol{x}_t}}) P(s_{t-1} = 2))\end{aligned} \tag{10}$$

The explicit maximization of this is $O(2^T)$ and intractable, but an EM algorithm may be used. In the E-step, the current iteration of parameters are used to compute the probability of the states throughout the process, and in the M-step, the most probable states assumed to occur, and the conditionally optimal parameters are computed from $\mathscr{L}$. Simpler Markov Switching Models were successfully estimated but without notable findings. MSM models may be applicable in various forms to describe the underlying environment and the dependency relationship or both simulatenously.

Panel Data Regressions

Our focus has been indices generated from the distribution of multiples, and the distribution by itself. However, individual firms may be identified in each cross section, motivating the definition of a general panel regression problem:

$$m_{ct} = \alpha_c + \boldsymbol{\beta}_c \cdot \boldsymbol{x}_t + \epsilon_{ct} \tag{11}$$

for firm c at time t, where ϵ_{ct} has a non-diagonal conditional covariance matrix and is (potentially) serially dependent. Firm specific variables $\boldsymbol{z}_c$ may be used to define time independent coefficients α_c and $\boldsymbol{\beta}_c$ using linear models:

$$\begin{aligned}\alpha_c &= \alpha_\alpha + \boldsymbol{\beta}_\alpha \cdot \boldsymbol{z}_c + \epsilon^c_\alpha \\ \beta^c_i &= \alpha_{\beta_i} + \boldsymbol{\beta}_{\beta_i} \cdot \boldsymbol{z}_c + \epsilon^c_{\beta_i}\end{aligned} \tag{12}$$

[6]The model is not constrained to distinguish between these two markets, but our hypothesis has been that these are the states that will emerge from the model.

This model is one out of a range of possible panel regression formulations that may be applicable for our case. The panel regression provides a bottom-up approach for understanding the multiple distribution and may be used to enhance predictions, and to gain a better understanding of distributional shifts. Furthermore, the panel perspective may be used in association with machine learning methods and system dynamic approaches. We have only provided a superficial brainstorm of the applications, and implementation would require a careful analysis of challenges associated with firms moving in and out of sample, stationarity, estimation tractability, etc., see Baltagi (2008) for elaboration on the methods.

Size-adjusted Multiple

In the introduction of Chapter 5, a size-adjusted β_t-multiple was introduced. Specifically, the form of the multiple was proposed to be:

$$M_t^* = \frac{V_t}{\pi_t^{\boldsymbol{\beta}_t}} \tag{5.8}$$

using β_t from Equation 5.2. This multiple is by definition linearly independent of the distributions of π_t and V_t, a property almost captured by the regular V_t/π_t multiple. The multiple might be used to render discussion concerning firm size as a factor by itself superflous, and analyses corresponding to the ones we conducted in Chapter 5 could be done, possibly with better accuracy because size-effects are completely mitigated.

Corresponding β_ts may be computed for other kinds of multiples (e.g., EV/EBIT, P/E or industry-specific multiples) in order to adjust them, but also to compare usefulness and explore and contrast properties of the multiples. The α_t values also contain information about the nature of the multiple in relation to the error term, with potential applications.

System Dynamics Approach

The relationship between the macroeconomic environment and pricing levels in capital markets is a complex system with distinct variables interacting through feedback loops - a change in one variable impacts other variables over time, which in turn impacts the original variable and so on. The VAR model of underlying variables developed in Chapter 6 resembles this kind of thinking, but do not necessarily incorporate all relevant effects. Forrester introduced a system dynamic method in the fifties (see Forrester (1997)), allowing for more intricate, non-linear relationships and better understanding of causality between multiples and macroeconomic factors. In particular, investment opportunity supply and demand might be modeled explicitly, allowing for a broader understanding of the complete system, and a granular description of individual factors and agents in the market.

Quantum Mechanical Approach

In Chapter 5, the underlying concern is to estimate the nature of propagation for a probability distribution. In fact, Quantum Mechanics are concerned with a similar problem of uncertainty propagation (Townsend 2000). Inspired by this, we propose a complex valued wave function $\Psi(m)$ describing the continuously valued multiple-state for a particular firm, in which $|\Psi(m)|^2 \in \mathbb{R}$ is the probability distribution of observing it at a multiple m, which would lead to a collapse of the wave function into a dirac delta function: $|\Psi(m')|^2 = \delta(m - m')$. In Quantum Mechanics, the time dependence of a position-state $\boldsymbol{r}$ is given by the Schrödinger (1926)-equation:

$$i\hbar \frac{\partial}{\partial t}\Psi(\boldsymbol{r}, t) = (\frac{-\hbar^2}{2\mu}\nabla^2 + V(\boldsymbol{r}, t))\Psi(\boldsymbol{r}, t) \tag{13}$$

where i is the imaginary unit and $\hbar$ is the reduced Planck constant.

Similarly, our quantum mechanical approach would require a definition of translational operators $\hat{O}_j$ for each of the j exogenous variables, working on the underlying variable-dependent wave function $\Psi(m, \boldsymbol{x}_t)$, possibly based on Hermitian operators related to observables. The boundary conditions of the system resemble the classical particle in a box problem. Hopefully, an adequate formulation would yield a mathematical framework that would facilitate statistical inference based on the empirical evidence obtained.

Analysis on Industries and Geographies

A reasonable hypothesis is that distinct industries and geographies are driven by distinct macroeconomic factors. Thus, we consider a bottom-up approach for understanding aggregate multiples based on the drivers of individual industries and geographies an interesting extension of our analyses. From an industry perspective, one could also explore the possibility of operating with sector-specific multiples that are more suitable and reflect underlying value creation, as business models and capital structures vary significantly across industries. Concerning geographies, country specific macroeconomic factors could be applied and thus, more accurate relationships may be derived. Expanding the dataset by incorporating all listed firms in the world would also be a possibility.

Leading Irrational Indicators

The data used for our analyses is structured data, available to most professional investors. Within this data, we were unable to find any significant leading indicators providing predictions of the multiple, as expected from an Efficient Market Hypothesis perspective. However, there is a vast landscape of semantic sentiment data in social media and other online services that may contain information not accounted for by the financial markets, but with latent effects on valuation - Keynes (1936)s "animal spirits". The problem is to synthesize the data to obtain accurate predictions. Big Data and machine learning provide tools that may be applicable: natural language processing algorithms is a cornerstone, clustering analyses and other unsupervised learning algorithms can be used to segment consumers, and regression tools would be used to obtain predictions with the resulting indicators. This is a vast field of study within Computer Science, and was regarded out of scope early in the process. See Pang et al. (2008) and Russell & Norvig (2003) for elaboration.

Advanced Parametric Distribution Methods

More flexible parametric models may be developed for describing the distribution of multiples. In Chapter 5, the possibilities of implementing a skewed generalized t distribution (Theodossiou 1998) or a generalized hyperbolic distribution (Barndorff-Nielsen 1977) were discussed, in order to better capture the dynamics of multiples distributions over time. With five parameters, the optimization problem becomes less tractable, and the risk of overfitting becomes prominent. The distribution may also be modeled using a composition of distributions, e.g., using extreme value theory or a generalized Pareto distribution for the $m > 30x$-tails and a bell-shaped parametric distribution for companies with multiples $< 30x$. It is natural to set boundary restrictions, e.g., by requiring a continuous or continuous and smooth transition. Moreover, spline or kernel based approaches, convolution formulations and methods involving transformations into new spaces are all imaginable, but tractability of the likelihood maximization becomes a pressing issue.

Long-memory Models

Tsay (2010) describes a long memory model with the following properties:

$$(1 - B)^d x_t = w_t \tag{14}$$

Long-memory models are capable of capturing non-exponential decay of autocorrelation, which turns out to be present in many time series models, both economic and physical. Usage of long-memory models is prevalent when time-series are near-integrated, as may be hypothesized to be the case for median EV/EBITDA. For our analyses, we deemed AR modeling of residuals to be sufficient, but long-memory models may be appropriate for a more profound understanding of the autocorrelation.

GARCH Volatility Modeling

Heteroskedasticity is often prevalent in time series, and has been shown for several of the models we are considering. Aiming to model this explicitly, it is natural to use a Generalized AutoRegressive Conditional Heteroskedasticity model (Alexander 2008*a*). We propose an ARMA(n,m)-GARCH(p,q) description, with the ARMA model:

$$y_t = \alpha + \sum_{i=1}^{n} \gamma_i y_{t-i} + \boldsymbol{\beta}_t \cdot \boldsymbol{x}_t + \eta_t + \sum_{j=1}^{m} b_j \eta_{t-i} \tag{15}$$

in which the error process η_t is normally distributed with conditional variance σ_t^2 given by the GARCH model:

$$\sigma_t^2 = \bar{\sigma}^2 + \sum_{i=1}^{p} c_i \mu_{t-i}^2 + \sum_{j=1}^{q} d_j \sigma_{t-j}^2 \tag{16}$$

The estimation of these models is usually relatively straight-forward based on MLE.

However, GARCH models are primarily used for high-frequecy financial data, because volatility clustering effects tend to be mitigated for long term series (Alexander 2008*a*), and the implementation would be most relevant in relation to other multiple time granularities.

Modeling Tail Distributions with New Independent Variables

In section 5.1, we conducted a subset regression on the principal components in *qPCA*, from which some new variables emerged. In Chapter 5, the underlying variables based on a subset regression associated with *Ap* were used to model the distribution tails and moments, ignoring other factors that might have more explanatory power. If the model is extended in this way, the complexity would increase, but it might pay of e.g., giving increased R^2 for the other principal components from *qPCA* or better accuracy for the ν-parameter in the generalized log Student-t distribution.

Estimation of Frequency Approach Likelihood Function

In Section 5.3, a likelihood function was derivated, whose maximization would yield the optimal set of α_b and $\boldsymbol{\beta}_b$-parameters:

$$\mathscr{L}(\boldsymbol{\beta}_{1:B}, \alpha_{1:B}; \boldsymbol{x}_{1:T}) = \sum_{t=1}^{T} |\Omega_t| \log \gamma_t - \sum_{t=1}^{T} \sum_{b=1}^{B} |\Omega_{bt}| \log(1 + e^{\alpha + \boldsymbol{\beta}_b \cdot \boldsymbol{x}_t}) \tag{5.23}$$

For our version of the optimization, the problem comprised 250 individual parameters, and we were unable to obtain optimized coefficients. However, the expression may be formulated into something with fewer parameters by applying appropriate constraints, e.g., by setting the parameters $\beta(b)$ to be a fixed continuous function. With suited optimization algorithms, the problem might become tractable even in its raw form, either exactly or by a heuristic.

Aiming to gain an a better approximation to the function, it is natural to include the bucket allocation problem in the likelihood optimization:

$$L(\boldsymbol{\beta}_{1:B}, \alpha_{1:B}, B, l_{1:B}, u_{1:B}; \boldsymbol{x}_{1:T}) = \prod_{t=1}^{T} \prod_{\omega \in \Omega_t} P(M_t = \omega) \tag{17}$$

but this would dramatically increase the search space of solutions. As with Equation 5.23, also this optimazion problem may become more tractable by applying constraints and powerful algorithms.

Regression based on Stationarity

The Johansen test eigenvalues define a cointegrating equation rendering the residuals maximally stationary. This corresponds to OLS, in which the RSS is minimized, and may be used to define a new type of regression, and may be interpreted the regression most able to capture cointegration effects. However, this is speculative, and merely an idea for further analyses.

Machine Learning

Some preliminary testing of machine learning models was conducted in Section 4.4, without notable findings. However, it was shown that the machine leaning models may be tuned to be at least as good as linear regression, and possibly significantly better on the index regression problem, by appropriately adjusting the model parameters and adopting the training data. Nevertheless, it is unlikely that the index regression problem is the area in which the applications are most relevant because the sample size is limited. Looking at the full sample distribution, $\Omega_{1:T}$, a substantial training data set emerges in the cross-section. A machine learning model, e.g., an ANN could perform well in predicting individual company development or other distributional aspects, related to the discussions in Chapter 5 and 6, in particular in the presence of additional data, such as classification data or unstructured data.

The Autocorrelated Continuous Distributional Residual Process, $\hat{\epsilon}(m)_h^s$

For an actual distribution Ω_t and a fitted log Student-t distribution $f(m)_t$, there are multiple ways of defining the distribution-residuals $\hat{\epsilon}(m)_t$. One approach may be to fit the MLE-Student-t distribution $f(m)_t^*$ to the observed values Ω_t, and to use the arithmetic differences to obtain a residual distribution: $\hat{\epsilon}(m)_t = f(m)_t - f(m)_t^*$. This approach might be criticized of putting too much emphasis on the Student-t distribution assumption, and in response to that critique, other approaches include using a bucket-based distribution $f(m)_t^b$ a fitted spline, definitions based on distributional parameters, distributional moments or quantiles. Independently of how the residual distribution is defined, it will adhere to the following property:

$$\int_0^\infty \hat{\epsilon}(m)_t dm = \int_0^\infty f(m)_t - f(m)_t^* dm = \int_0^\infty f(m)_t dm - \int_0^\infty f(m)_t^* dm = 0 \tag{18}$$

The historical distribution residuals $\hat{\epsilon}(m)_t$ will be used to generate new distribution residual processes for convolution with the VAR-residuals, but it is unlikely that the distribution residual process is a continuous white-noise process. This consideration will entail some definition of independence for these kinds of processes. Furthermore, some dynamic in the process must be postulated, obtaining a new set of independent continuous distribution residuals, from which individual simulation residuals can be drawn, generating a correlated residual process, $\hat{\epsilon}(m)_{T+1:H}^{1:S}$, that may be applied to the individual simulations, estimating the error of the distribution in a proper way.

Conjugate Pairs of AR(n) Characteristic Equation

Huerta & West (1999) and Tsay (2010) describe a method for identifying inherent cyclicalities in time series involving the conjugate pairs $a \pm ib$ of roots for the AR(n) characteristic equation. The conjugate pairs give rise to a stochastic period, by the following equation:

$$k = \frac{2\pi}{\cos^{-1}(a/\sqrt{a^2+b^2})} \tag{19}$$

k may be used to identify cyclic trends in the variables. The conjugate pairs and the corresponding period lengths k were found for the roots from Section 6.1, but were of no practical interest.

E.II CBOE Volatility Index vs. Global Economic Policy Uncertainty Index

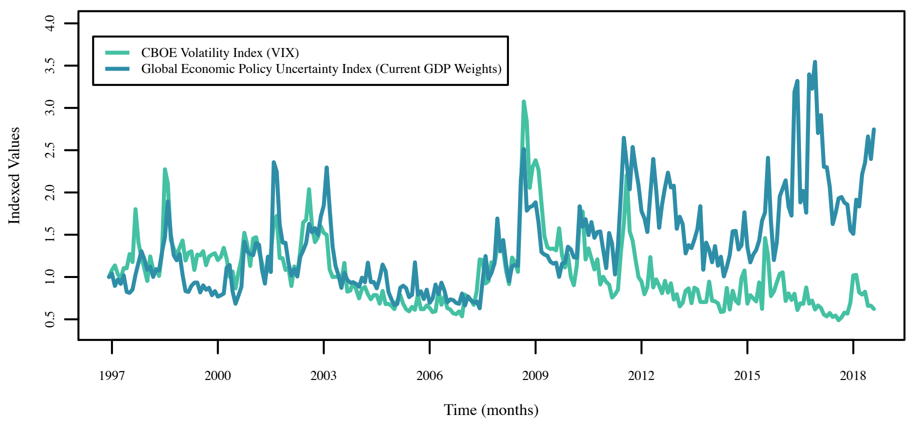


Figure 31: Indexed values for CBOE Volatility Index vs. Global Economic Policy Uncertainty Index

Explanation of Economic Policy Uncertainty Index developed by Baker et al. (2016):
"The Economic Policy Uncertainty Index (EPUI) reflects the relative frequency of own-country newspaper articles that contain a trio of terms pertaining to the economy (E), policy (P) and uncertainty (U). In other words, each monthly national EPU index value is proportional to the share of own-country newspaper articles that discuss economic policy uncertainty in that month."

Table 23: Correlation coefficients for relevant time intervals

	Time interval	
	2000-2012	2012-2018
Pearson Correlation	0.73	−0.08
Spearman Correlation	0.76	−0.04
Kendall Correlation	0.57	−0.02

The common Pearson correlation coefficient evaluates linear relationships, and Spearman evaluates monotonic relationship, while Kendall evaluates non-linear dependencies between variables. In Table 23 the results are shown for two time periods: 2000-2012 and 2012-2018. In conclusion, volatility and economic policy uncertainty have been correlated historically, while in recent years there is almost no correlation (slightly negative).