Empirical Business Valuation and Asset Pricing:

An Analysis from an Economic Perspective

Inaugural - Dissertation

zur

Erlangung der wirtschaftswissenschaftlichen Doktorwürde

des Fachbereichs Wirtschaftswissenschaften

der Philipps-Universität Marburg

eingereicht von:

Thomas Otto (M.Sc.)

aus Bad Nauheim

Erstgutachter/in:	Prof. Dr. Bernhard Nietert
Zweitgutachter/in:	Prof. Dr. Sascha H. Mölls
Einreichungstermin:	20. September 2019
Prüfungstermin:	18. Dezember 2019
Erscheinungsort:	Marburg
Hochschulkennziffer:	1180

Table of Contents

List of Abbreviations

a	arithmetic mean
APT	Arbitrage Pricing Theory
B	Book Value Of Common Equity
BRIC	Brazil, Russia, India, China
CAPM	Capital Asset Pricing Model
CART	Classification and Regression Trees
CE	Cash & Short Term Investments
D	Ordinary Cash Dividends
E	Earnings
e.g.	for example (abbreviation for exempli gratia)
EBIT	Earnings Before Interest And Taxes
EBITDA	Earnings Before Interest, Taxes & Depreciation
EBT	Earnings Before Taxes
etc.	and other similar things (abbreviation for et cetera)
f.	the following page
ff.	and the following pages
g	geometric mean
GAAP	Generally Accepted Accounting Principles
GI	Gross Income

h	harmonic mean
i.e.	that is (abbreviation for id est)
IAS	International Accounting Standards
IC	Invested Capital
ICB	Industry Classification Benchmark
IFRS	International Financial Reporting Standards
m	median
M	Model
NAME	Company Name
NAT	Nation
No.	number (abbreviation for numero)
OCF	Operating Cash Flow
OLS	Ordinary Least Squares
P	Market Capitalization – Fiscal Period End
p.	page
PCR	Principal Component Regression
PLS	Partial Least Squares
pp.	pages
Quantile (0.25)	Quantile Regression with $\tau = 0.25$
Quantile (0.50)	Quantile Regression with $\tau = 0.50$
Quantile (0.75)	Quantile Regression with $\tau = 0.75$

R&D	Research and Development
R^2	coefficient of determination
ros	ratio of sums
s.t.	subject to
SA	Net Sales Or Revenues
SRI	Socially Responsible Investments
TA	Total Assets
U.S.	United States
UK	United Kingdom
WLS	Weighted Least Squares
WPID	Worldscope Permanent I.D.

List of Symbols

$A_{i,j,t,s}$	characteristic j of asset i at time t in state s
$a_{j,t,s}$	characteristic j of the asset to be valued at time t in state s
$f(.)$ and $g(x)$	functions that determine the portfolio holdings constraints
i	index for assets (e.g., companies)
j	index for characteristics (e.g., accounting figures)
L	Lagrange function
m	number of characteristics
n	number of assets
p	index for norms
$P_{C,A_{C,j},x}$	(estimated) price of company C using accounting figure $A_{C,j}$ and multiple β_{x,A_j} with $x \in \{a, g, h, med, roa\}$
$P_{i,t}$	price of asset i at time t
$Prop(\bar{R})$	percentage of assets which have a Ratio (3.1) less than or equal to the value $\bar{R}$
q	index for dual norms, i.e., $\frac{1}{p} + \frac{1}{q} = 1$
$R_{i,t}$	return of asset i at time t
$R_{i,t} - r_t$	return differential of asset i at time t to the riskless rate
$ratio_{C_i,j}$	Ratio (3.1) for asset i based on the statistical method j
RM_k	risk measure$_k$

$s(.)$	function that integrates synergies into the buyer's/seller's decision problem
W_t	investors' wealth at time t
x	scaling factor
$\|x\|_p$	L_p-norm
y_{C_i}	price/return of asset i
$\hat{y}_{C_i,stat.meth._j}$	estimated price/return of asset i based on the statistical method j
$\hat{y}_{C_i,stat.meth_{ref}}$	estimated price/return of asset i based on the reference statistical method ref
y_i	price/return of asset i
$\hat{y}_i$	estimated price/return of asset i
β_0	intercept parameter
β_a	multiple based on the arithmetic mean
β_g	multiple based on the geometric mean
β_h	multiple based on the harmonic mean
β_j	regression coefficient of characteristic j
β_m	multiple based on the median
β_{ros}	multiple based on the ratio of sums
γ_i^+	Lagrange multiplier
γ_i^-	Lagrange multiplier
ε_i	error measure of asset i

ε_i^-	error measure of asset i in the case of an underestimation
ε_i^+	error measure of asset i in the case of an overestimation
ε_i^{log}	logarithmic error of asset i, i.e., $\varepsilon_i^{log} = \ln\left(\frac{\hat{y}_i}{y_i}\right)$
ε_i^{pct}	percentage error of asset i, i.e., $\varepsilon_i^{pct} = \frac{\hat{y}_i - y_i}{y_i}$
θ_j	multiple of characteristic j
λ_i^- and $N_{i,t}^-$	number of sales of asset i at time t
λ_i^- and $w_{i,t-1}^-$	portfolio weight of sales of asset i at time $t-1$
λ_i^+ and $N_{i,t}^+$	number of purchases of asset i at time t
λ_i^+ and $w_{i,t-1}^+$	portfolio weight of purchases of asset i at time $t-1$
μ_i^-	upper limit on underestimations for asset i
μ_i^+	upper limit on overestimations for asset i
ν_i^+	Lagrange multiplier
ν_i^-	Lagrange multiplier
τ	weight given to underestimations/quantile in quantile regressions
ω_i	weight given to asset i

List of Figures

List of Tables

Chapter I: Introduction

1 Introduction to the Problem

Common basis of all empirical accounting-based asset pricing models is their attempt to explain today's asset prices or returns with accounting characteristics that are observable today. Technically, empirical accounting-based asset pricing is implemented in the literature with a wide variety of statistical methods: regression approaches, method of multiples, and error measures, a fact that results in several problems.

<u>First problem</u>

Given that regression approaches, method of multiples, and error measures deal with empirical asset pricing, the multitude of conceptually different and non-connected approaches is puzzling and gives rise to two questions:

(i) If regression approaches, method of multiples, and error measures are applied empirically, they might lead to vastly different valuation results. Therefore, wouldn't it be useful to elaborate conceptual similarities and differences between these statistical methods and even find a superordinate category?

(ii) With respect to regression approaches, the existing literature uses just a small subset of possible statistical methods for empirical asset pricing, i.e., ordinary least squares, weighted least squares, or quantile regressions. Wouldn't it be rational to enlarge this subset of regression approaches by using other functions of the residuals, e.g., higher (and not first or second) order of absolute values of residuals or the maximum error?
With respect to the method of multiples, wouldn't it be useful to possess a pricing formula that can integrate different methods of computing means as well as using several accounting figures?

> With respect to error measures, wouldn't it be reasonable to have a pricing framework (= objective function) that is consistent with the error measure (= quality assessment).

Given these questions, the first objective of this thesis in Chapter II is to analyze which of the existing empirical asset pricing approaches are conceptually similar, i.e., can be summarized to a superordinate category and present statistical methods that can be considered as quasi-natural extensions to existing empirical asset pricing models.

Second problem

Based on this overview over empirical asset pricing models and the literature, it can be strongly assumed that the chosen factors (numbers and specific selection of explanatory variables) as well as the specific statistical method used (e.g., ordinary least squares regression, quantile regression) have an important influence on the explanatory power of an empirical analysis. Since the only concern of the majority of existing papers is the previously mentioned explanatory power, they can be regarded as dealing with statistical significance of factors/specific statistical methods, whereas the economic relevance is far less analyzed.

Since price differences are the decisive aspect of valuation models in practice and not statistical significance, analyzing their economic significance is essential and inevitable. Nobody will pay a higher price for a company just because a specific valuation method produces a high out-of-sample R^2. Moreover, business decisions should not be based only on whether a p-value passes a specific threshold because statistical significance (p-value) cannot measure the size of an effect or the importance of a result.

Therefore, it is the second objective of this thesis in Chapter III to analyze the economic significance of different factors/specific statistical methods.

Third problem

If, however, different factors/specific statistical methods lead to economically significant differences in value, a model-selection criterion is needed that is based on economic instead of statistical criteria. While arbitrage theory provides a general guideline for economic model evaluation for theoretical asset pricing models (i.e., prices must be a linear function of their future cash flows), empirical asset pricing models do not rely on present values of cash flows, but on assumed relations between accounting characteristics/factor returns and company prices/returns. For that reason, no theoretical guidelines regarding the components of the model exist. In particular, there are neither hints regarding the number and type of explanatory variables nor the specific statistical approach.

Given this high need for an economic model evaluation criterion, the third objective of this thesis in Chapter IV is to develop an economic model evaluation criterion and come up with an economic ranking of different empirical models.

<u>Fourth problem</u>

From the perspective of asset pricing theory such a model evaluation criterion is superfluous because the correct business valuation model is clear: the present value of future cash flows. Practically, forecasts of the future are difficult and, in particular, the determination of discount factors proves problematic. Therefore, it might be better to use a theoretically less convincing but easier applicable model—e.g., use of accounting characteristics—instead of a theoretically superior but inadequately implementable model—present value. However, the superior practicability of existing accounting-based valuations comes at a high cost: a relatively weak foundation in asset pricing theory:

(i) Multiples

Multiples essentially argue that similar accounting characteristics should result in similar prices.

Problems from the perspective of asset pricing theory: While such a valuation statement is intuitive, it is not backed up by asset pricing/arbitrage theory that states: Identical cash flow streams must possess identical prices. In other words, there are three differences between multiples and arbitrage theory. First, accounting characteristics are considered instead of cash flow streams. Second, similar instead of identical positions are examined. Third, one accounting characteristic is regarded as enough to characterize a company completely.

(ii) Implementing discounted cash flow models with the help of accounting characteristics
In literature, there are discounted cash flow models that use (functions of) accounting figures in order to express cash flows, the horizon value and/or the discount rate.
Problems from the perspective of asset pricing theory: Irrespective of the specific inclusion of the accounting characteristics in the discounted cash flow models, they can only serve as an approximation, i.e., the models contain assumptions that do not generally hold in reality.

(iii) Empirical accounting-based approaches
Empirical accounting-based approaches explain stock prices with the help of accounting characteristics.
Problems from the perspective of asset pricing theory: These empirical accounting-based approaches belong to the field of value relevance studies and, thus, are only interested in statistical significance of accounting characteristics, but not economic significance, i.e., they do not derive pricing statements. In principle, the regression coefficients of value relevance studies can also be used to obtain business values. However, valuation differences between different regression approaches are huge and these models have a weak economic backing when contrasted with the economic principle.

All these problems underline the trade-off between asset pricing rigor and practicability of models: Present value models are theoretically superior, but their practical implementation in form of constant discount rates and horizon models is far from economically convincing. Accounting-based models are characterized by less asset pricing theory rigor, however, can be implemented without sacrificing much of their theoretical basis. Obtaining better asset pricing models, hence, means either improve the implementation of present value models or the theoretical foundations of accounting-based models. Two reasons favor the improvement of the asset pricing foundation of empirical accounting-based models. On the one hand, the accounting literature so far has not fully exploited the asset pricing potential of accounting-based valuation models: It can be increased visibly without sacrificing practicability. On the other hand, purely empirical models always create a justification problem: Who would pay a higher price for a company because sales multiples result in higher prices than earnings multiples? Who would pay a higher price for a company because a lower discount rate for earnings is used? Who would pay a higher price for a company because an empirical estimation procedure, which possesses a higher R^2, recommends a higher price than other empirical estimation procedures?

Therefore, it is the fourth objective of this thesis in Chapter V to connect the practicability of accounting-based valuation models with the theoretical rigor of asset pricing theory.

2 Organization of the Thesis

The remainder of this thesis is organized as follows: Chapter II gives an overview of existing empirical assets pricing approaches and condenses them into a common framework. Chapter III analyzes the economic significance of valuation differences of different regression approaches. Chapter IV defines an economic model evaluation criterion and applies it to selected empirical asset pricing models. Finally, based on these results, Chapter V develops an accounting-based empirical business valuation model: the optimize-the-price approach.

Chapter II: Overview of Existing Approaches and Elaboration of a Common Framework

1 Introduction

Empirical asset pricing models have one thing in common: They try to explain today's asset prices (or returns) with value drivers that are observable today. However, at this point the common ground ends because three conceptually different approaches—we call them different categories of statistical methods—exist: regression approaches, method of multiples, and error measures.

Regression approaches are the most prominent in the academic literature and are associated with factor models/predictability of stock returns or value relevance studies in accounting. However, regression approaches seem to get more diverse in recent times. On the one hand, the number and specification of factors is increased (e.g., the overview in Harvey/Liu/Zhu (2016)), on the other hand, different regression methods like quantile regressions (e.g., Allen/Singh/Powell (2011)), weighted least squares regressions (e.g., Easton/Sommers (2003)), or generalized least squares regressions (e.g., Lewellen/Nagel/Shanken (2010)) are employed. Finally, regression approaches stand unconnected to alternative pricing approaches like the method of multiples and error measures.

The method of multiples is extremely popular in business valuation because of its easy implementation (see Coenenberg/Schultze (2002), p. 697). Perhaps due to its popularity, many approaches exist ranging from arithmetic, geometric, and harmonic mean to median as well as the ratio of averages (see Agrrawal/Borgman/Clark/Strong (2010), pp. 12 ff.). Even pricing results for several accounting figures (e.g., EBIT and sales multiples) are averaged (see Beatty/Riffe/Thompson (1999), p. 26 and Cheng/McNamara (2000), p. 352).—All these approaches are primarily unrelated and, in particular, not connected with regression analysis.

Finally, error measures are used to provide a suitable criterion for assessing the results of empirical pricing approaches (see Dittmann/Maug (2008), pp. 1 ff.). However, they suffer to some degree from consistency issues: the quality assessment (error measure) does not fit to the framework of the pricing model because both use different objective functions. Moreover, they are not related to regression analysis.

Given that all three categories of statistical methods, regression approaches, method of multiples, and error measures circle around the same problem, namely empirical asset pricing, the multitude of conceptually different and non-connected approaches is puzzling and gives rise to two questions:

(i) If regression approaches, method of multiples, and error measures are applied empirically, they might lead to vastly different valuation results (e.g., Nietert/Otto (2018) for multiples and Chapter III for regression approaches). Wouldn't it then be useful to understand why valuation results are different or even identify superior statistical methods? In other words, wouldn't it be helpful to elaborate conceptual similarities and differences between statistical methods and even find a superordinate category?

(ii) Within regression approaches, the literature uses just a small subset of possible statistical methods for empirical asset pricing, i.e., ordinary least squares, weighted least squares, or quantile regressions. "Why not minimize some other function of the residuals", as Wooldridge in his famous textbook on econometrics (see Wooldridge (2012), p. 31) asks? Potential candidates would be higher (and not first or second) order of absolute values of residuals or the maximum error.

With respect to the method of multiples, wouldn't it be useful to possess a pricing formula that can integrate different methods of computing means as well as using several accounting figures?

With respect to error measures, wouldn't it be reasonable to have a pricing framework (= objective function) that is consistent with the error measure (= quality assessment).

Given these questions, the objectives of this Chapter II are (i) to analyze which of the existing empirical asset pricing approaches are conceptually similar, i.e., can be summarized to a superordinate category; (ii) to present statistical methods that can be considered as quasi-natural extensions to existing empirical asset pricing models.

To achieve these objectives, a two-step procedure is followed. In a first step, the three categories of statistical methods (regression approaches, method of multiples, and error measures) are analyzed with regard to whether they can be aggregated to a general statistical method, i.e., to one superordinate category. In a second step, the general statistical method is used to check whether other statistical methods can be subsumed under the general statistical method. If this is the case, a quasi-natural extension to existing empirical asset pricing models will be found.

The results of this chapter can be summarized as follows: First, regression approaches and error measures can be combined to one superordinate category because they (can be formulated to) minimize functions of residuals. The method of multiples, however, remains a separated category since the multiple, the factor loading, is not determined from an optimization problem. Second, quasi-natural extensions of existing

(i) regression approaches combine higher orders of residuals (L_p-norms) with different penalties on over- and underestimations (quantile regressions), and dependence structures between error terms of different observations (generalized least squares regressions).

(ii) methods of multiples compute prices as weighted average of prices arising from different methods of computing means using different accounting figures.

(iii) error measures allow for the computation of factor loadings from an objective function that is consistent with the error measure (= quality assessment) used.

Compared to the literature, this chapter provides two contributions: First, it analyzes empirical asset pricing models across categories. The literature analyzes factor models/predictability of stock returns, value relevance, and multiples completely separately even though all three categories deal with asset pricing. Moreover, the literature in both factor models/predictability of stock returns (e.g., Harvey/Liu/Zhu (2016) and Appendix 4) and value relevance (e.g., Mölls/Strauß (2007) and Appendix 4) is almost exclusively interested in discussion of factors, but does not touch the issue of different statistical methods. The only exception is Allen/Singh/Powell (2011) who examine the integration of quantile regression into the Fama/French (1993) three-factor model. Regarding the analysis of factors, the literature is strictly empirical. The most advanced paper by Barillas/Shanken (2018) employs an empirical nesting approach. If, e.g., the CAPM and the Fama/French (1993) three factor model were equivalent regarding the intercept (alpha is equal to zero), the CAPM would be favored because it was the more parsimonious model. We, on the other hand, propose a theoretical nesting approach by showing that different statistical methods can be nested into a superordinate category.—Such a theoretical approach is able to compare, e.g., quantile regressions, generalized least squares regressions, and multiples, which cannot be done using the nesting approach of Barillas/Shanken (2018).

Second, by proposing quasi-natural extensions to empirical asset pricing models, it partially provides an answer to Wooldridge who asks: “Why not minimize some other function of the residuals?” (see Wooldridge (2012), p. 31).

With respect to regression models the empirical asset pricing literature so far has extended ordinary least squares regressions regarding quantile regressions (e.g., Allen/Singh/Powell (2011), p. 176), weighted least squares regressions (e.g., Easton/Sommers (2003), Formula (2), p. 42), and generalized least squares regressions (e.g., Lewellen/Nagel/Shanken (2010), p. 183). This chapter adds higher orders of residuals (L_p-norms) on an isolated basis and together with different penalties on over- and underestimations (quantile regressions) and dependence structures between error terms of different observations (generalized least squares regressions).

With respect to multiples Beatty/Riffe/Thompson (1999) and Cheng/McNamara (2000) discuss how prices arising from the use of different accounting figures can be weighted to obtain a final price. Prices arising from different methods of computing the mean, however, are not examined.—This chapter closes this gap and shows how to combine price estimates arising from different methods of computing means with those using several accounting figures.

In the context of error terms, the so far missing computation of consistent factor loadings is provided, consistent in the sense that the objective function becomes consistent with the error measure (= quality assessment).

The remainder of this Chapter II is organized as follows: Section 2 gives an overview of existing empirical asset pricing approaches and proposes quasi-natural extensions. In Section 3 the superordinate category of the presented statistical methods is elaborated. Section 4 concludes this chapter.

2 Overview of Existing Approaches[1]

2.1 Regression Approaches

2.1.1 Basic Principle

Regression approaches attempt to explain a dependent variable (asset prices or returns) with the help of explanatory variables (= factors) best possible. Best possible means that the unknown model parameters (= regression coefficients, i.e., factor loadings) are determined so that a function of residuals (objective func-

[1] Note that model calibration—see, e.g., the introduction to this approach in Cochrane (2005)—does not belong to the class of empirical asset pricing models. Model calibration implements theoretical models empirically, whereas empirical asset pricing models directly refer to the empirical relation without taking the detour over a theoretical model.

tion) is minimized. The objective function of the regression serves at the same time as quality assessment of the model's explanatory power. Therefore, regression approaches are consistent since regression coefficients and quality assessment are derived from the same objective function. Existing regression approaches only differ in the way how this objective function is defined.

Regression analysis can be conducted with cross-sectional as well as with time series data both with prices and returns. Characteristic examples for cross-sectional regressions of prices are value relevance studies (e.g., Easton/Harris (1991) and Appendix 4 for a rather comprehensive overview) that explain stock prices with the help of earnings per share and book value per share. Characteristic examples for time series regressions of returns are the three-, four-, five-, and six-factor models by Fama/French (1993), Carhart (1997), Fama/French (2015), and Fama/French (2018) which explain stock returns with the help of market risk, size, value, momentum etc.—Again use Appendix 4 for an overview.— In the financial empirical asset pricing literature, time series regressions play the leading role, whereas cross-sectional regressions dominate in the empirical accounting literature.

Formally, regression approaches work as follows:

(2.1)

$$y_i = \beta_0 + \sum_{j=1}^{m} \beta_j \cdot A_{i,j} - \varepsilon_i$$

where y_i denotes observation i of the dependent variable, $A_{i,j}$ characteristic j of observation i, β_j the regression coefficient of characteristic j, β_0 the intercept parameter, and ε_i the residual of asset i.

The relation of (2.1) to all four approaches (cross section of prices as well as returns and time series of prices as well as returns) is described in more detail:

Prices of companies in a cross section are the standard case of value relevance studies in accounting (e.g., the survey paper of Mölls/Strauß (2007), p. 958 or Appendix 4). In this standard case the explanatory relation reads, e.g.,

$$P_{i,t} = \beta_{0,t} + \beta_{EBIT,t} \cdot EBIT_{i,t} + \cdots - \varepsilon_{i,t}$$

for all companies i in the sample at time t.

In other words, the variable y_i in (2.1) is equal to the price of company i at time t. The variables $A_{i,j}$ are accounting figures, e.g., $EBIT_{i,t}$ (but not their growth rates) at time t.

Returns of companies in the cross section can be found in two different strands of the literature: on the one hand, in value relevance studies in accounting, e.g., the survey paper Mölls/Strauß (2007), Kothari/Zimmerman (1995), Harris/Muller (1999) or Appendix 4; on the other hand, in the second step of the two-pass regressions of Black/Jensen/Scholes (1972) and Fama/MacBeth (1973).

The regression equation that describes the relation between dependent and independent variables reads

– in accounting (e.g., Harris/Muller (1999), Formula 2, p. 299 or Kothari/ Zimmerman (1995), Formula 2, p. 159)

$$R_{i,t} or \ \frac{Market\ Value_{i,t}}{Market\ Value_{i,t-1}} = \beta_{0,t} + \beta_{earnings,t} \cdot \frac{Earnings_{i,t}}{Market\ Value_{i,t-1}} + \cdots - \varepsilon_{i,t}$$

– in two-pass regressions (Black/Jensen/Scholes (1972) and Fama/MacBeth (1973))

$$R_{i,t} - r_t = \gamma_{0,t} + \gamma_{Factor_1,t} \cdot Beta_{Factor_1,i,t} + \cdots - \varepsilon_{i,t}$$

for all companies i in the sample at time t.

In other words, the variable y_i in (2.1) is equal to the return or the return differential of company i at time t to the riskless rate. The variables $A_{i,j}$ are either relative accounting figures or beta factors determined from time series regressions (first step of the two-pass regression) at time t.

Prices of companies in time series can be found in, e.g., Kothari/Zimmerman (1995), p. 175 and Appendix 4 and look like

$$P_{i,\tau} = \beta_{0,i} + \beta_{1,i} \cdot earnings_{i,\tau} + \cdots - \varepsilon_{i,\tau}$$

for company i at all points in time τ.

In other words, the variable y_i in (2.1) is equal to the price of a company i at time τ. The variable $A_{i,j}$ denote accounting figure j of company i at different points in time τ.

Returns of companies in time series can be found in the first step of the two-pass regressions of Black/Jensen/Scholes (1972) and Fama/MacBeth (1973):

$$R_{i,\tau} - r_\tau = \beta_i + \beta_{Factor_1,i,\tau} \cdot R_{Factor_1,\tau} + \cdots - \varepsilon_{i,\tau}$$

for company i at all points in time τ.

In other words, the variable y_i in (2.1) is equal to the return or return differential to the riskless rate of company i at time τ. The variable $A_{i,j}$ denote returns of factor j at different points in time τ. In Fama/French (2015), Formula 4, p. 2, e.g., factor returns are specified as return of a portfolio (of small stocks, stocks with high book-to-market ratio etc.) minus the return on another portfolio (of big stocks, stocks with high book-to-market ratio etc.).

To keep the complexity of the notation in check we will use the regression Formula (2.1) in the remainder of this Chapter II and suppress, in addition, the time subscript t. That way, we can use one formula that holds for cross section of prices as well as returns and time series of prices as well as returns. However, when we will analyze the implicit economic assumptions of empirical asset pricing approaches in Chapter IV, we will thoroughly distinguish between cross section of prices as well as returns and time series of prices as well as returns.

2.1.2 Existing Approaches in the Literature

Our criterion used to structure existing regression approaches is their respective objective function. For each objective function one characteristic paper is cited. Hence, this section does not aim at providing an overview of all papers in empirical asset pricing that use regression approaches. Instead it prepares for the analysis of differences and similarities in regression approaches and, thus, the identification of a superordinate category.

2.1.2.1 Ordinary Least Squares Regression

Ordinary least squares regressions have the highest degree of dissemination in empirical asset pricing studies both the field of value relevance and factor models/predictability (see Appendix 4).

Ordinary least squares regression minimizes the sum of the squared residuals with respect to regression coefficients. In this connection, ordinary least squares regression does not penalize underestimations differently from overestimations. Formally,

(2.2)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \sum_{i=1}^{n} \left(\left[\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j \right] - y_i \right)^2$$

where residual ε_i reads

(2.3)

$$\varepsilon_i = \left(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j}_{=\widehat{y_i}} \right) - y_i$$

and $\widehat{y_i}$ denotes the estimated price/return of observation i.

Alternatively,

(2.4)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \sqrt{\sum_{i=1}^{n} \left(\left[\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j \right] - y_i \right)^2}$$

can be used as objective function. Since the square root function is monotonously increasing in its arguments, Problems (2.2) and (2.4) deliver the same minimum for $\beta_0, \beta_1, \ldots, \beta_m$. Only the value of the objective functions differs (see Boyd/Vandenberghe (2009), p. 131).

2.1.2.2 Weighted Least Squares Regression

Weighted least squares regressions are the second most used statistical method (see Appendix 4). Weighted least squares regression multiplies all inputs (observations y_i and characteristics $A_{i,j}$) by a weight ω_i before they enter the optimization problem. A typical application of weighted least squares regression is the case of heteroscedastic error terms, i.e., error terms are uncorrelated but not identically distributed (see Rao/Toutenburg/Shalabh/Heumann (2008), p. 156). If the weights are correctly specified, weighted least squares regression results in lower standard errors than the traditional ordinary least squares regression (see Wooldridge (2012), pp. 280 ff.).

Then, the sum of the squared residuals is minimized with respect to regression coefficients. In this connection, weighted least squares regression does not penalize underestimations differently from overestimations. Formally,

(2.5)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \sum_{i=1}^{n} \left(\left[\omega_i \beta_0 + \sum_{j=1}^{m} \omega_i A_{i,j} \beta_j \right] - \omega_i y_i \right)^2$$

or, alternatively,

(2.6)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \sqrt{\sum_{i=1}^{n} \left(\left[\omega_i \beta_0 + \sum_{j=1}^{m} \omega_i A_{i,j} \beta_j \right] - \omega_i y_i \right)^2}$$

where residual ε_i reads

(2.7)

$$\omega_i \varepsilon_i = \left(\underbrace{\omega_i \beta_0 + \sum_{j=1}^{m} \omega_i A_{i,j} \beta_j}_{=\omega_i \hat{y}_i} \right) - \omega_i y_i$$

Note that there are two different approaches of weighting in the literature: Easton/Sommers (2003), Formula (2), p. 42 use $P_{i,t}$ as weight ω_i, whereas, e.g., Brown/Lo/Lys (1999), Formula (15), p. 105 employ $P_{i,t-1}$.

2.1.2.3 Quantile Regression

Quantile regressions are only employed by Allen/Singh/Powell (2011) in the field of empirical asset pricing. Quantile regressions are able to penalize overestimations differently from underestimations by using a weight τ for underestimations and $1 - \tau$ for overestimations where it holds $0 < \tau < 1$ (see Koenker (2005), p. 5). For $\tau < 0.50$, overestimations are penalized more strongly in the objective function, for $\tau > 0.50$ underestimations. For $\tau = 0.50$ (median regression), over- and underestimations are treated equally. The cases $\tau = 0$ and $\tau = 1$ are excluded since otherwise there would be no trade-off between over- and underestimations. Consequently, $\hat{y}_i$ would be set arbitrarily low for penalized overestimations so that never an overestimation in the sense $\hat{y}_i > y_i$ occurs (high for penalized underestimations so that never an underestimation in the sense $\hat{y}_i < y_i$ occurs) and an infinite number of admissible solutions for the regression coefficients would result. The higher τ, the more optimization reduces underestimations due to their high penalty and increases overestimations due to their low penalty, a fact that leads to higher estimated prices/returns $\hat{y}_i$. Therefore, the weighting factor τ corresponds to the estimated quantile of the dependent variable because the sample is divided in a way such that τ percent are below and $(1 - \tau)$ percent are above the estimated price/return $\hat{y}_i$ (see Koenker (2005), p. 7).

Formally, quantile regression can be written (see Koenker (2005), Formula (1.19), p. 10)

(2.8)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} (1-\tau)\cdot \underbrace{\sum_{\substack{i=1\\ y_i<\beta_0+\sum_{j=1}^m A_{i,j}\beta_j}}^{n} \left| \left[\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j\right] - y_i \right|}_{overestimation}$$

$$+\tau\cdot \underbrace{\sum_{\substack{i=1\\ y_i\geq\beta_0+\sum_{j=1}^m A_{i,j}\beta_j}}^{n} \left| \left[\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j\right] - y_i \right|}_{underestimation}$$

where residual ε_i reads

(2.9)

$$\varepsilon_i = \left(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j}_{=\hat{y}_i}\right) - y_i$$

Several quantiles τ can be considered. Therefore, a family of regression lines is available for interpretation and, thus, a more complete view of the relationship between the variables is obtained compared to ordinary least squares regression. Ordinary least squares regression only looks at one regression line (see Koenker (2005), pp. 17, 25) because its regression line is based on the conditional mean in the sense of $E\{y_i|A_1,\ldots,A_m\} = \beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j$.

2.1.2.4 Generalized Least Squares Regression

Generalized least squares regressions in the field of empirical asset pricing were pioneered by Sami/Zhou (2004) and massively advocated by Lewellen/Nagel/Shanken (2010). However, overall they are still used sparingly (see Appendix 4).

The intuition behind generalized least squares regressions is that error terms might not be uncorrelated and identically distributed. Instead, error terms of different observations might depend on each other.

This dependence structure can be characterized as follows: the variance of error terms reads $\sigma^2 W$ where W is a positive definite matrix instead of an identity matrix as in the classical regression model (see Rao/Toutenburg/Shalabh/Heumann (2008), p. 143). Special cases of this dependence structure are: (i) heteroscedasticity where error terms are uncorrelated but possess a different variance; W then becomes a diagonal matrix and generalized least squares regression simplifies to weighted least squares regression (see Rao/Toutenburg/Shalabh/Heumann (2008), p. 156); (ii) serial correlation where error terms exhibit serial correlation but have an identical variance; W then becomes a matrix that contains the autocorrelation coefficients of first and higher order (see Rao/Toutenburg/Shalabh/Heumann (2008), p. 159).—Recall, regression approaches can be formulated as cross-sectional and time series regressions (see Section 2.1.1) and, hence, both heteroscedasticity and serial correlation might be relevant.

Formally, dependence structures of error terms can be captured by transforming the original variables of the regression model (see Rao/Toutenburg/Shalabh/Heumann (2008), pp. 143 f., 151):

(2.10)

$$\begin{pmatrix} y_1^* \\ \vdots \\ y_n^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} A_{1,1}^* & \cdots & A_{1,m}^* \\ \vdots & \ddots & \vdots \\ A_{n,1}^* & \cdots & A_{n,m}^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{pmatrix}$$

$$\begin{pmatrix} \varepsilon_1^* \\ \vdots \\ \varepsilon_n^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

where ω is a square and regular matrix and it holds $W^{-1} = \omega\omega$.

Generalized least squares regression minimizes the sum of the squared residuals with respect to regression coefficients. In this connection, it does not penalize underestimations differently from overestimations. Formally,

(2.11)

$$\min_{\beta_0,\beta_1,\dots,\beta_m} \sum_{i=1}^{n} \left(\left[\beta_0 + \sum_{j=1}^{m} A_{i,j}^{*} \beta_j \right] - y_i^{*} \right)^2$$

or, alternatively,

(2.12)

$$\min_{\beta_0,\beta_1,\dots,\beta_m} \sqrt{\sum_{i=1}^{n} \left(\left[\beta_0 + \sum_{j=1}^{m} A_{i,j}^{*} \beta_j \right] - y_i^{*} \right)^2}$$

where residual ε_i^* reads

(2.13)

$$\varepsilon_i^* = \left(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}^{*} \beta_j}_{=\widehat{y_i^*}} \right) - y_i^*$$

2.1.3 Possible Extensions to Regression Approaches

“Possible extensions to regression models” could in principle comprise any statistical method of the literature. Therefore, we must be more precise and confine ourselves to what we would like to call quasi-natural extensions. Quasi-natural extensions integrate separated features of the statistical methods introduced in Section 2.1.2 into one general statistical method, but do not modify the core of regression approaches in that they can be subsumed under the minimization of a function of residuals.

Separated features of the statistical methods in Section 2.1.2 are:

(i) The order of the function of residuals; so far: absolute values (first order) and quadratic functions (second order).

(ii) The different weighting of over- and underestimations; so far: quantile regressions.

Based on the separated features (i) to (ii) we suggest the following possible (quasi-natural) extensions.

2.1.3.1 L_p-Norms as Objective Functions

The intuition behind this extension is that not only absolute values (first order) or quadratic functions (second order) of residuals could be minimized. Instead, the absolute value of higher orders of residuals could also be considered. The higher the order, the less (more) influence have small (large) residuals on the objective function and vice versa. Therefore, in the limiting case, where the order of the function approaches infinity, only the maximum residual becomes relevant.

Technically, the order of the residuals' function is captured by means of L_p-norms, which are defined as (see Collatz (1964), pp. 132 ff.):

(2.14)

$$\|\varepsilon\|_p = \left[\sum_{i=1}^{n} |\varepsilon_i|^p\right]^{\frac{1}{p}}$$

The limiting case $p = \infty$ results from:

(2.15)

$$\|\varepsilon\|_\infty = \max\left\{|\varepsilon_i| \,\middle|\, i = 1, \ldots, n\right\}$$

Minimizing L_p-norms means that the p^{th} root of the sum of the p^{th} power of the absolute values of residuals is minimized. In this connection, L_p-norms do not penalize underestimations differently from overestimations. Formally,

(2.16)

$$\min_{\beta_0,\beta_1,\dots,\beta_m}\left(\sum_{i=1}^{n}\left|\left[\beta_0+\sum_{j=1}^{m}A_{i,j}\beta_j\right]-y_i\right|^p\right)^{\frac{1}{p}}$$

where residual ε_i reads

(2.17)

$$\varepsilon_i=\left(\underbrace{\beta_0+\sum_{j=1}^{m}A_{i,j}\beta_j}_{=\hat{y}_i}\right)-y_i$$

2.1.3.2 Different Weighting of Over- and Underestimations and L_p-Norms as Objective Functions

This extension is motivated by a combination of L_p-norms with the idea of an asymmetric penalty on over- and underestimations.

Then, the minimization of the L_p-norm with different weighting of over- and underestimations means that the p^{th} root of the sum of the p^{th} power of the absolute values of the residuals is minimized. In this connection a weight τ for underestimations and $1-\tau$ for overestimations is used with $0<\tau<1$. Formally,

(2.18)

$$\min_{\beta_0,\beta_1,\dots,\beta_m}\left[(1-\tau)\cdot\underbrace{\sum_{\substack{i=1\\ y_i<\beta_0+\sum_{j=1}^{m}A_{i,j}\beta_j}}^{n}\left|\left[\beta_0+\sum_{j=1}^{m}A_{i,j}\beta_j\right]-y_i\right|^p}_{overestimation}\right.$$
$$\left.+\tau\cdot\underbrace{\sum_{\substack{i=1\\ y_i\geq\beta_0+\sum_{j=1}^{m}A_{i,j}\beta_j}}^{n}\left|\left[\beta_0+\sum_{j=1}^{m}A_{i,j}\beta_j\right]-y_i\right|^p}_{underestimation}\right]^{\frac{1}{p}}$$

where residual ε_i reads

(2.19)

$$\varepsilon_i = \left(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j}_{=\hat{y}_i}\right) - y_i$$

2.1.3.3 Combining the Extensions from Sections 2.1.3.1 to 2.1.3.2

Combing L_p-norms with different penalties on over- and underestimations (quantile regressions) and integrating them into the generalized least squares regressions framework (2.11) to (2.13) yields

(2.20)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \left[(1-\tau) \cdot \underbrace{\sum_{\substack{i=1 \\ y_i^* < \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j}}^{n} \left| \left[\beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j \right] - y_i^* \right|^p}_{overestimation} \right.$$
$$\left. + \tau \cdot \underbrace{\sum_{\substack{i=1 \\ y_i^* \geq \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j}}^{n} \left| \left[\beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j \right] - y_i^* \right|^p}_{underestimation} \right]^{\frac{1}{p}}$$

with

(2.21)

$$\begin{pmatrix} y_1^* \\ \vdots \\ y_n^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} A_{1,1}^* & \cdots & A_{1,m}^* \\ \vdots & \ddots & \vdots \\ A_{n,1}^* & \cdots & A_{n,m}^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{pmatrix}$$

$$\begin{pmatrix} \varepsilon_1^* \\ \vdots \\ \varepsilon_n^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

where ω is a regular an symmetric matrix and it holds $W^{-1} = \omega\omega$

and where residual ε_i^* reads

(2.22)

$$\varepsilon_i^* = \Bigg(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j}_{=\widehat{y_i^*}}\Bigg) - y_i^*$$

2.2 Method of Multiples

2.2.1 Basic Principle

A multiple is defined as company price divided by the accounting figure of interest, e.g., the price earnings ratio. It is obtained from an average of a group of comparable companies (see Peemöller/Meister/Beckmann (2002), pp. 197 f.).—This is the reason why the method of multiples is often called comparable company approach (see Peemöller/Meister/Beckmann (2002), pp. 197 f.).—Therefore, the method of multiples does not include the optimization of an objective function and, hence, there is no associated quality assessment of the explanatory power of multiples.

The importance of this method is made clear by the fact that 99 percent of analysts' reports rely on multiple-based valuations (see Asquith/Mickhail/Au (2005), p. 257). Moreover, the comparable company approach is used in nearly all initial public offerings (see Beckmann/Meister/Meitner (2003), pp. 103 f.).

Technically, multiples could be applied to both prices and returns. However, given the scope of their application—valuation of corporations—using multiples with returns is not used in both the literature and the industry.

Similarly, multiples can, in principle, be applied to cross-sectional or time series data. However, the defining feature of multiples is the comparison with similar companies and not with the history of the company under consideration. Hence, multiples are solely associated with a cross-sectional analysis.

Pricing by multiplies means that a multiple is multiplied by the corresponding accounting figure at time t of the valuation object to determine its price at time t.

Such a procedure implies three things:

(i) Only positive multiples can be interpreted in economic terms. Negative multiples would revert the ordering of companies in that companies with a, e.g., higher positive EBIT or sales would be regarded as inferior to companies with smaller EBIT or sales.

(ii) Only positive accounting figures for both the company to be valued and the group of comparable companies will yield to meaningful economic interpretations. A negative accounting figure combined with a positive multiple results in a negative price.

(iii) Combinations of negative multiples with negative accounting numbers lead to completely implausible results: Companies with negative EBITs would realize higher prices than companies with positive EBITs.

2.2.2 Existing Approaches in the Literature

Principally, approaches with one and those with several multiples exist.

2.2.2.1 Approaches with One Multiple

There are five different ways of estimating multiples from comparable companies (see Agrrawal/Borgman/Clark/Strong (2010), p. 12 ff.): the arithmetic mean β_a (e.g., Baker/Rubak (1999) or Liu/Nissim/Thomas (2002)), the geometric mean β_g (e.g., Kim/Ritter (1999)), the harmonic mean β_h (e.g., Baker/Rubak (1999) or Liu/Nissim/Thomas (2002)), the median β_m(e.g., Alford (1992), Cheng/McNamara (2000), or Kim/Ritter (1999)), and the ratio of averages β_{roa} (e.g., Beatty/Riffe/Thompson (1999)).

2.2.2.1.1 Arithmetic Mean

Using the arithmetic mean, the multiple is determined as

(2.23)

$$\beta_a = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i}{A_i}\right)$$

where $\beta_a > 0$ by construction.

2.2.2.1.2 Geometric Mean

Using the geometric mean, the multiple is determined as

(2.24)

$$\beta_g = \left(\prod_{i=1}^{n}\left(\frac{y_i}{A_i}\right)\right)^{\frac{1}{n}} = \exp\left\{\frac{1}{n}\sum_{i=1}^{n}\ln\left(\frac{y_i}{A_i}\right)\right\}$$

where $\beta_g > 0$ by construction.

2.2.2.1.3 Harmonic Mean

Using the harmonic mean, the multiple is determined as

(2.25)

$$\beta_h = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\left(\left(\frac{y_i}{A_i}\right)^{-1}\right)}$$

where $\beta_h > 0$ by construction.

2.2.2.1.4 Median

Using the median, the multiple is determined as

(2.26)

$$\beta_{med} = \mathrm{Median}\left(\frac{y_i}{A_i}\right)$$

where $\beta_{med} > 0$ by construction.

2.2.2.1.5 Ratio of Averages

Using the ratio of averages, the multiple is determined as

(2.27)

$$\beta_{roa} = \frac{\frac{1}{n}\sum_{i=1}^{n} y_i}{\frac{1}{n}\sum_{i=1}^{n} A_i}$$

where $\beta_{roa} > 0$ by construction.

Therefore, the ratio of averages deviates from the computational procedures of the other multiples: It computes averages first and then computes a ratio. The other multiples calculate the ratio in the first step and average over these ratios in the second step.

2.2.2.2 Approaches with Several Multiples

Existing models that use several multiples define the estimated prices as the arithmetic mean of the separate single-factor price estimates (see Beatty/ Riffe/Thompson (1999), p. 26 and Cheng/McNamara (2000), p. 352). In this connection, it is averaged over prices estimated with the help of different accounting figures $A_{i,j}$ (e.g., EBIT and sales) that use the same method of computing means (e.g., arithmetic means).

Formally, the price of company C to be valued reads using its m accounting figures $A_{C,j}$ and multiples β_{x,A_j} with $x \in \{a, g, h, med, roa\}$:

(2.28)

$$\begin{gathered} P_{C,A_{C,1}} = \beta_{x,A_1} \cdot A_{C,1} \\ \vdots \\ P_{C,A_{C,m}} = \beta_{x,A_m} \cdot A_{C,m} \end{gathered}$$

where $P_{C,A_{C,j}}$ is the (estimated) price of company C using accounting figure $A_{C,j}$.

Then, the final price of company C can be computed as

(2.29)

$$P_C = \frac{1}{m}\sum_{j=1}^{m} P_{C,A_{C,j}} = \frac{1}{m}\sum_{j=1}^{m} \beta_{x,A_j} \cdot A_{C,j}$$

where m is the number of different accounting figures (like EBIT, sales etc.).

2.2.3 Possible Extensions to the Method of Multiples

“Possible extensions to the method of multiples” should follow the guidelines of quasi-natural extensions developed in Section 2.1.3. Quasi-natural extensions do not modify the core of multiples, i.e., that pricing by multiplies means that a multiple (i.e., company price divided by the accounting figure of interest, e.g., the price earnings ratio) is multiplied by the corresponding accounting figure of the valuation object.

However, possible extensions do not just average over different prices computed with the help of different accounting figures using equally weighting, but also

(i) average over different methods to compute means.—Nietert/Otto (2018) show that different methods to compute means results in huge (computed) price differences.

(ii) average with an arbitrary weighting scheme.

2.2.3.1 Averaging over Prices Arising from Different Methods of Computing Means

Formally, the price of company C to be valued reads using its accounting figure $A_{C,j}$ and multiple β_{x,A_j} with $x \in \{a, g, h, med, roa\}$

(2.30)

$$P_{C,A_{C,j},a} = \beta_{a,A_j} \cdot A_{C,j}$$
$$P_{C,A_{C,j},g} = \beta_{g,A_j} \cdot A_{C,j}$$
$$P_{C,A_{C,j},h} = \beta_{h,A_j} \cdot A_{C,j}$$

$$P_{C,A_{C,j},med} = \beta_{med,A_j} \cdot A_{C,j}$$
$$P_{C,A_{C,j},roa} = \beta_{roa,A_j} \cdot A_{C,j}$$

Then, the final price of the company can be computed as

(2.31)

$$P_C = \frac{1}{5} \sum_{x \in \{a,g,h,med,roa\}} P_{C,A_{C,j},x} = \frac{1}{5} \sum_{x \in \{a,g,h,med,roa\}} \beta_{x,A_j} \cdot A_{C,j}$$

or with an arbitrary weighting scheme $\omega_{C,A_{C,j},x}$ for the different methods of computing means

(2.32)

$$P_C = \sum_{x \in \{a,g,h,med,roa\}} \omega_{C,A_{C,j},x} \cdot P_{C,A_{C,j},x} = \sum_{x \in \{a,g,h,med,roa\}} \omega_{C,A_{C,j},x} \cdot \beta_{x,A_j} \cdot A_{C,j}$$

where $\sum_{x \in \{a,g,h,med,roa\}} \omega_{C,A_{C,j},x} = 1$.

2.2.3.2 Averaging over Prices Arising from Different Accounting Figures and Different Methods of Computing Means

The most general form of the method of multiples combines averaging over prices arising from using different accounting figures (Section 2.2.2.2) with those arising from different methods of computing means using arbitrary weighting schemes (Section 2.2.3.1). Then, it is obtained

– Prices from different accounting figures with arbitrary weighting scheme (Section 2.2.2.2)

  (2.29)

$$P_C = \sum_{j=1}^{m} \omega_{C,A_{C,j},x} \cdot P_{C,A_{C,j},x} = \sum_{j=1}^{m} \omega_{C,A_{C,j},x} \cdot \beta_{x,A_j} \cdot A_{C,j}$$

– Prices from different methods computing means with arbitrary weighting scheme (Section 2.2.3.1)

(2.32)

$$P_C = \sum_{x \in \{a,g,h,med,roa\}} \omega_{C,A_{C,j},x} \cdot P_{C,A_{C,j},x} = \sum_{x \in \{a,g,h,med,roa\}} \omega_{C,A_{C,j},x} \cdot \beta_{x,A_j} \cdot A_{C,j}$$

Hence, it finally holds

(2.33)

$$\begin{aligned} P_C &= \sum_{x \in \{a,g,h,med,roa\}} \sum_{j=1}^{m} \omega_{C,A_{C,j},x} \cdot P_{C,A_{C,j},x} \\ &= \sum_{x \in \{a,g,h,med,roa\}} \sum_{j=1}^{m} \omega_{C,A_{C,j},x} \cdot \beta_{x,A_j} \cdot A_{C,j} \end{aligned}$$

where $\sum_{x \in \{a,g,h,med,roa\}} \sum_{j=1}^{m} \omega_{C,A_{C,j},x} = 1$.

2.3 Error Measures

2.3.1 Basic Principle

Error measures serve to evaluate and compare different model results that are estimated with the help of regression approaches or the method of multiples (e.g., Dittmann/Maug (2008)). As such, error measures do not determine factor loadings but provide just a quality assessment of explanatory power. Consequently, the use of error measures often creates an inconsistency since factor loadings and quality assessment are derived from different objective functions.

Since error measures evaluate the outcome of regression approaches or the method of multiples, they can be applied to both cross-sectional and time series data as well as prices and returns. Due to missing factor loadings, however, error measure cannot price assets.

2.3.2 Existing Approaches in the Literature

Error measures proceed as follows. In a first step, errors are calculated for each asset i. In a second step, these errors are expressed as either percentage or logarithmic errors (e.g., Dittmann/Maug (2008)). In a third step, the resulting distribution of percentage or logarithmic errors is evaluated, e.g., based on descriptive statistics (e.g., Dittmann/Maug (2008)). In other words, error measures do not prescribe a certain descriptive statistic as evaluation criterion. Instead decision makers possess full flexibility as to which descriptive statistic they regard as relevant.—From that perspective, one can begin to fathom a certain connection to regression analysis because ordinary least squares regressions use the variance of errors as descriptive statistics.

2.3.2.1 Percentage Error

The percentage error ε_i^{pct} is defined as the difference between the estimated price (return) $\hat{y}_i$ and the actual price (return) y_i divided by the actual price (return) y_i (see Dittmann/Maug (2008), p. 6):

(2.34)

$$\varepsilon_i^{pct} = \frac{\hat{y}_i - y_i}{y_i}$$

2.3.2.2 Logarithmic Error

The logarithmic error ε_i^{log} is defined as the natural logarithm of the quotient of the estimated price (return) $\hat{y}_i$ and the actual price (return) y_i (see Dittmann/Maug (2008), p. 6):

(2.35)

$$\varepsilon_i^{log} = \ln\left(\frac{\hat{y}_i}{y_i}\right)$$

2.3.3 Possible Extensions to Error Measures

"Possible extensions to error measures" should follow the guidelines of quasi-natural extensions developed in Section 2.1.3, i.e., do not modify the core of error measures: the use of percentage or logarithmic errors.

In this connection, quasi-natural extensions rest upon two ideas. Frist, the consistent use of objective functions and error measures is recommended. This means that an error measure should be minimized to obtain factor loadings so that the objective function from which factor loadings are obtained and the quality assessment coincide. The determination of factor loadings in turn means that pricing will become possible. Second, a meaningful descriptive statistic might be selected.

2.3.3.1 L_p-Norms of Percentage Errors

Since the percentage error is defined as $\frac{\hat{y}_i - y_i}{y_i} = \frac{\varepsilon_i}{y_i}$, each relation between explanatory variables A_k and dependent variable y_i fits the idea of a percentage error when the error is additively connected to the explanatory variables, i.e.,

(2.36)

$$y_i = f\left(A_1, \dots, A_j, \dots, A_m\right) - \varepsilon_i$$

and

(2.37)

$$\hat{y}_i = f\left(A_1, \dots, A_j, \dots, A_m\right)$$

This can be seen as follows:

(2.38)

$$\frac{\hat{y}_i - y_i}{y_i} = \frac{f\left(A_1, \dots, A_j, \dots, A_m\right) - \left(f\left(A_1, \dots, A_j, \dots, A_m\right) - \varepsilon_i\right)}{y_i} = \frac{\varepsilon_i}{y_i}$$

Thus (2.38) $\frac{\varepsilon_i}{y_i}$ is identical to ε^{pct}.

However, an application argument favors a more specific, i.e., linear model because the determination of factor loadings is vastly simplified in a linear model.

Hence, it is specified

(2.39)

$$y_i = \beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j - \varepsilon_i$$

and

(2.40)

$$\hat{y}_i = \beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j$$

To be able to derive factor loadings, the asset-specific percentage errors must be aggregated to make statements about the size of total mispricing. Then, it becomes, however, necessary avoiding that positive and negative deviations compensate each other. Combining both requirements means L_p-norms of percentage errors $\|\varepsilon^{pct}\|_p$ must be used.

Based on the specification (2.40), factor loadings β_j are then determined by minimizing L_p-norms of percentage errors $\|\varepsilon^{pct}\|_p$, i.e.,

- if underestimations are not penalized differently from overestimations

  (2.41)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \left[\sum_{i=1}^{n} \left| \frac{\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j - y_i}{y_i} \right|^p \right]^{\frac{1}{p}}$$

- if underestimations are penalized differently from overestimations

  (2.42)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m} \left[(1-\tau) \cdot \underbrace{\sum_{\substack{i=1 \\ y_i < \beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j}}^{n} \left| \frac{\beta_0 + \sum_{j=1}^{m} A_{i,j}\beta_j - y_i}{y_i} \right|^p}_{overestimation} \right.$$

$$+\tau\cdot \underbrace{\sum_{\substack{i=1 \\ y_i \geq \beta_0+\sum_{j=1}^{m} A_{i,j}\beta_j}}^{n} \left|\frac{\beta_0+\sum_{j=1}^{m} A_{i,j}\beta_j - y_i}{y_i}\right|^p}_{underestimation} \Bigg]^{\frac{1}{p}}$$

2.3.3.2 L_p-Norms of Logarithmic Errors

Using logarithmic errors means that over- and underestimations are not treated equally because the natural logarithm is a concave function that weights/penalizes underestimations more strongly than overestimations. From that perspective, there will be always an asymmetric penalty irrespective of whether an additional penalty term is supposed on over- and underestimations (quantile regressions).

With this general remark in mind, logarithmic errors can be approached in a similar vein as percentage errors: Since the logarithmic error is defined as $\ln\left(\frac{\hat{y}_i}{y_i}\right) = \varepsilon_i$, each model fits logarithmic errors when the error is an exponential function that is multiplicatively connected with the explanatory variables A_k, i.e.,

(2.43)

$$y_i = f\left(A_1, \dots, A_j, \dots, A_m\right) \cdot \exp\{-\varepsilon_i\}$$

and

(2.44)

$$\hat{y}_i = f\left(A_1, \dots, A_j, \dots, A_m\right)$$

This can be seen as follows:

(2.45)

$$\ln\left(\frac{\hat{y}_i}{y_i}\right) = \ln\left(\frac{f\left(A_1, \dots, A_j, \dots, A_m\right)}{f\left(A_1, \dots, A_j, \dots, A_m\right) \cdot \exp\{-\varepsilon_i\}}\right) = \ln\left(\frac{1}{\exp\{-\varepsilon_i\}}\right)$$
$$= \ln(1) - \ln(\exp\{-\varepsilon_i\}) = \varepsilon_i$$

Thus (2.45) ε_i is identical to ε^{log}.

However, an application argument favors a more specific, i.e., exponential model: the determination of factor loadings is vastly simplified if a linear model—after taking logarithm—is used.

Hence, it is specified

(2.46)

$$y_i = \exp(\beta_0) \cdot \prod_{j=1}^{m} \left(A_{i,j}\right)^{\beta_j} \cdot \exp(-\varepsilon_i)$$

and

(2.47)

$$\hat{y}_i = \exp(\beta_0) \cdot \prod_{j=1}^{m} \left(A_{i,j}\right)^{\beta_j}$$

To be able to derive factor loadings, the asset-specific logarithmic errors must be aggregated to make statements about the size of total mispricing. For that reason, L_p-norms of logarithmic errors $\|\varepsilon^{log}\|_p$ are used as objective function.

Based on the specification (2.47), factor loadings β_j are then determined by minimizing L_p-norms of logarithmic errors $\|\varepsilon^{log}\|_p$, i.e.,

- if underestimations are not penalized directly differently (but only indirectly by means of the logarithm) from overestimations

  (2.48)

$$\min_{\beta_0,\beta_1,\dots,\beta_m} \left[\sum_{i=1}^{n} \left|\ln\left(\frac{\hat{y}_i}{y_i}\right)\right|^p\right]^{\frac{1}{p}}$$

$$= \min_{\beta_0,\beta_1,\dots,\beta_m} \left[\sum_{i=1}^{n} |\ln(\hat{y}_i) - \ln(y_i)|^p\right]^{\frac{1}{p}}$$

$$= \min_{\beta_0,\beta_1,\dots,\beta_m} \left[\sum_{i=1}^{n} \left|\ln\left(\exp(\beta_0) \cdot \prod_{j=1}^{m} \left(A_{i,j}\right)^{\beta_j}\right) - \ln(y_i)\right|^p\right]^{\frac{1}{p}}$$

  i.e.,

(2.49)

$$\min_{\beta_0,\beta_1,\dots,\beta_m}\left[\sum_{i=1}^{n}\left|\beta_0+\sum_{j=1}^{m}\beta_j\ln\left(A_{i,j}\right)-\ln(y_i)\right|^p\right]^{\frac{1}{p}}$$

– if underestimations are penalized directly differently (and indirectly by means of the logarithm) from overestimations

(2.50)

$$\min_{\beta_0,\beta_1,\dots,\beta_m}\left[(1-\tau)\cdot\underbrace{\sum_{\substack{i=1\\ \ln(y_i)<\beta_0+\sum_{j=1}^{m}\beta_j\ln\left(A_{i,j}\right)}}^{n}\left|\beta_0+\sum_{j=1}^{m}\beta_j\ln\left(A_{i,j}\right)-\ln(y_i)\right|^p}_{overestimation}\right.$$

$$\left.+\tau\cdot\underbrace{\sum_{\substack{i=1\\ \ln(y_i)\geq\beta_0+\sum_{j=1}^{m}\beta_j\ln\left(A_{i,j}\right)}}^{n}\left|\beta_0+\sum_{j=1}^{m}\beta_j\ln\left(A_{i,j}\right)-\ln(y_i)\right|^p}_{underestimation}\right]^{\frac{1}{p}}$$

3 Elaboration of a Common Framework for Empirical Asset Pricing Models

Taking the overview of empirical models developed in Section 2 as starting point, a common framework for empirical asset pricing models can be developed. In particular, it is determined which of the three different categories of statistical methods—regression approaches, method of multiples, and error measures—are conceptually similar, i.e., can be summarized to a superordinate category and which approaches are conceptually different.

3.1 Regression Approaches

3.1.1 Most General Regression Model

Problem

(2.20)

$$\min_{\beta_0,\beta_1,\ldots,\beta_m}\left[(1-\tau)\cdot\underbrace{\sum_{\substack{i=1\\ y_i^*<\beta_0+\sum_{j=1}^m A_{i,j}^*\beta_j}}^{n}\left|\left[\beta_0+\sum_{j=1}^{m}A_{i,j}^*\beta_j\right]-y_i^*\right|^p}_{overestimation}\right.$$

$$\left.+\tau\cdot\underbrace{\sum_{\substack{i=1\\ y_i^*\geq\beta_0+\sum_{j=1}^m A_{i,j}^*\beta_j}}^{n}\left|\left[\beta_0+\sum_{j=1}^{m}A_{i,j}^*\beta_j\right]-y_i^*\right|^p}_{underestimation}\right]^{\frac{1}{p}}$$

with

(2.21)

$$\begin{pmatrix} y_1^* \\ \vdots \\ y_n^* \end{pmatrix}=\underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv\omega}\cdot\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} A_{1,1}^* & \cdots & A_{1,m}^* \\ \vdots & \ddots & \vdots \\ A_{n,1}^* & \cdots & A_{n,m}^* \end{pmatrix}=\underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv\omega}\cdot\begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{pmatrix}$$

$$\begin{pmatrix} \varepsilon_1^* \\ \vdots \\ \varepsilon_n^* \end{pmatrix}=\underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv\omega}\cdot\begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

where ω is a regular an symmetric matrix and it holds $W^{-1}=\omega\omega$

and where residual ε_i reads

(2.22)

$$\varepsilon_i^* = \underbrace{\left(\beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j\right)}_{=\widehat{y_i^*}} - y_i^*$$

is the most general problem since it combines L_p-norms with different penalties on over- and underestimations (quantile regressions) with dependence structures between error terms (generalized least squares regressions).

All regression approaches—whether existing approaches or extensions—can be derived as special cases from (2.20).

3.1.2 Identifying Other Regression Models as Special Cases of the Most General Regression Model

3.1.2.1 Existing Approaches in the Literature

– Ordinary least squares regression

  $\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{identity matrix}$

  $p = 2$

  $\tau = 1$

  and summation from $i = 1$ to n without distinguishing between $y_i^* < \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j$ and $y_i^* \geq \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j$.

– Weighted least squares regression

  $\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{diagonal matrix}$

  $p = 2$

  $\tau = 1$

  and summation from $i = 1$ to n without distinguishing between $y_i^* < \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j$ and $y_i^* \geq \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j$.

– Quantile regression

$$\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{identity matrix}$$

$p = 1$

3.1.2.2 Possible Extensions to Regression Approaches

– L_p-norm

$$\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{identity matrix}$$

$\tau = 1$

and summation from $i = 1$ to n without distinguishing between $y_i^* < \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j$ and $y_i^* \geq \beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j$.

– L_p-norm with different weighting of over- and underestimations

$$\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{identity matrix}$$

3.1.3 An Alternative Formulation of the General Regression Model (2.20)

Problem (2.20) makes mispricing only implicitly visible by means of the objective function. If mispricing is to be made explicit, it is recommended introducing upper limits for over- and underestimations. This can be achieved as follows:

Based on (2.22) an overestimation can be identified as

$$\varepsilon_i^{*+} = \left(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}^* \beta_j}_{=\widehat{y_i^*}} \right) - y_i^* > 0$$

and an underestimation as

$$\varepsilon_i^{*-} = \left(\underbrace{\beta_0 + \sum_{j=1}^{m} A_{i,j}^{*}\beta_j}_{=\widehat{y_i^*}}\right) - y_i^* < 0$$

which implies

$$\varepsilon_i^{*+} > 0 => \varepsilon_i^{*-} = 0$$

$$\varepsilon_i^{*-} < 0 => \varepsilon_i^{*+} = 0$$

Now different upper bounds for over- and underestimations can be defined

(2.51)

$$\varepsilon_i^{*+} \leq \mu_i^+$$

$$\varepsilon_i^{*-} \geq -\mu_i^- \text{ or } -\varepsilon_i^{*-} \leq \mu_i^-$$

where $\mu_i^+ \geq 0$ and $\mu_i^- \geq 0$. Alternatively,

$$|\varepsilon_i^{*+}| \leq \mu_i^+$$

$$|\varepsilon_i^{*-}| \leq \mu_i^-$$

The upper limits on over- and underestimations allow reinterpreting problem (2.20) in a (slightly) more intuitive way: Factor loadings β_j are determined by minimizing the L_p-norms of residuals where the upper limits of over- and underestimations should be chosen as tight as possible. Finally, a scaling factor $x \geq 0$ can be integrated into the objective function because such a scaling factor does not change the outcome. With these modifications in mind, problem (2.20) can be re-formulated as:

(2.52)

$$\min_{\substack{\mu_1^+,\mu_1^-,\dots,\mu_n^+,\mu_n^-,\\ \beta_0,\beta_1,\dots,\beta_m}} x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}}$$

s.t.

(2.53)

overestimation: $\varepsilon_i^{*+} = \beta_0 + \sum_{j=1}^{m} A_{1,j}^{*}\beta_j - y_1^{*} > 0$

$$\varepsilon_1^{*+} \leq \mu_1^{+}$$
$$\vdots$$
$$\varepsilon_n^{*+} \leq \mu_n^{+}$$

underestimation: $\varepsilon_i^{*-} = \beta_0 + \sum_{j=1}^{m} A_{1,j}^{*}\beta_j - y_1^{*} < 0$

$$\varepsilon_1^{*-} \geq -\mu_1^{-} \text{ or } -\varepsilon_1^{*-} \leq \mu_1^{-}$$
$$\vdots$$
$$\varepsilon_n^{*-} \geq -\mu_n^{-} \text{ or } -\varepsilon_n^{*-} \leq \mu_n^{-}$$

(2.54)

$$\mu_1^{+} \geq 0, \mu_1^{-} \geq 0, \dots, \mu_n^{+} \geq 0, \mu_n^{-} \geq 0, \beta_0 \in \mathbb{R}, \beta_1 \in \mathbb{R}, \dots, \beta_m \in \mathbb{R}$$

where x is greater than zero and denotes a scaling factor and there is either an over- ($\varepsilon_i^{*+} > 0$) or an underestimation ($\varepsilon_i^{*-} < 0$) implying $\varepsilon_i^{*+} > 0 \Rightarrow \varepsilon_i^{*-} = 0$ and $\varepsilon_i^{*-} < 0 \Rightarrow \varepsilon_i^{*+} = 0$.

3.2 Method of Multiples

3.2.1 Most General Model of the Method of Multiples

The most general model of the method of multiples is

(2.33)

$$P_C = \sum_{x \in \{a,g,h,med,roa\}} \sum_{j=1}^{m} \omega_{C,A_{C,j},x} \cdot P_{C,A_{C,j},x}$$
$$= \sum_{x \in \{a,g,h,med,roa\}} \sum_{j=1}^{m} \omega_{C,A_{C,j},x} \cdot \beta_{x,A_j} \cdot A_{C,j}$$

3.2.2 Identifying Other Methods of Multiples as Special Cases of the Most General Model of the Method of Multiples

If there is just one method of computing means, (2.33) will simplify to (2.29), if there is just one accounting figure (2.33) simplifies to (2.32), and if there is just one accounting figure and one method of computing means, (2.33) will simplify to (2.28).

3.3 Error Measures

Both L_p-norms of percentage errors (Problem (2.42)) and of logarithmic errors (Problem (2.50)) can be written as special cases of the general regression model (2.20).

- L_p-norm of percentage errors
  $\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{identity matrix with } \omega_{i,i} = \frac{1}{P_i} \ \forall\, i = 1, \dots, n.$

- L_p-norm of logarithmic errors
  $\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix} = \text{identity matrix}$
  and variable transformation: $\ln(y_i)$ instead of y_i and $\ln(A_{i,j})$ instead of $A_{i,j}$.

3.4 Identification of the Superordinate Category

In connecting the method of multiples with regression approaches it can be observed that the core multiple Formula (2.33) leads to a structure that is formally similar to pricing in a regression context because $\omega_{C,A_{C,j},x} \cdot \beta_{x,A_j}$ describe the (general) factor loadings and $A_{C,j}$ the firm-specific factors. From that perspective, regression approaches and the method of multiples are comparable.

Regression and multiples can, however, be only subsumed under one superordinate category if their approach in determining factor loadings will be comparable as well. The method of multiples does not involve an optimization to determine multiples and, thus, can be understood as using an arbitrary objective function. In particular, the objective function of the general regression model (2.52) including its constraint (2.53) can be applied.—Again compatibility is obtained so far. However, multiples are determined based on averaging and require in addition a non-negativity constraint on factor loadings (see Equations (2.23) to (2.27)). These two constraints are absent with regression approaches.

Therefore, regression approaches and the method of multiples are structurally different and cannot be aggregated to one superordinate category and the method of multiples forms a category of its own.

Regression approaches and error measures can be subsumed under Problem (2.20) meaning they can be summarized under one superordinate category.

4 Conclusion

In empirical asset pricing three different categories of statistical methods—regression approaches, method of multiples, and error measures—are used. However, these categories give rise to vastly different empirical price/return estimates. Hence, two immediate questions arise: (i) Wouldn't it then be useful to understand why valuation results are different or even identify superior statistical methods? (ii) Wouldn't other statistical methods, which are not currently used by empirical asset pricing, like minimizing maximum error or generalized least squares regression lead to better pricing results?

Given these questions, the objectives of this Chapter II are (i) to analyze which of the existing empirical asset pricing approaches are conceptually similar, i.e., can be summarized to a superordinate category; (ii) to present statistical methods that can be considered as quasi-natural extensions to existing empirical asset pricing models.

The results of this chapter can be summarized as follows: First, regression approaches and error measures can be combined to one superordinate category because they (can be formulated to) minimize functions of residuals. The method of multiples, however, remains a separated category since the multiple—the factor loading—is not determined from an optimization problem.
Second, quasi-natural extensions of existing

(i) regression approaches combine higher orders of residuals (L_p-norms) with different penalties on over- and underestimations (quantile regressions), and dependence structures between error terms of different observations (generalized least squares regressions).

(ii) methods of multiples compute prices as weighted average of prices arising from different methods of computing means using different accounting figures.

(iii) error measures allow for the computation of factor loadings from an objective function that is consistent with the error measure (= quality assessment) used.

The practical implications of this chapter are twofold: First, it serves as an intermediate step towards the evaluation of the implicit economic assumptions of the empirical asset pricing approaches in Chapter IV. With the superordinate category identified, only the superordinate category needs to be analyzed with respect to its implicit economic assumptions. An examination of the whole variety of statistical methods is no longer needed. Second, its quasi-natural extensions to existing empirical asset pricing models have the potential of improving empirical asset pricing models.

Chapter III: Economic Significance of Valuation Differences of Different Regression Models

1 Introduction

The explanatory power of each empirical analysis depends on the chosen factors (numbers and specific selection of explanatory variables) as well as the specific statistical method used (e.g., ordinary least squares regression, quantile regression). The literature is aware of the importance of number and/or specification of factors. E.g., the literature survey of Harvey/Liu/Zhu (2016) lists 316 predictors for asset returns, Harvey/Liu (2019) even more than 400 factors (finance papers), Appendix 4 contains an overview of factors analyzed in value relevance studies (accounting papers). Less examined, but still adequately reflected in the literature is the question regarding the effect of specific statistical methods. Allen/ Singh/Powell (2011) raise the question how using quantile regression instead of ordinary least squares regression will change the explanatory power of the factors identified in Fama/French (1993). Brown/Lo/Lys (1999) and Easton/ Sommers (2003) express their concern about a scale effect that might bias price regressions and, hence, prefer weighted least squares over ordinary least squares regression. Lewellen/Nagel/Shanken (2010) recommend using generalized least squares regression to improve empirical models statistically. Finally, Barillas/Shanken (2018) employ an empirical nesting approach. If, e.g., the CAPM and the Fama/French (1993) three factor model were equivalent regarding the intercept (alpha is equal to zero), the CAPM would be favored because it was the more parsimonious model.—All these papers analyzing factors and specific statistical methods are concerned with explanatory power of the statistical methods why they can be regarded as dealing with statistical significance of factors/specific statistical methods.

Economic relevance of factors/specific statistical methods, on the other hand, is far less analyzed and, hence, understood. Economic significance regarding different numbers of factors/specific statistical methods comprises on the one hand the question how and not just whether (as with statistical significance) the choice of different numbers of factors and/or specific empirical models changes stock prices or returns; on the other hand the interplay between factors and specific statistical methods, i.e., whether some explanatory factors induce greater price changes when combined with specific statistical methods than other factors. The literature on socially responsible investments (SRI) e.g., Bauer/Koedijk/Otten (2005), Bollen (2007), Renneboog/Ter Horst/Zhang (2008), Hong/Kacperczyk (2009), Nofsinger/Varma (2014), and Ibikunle/Steffen (2017), indirectly addresses economic relevance of factors. It uses with the CAPM, Fama/French (1993), and Carhart (1997) three different empirical models to identify return differences between conventional and socially responsible investments to take the influence of different factors on returns into account. The literature on multiples is more explicit regarding economic relevance of factors and specific statistical methods: Beatty/Riffe/Thompson (1999), Cheng/McNamara (2000), and Schreiner (2007) average valuation results for several accounting figures (e.g., EBIT and sales multiples) because they are aware that different factors translate into different company prices. Nietert/Otto (2018) analyze valuation differences that arise from using different key statistics (e.g., EBIT or sales), the criterion of finding peers, and the method how multiples of comparable companies are aggregated (e.g., arithmetic or geometric average).

This less than desirable analysis of the economic significance of factors/specific statistical methods is somewhat puzzling: on the one hand, because price differences are the decisive aspect of valuation models in practice and not statistical significance. Nobody will pay a higher price for a company just because a specific valuation method produces a high out-of-sample R^2. On the other hand, because The American Statistical Association (2016) points out that business decisions should not be based only on whether a p-value passes a specific threshold since statistical significance (p-value) cannot measure the size of an effect or the importance of a result.

Therefore, it is the objective of this Chapter III to analyze the economic significance of different factors/specific statistical methods.

To achieve this objective, cross-sectional regression models with accounting figures as explanatory variables are used. More specifically, the factors from the value relevance and multiple literature are taken and combined with the standard statistical methods of the empirical asset pricing and value relevance literature, i.e., ordinary least squares, weighted least squares, and quantile regression. In addition, the role of statistical methods is analyzed with equal importance to the role of factors and the interplay between factors and statistical methods.—Regarding statistical significance the literature puts visibly more importance on factors than on statistical methods.

The results of this chapter can be summarized as follows:

First, economic significance regarding different factors/specific statistical methods addresses the question how and not just whether (as with statistical significance) the choice of different factors/specific statistical methods influences company prices/returns and consists of two components: "magnitude" and "similarity". "Magnitude" focuses on the size of differences between prices/returns that different factors/specific statistical methods produce. "Similarity" condenses the cumulative relative frequency distribution of price/return differences into one number and addresses the problem that moderate price/return differences do not necessarily mean similar empirical models.

Second, "magnitude" shows for our data basis, i.e., company prices in the cross section, that price differences are generally large. Only 13% of all factors/specific statistical methods belong to the best category (price differences of 10% or less). These price differences are primarily caused by specific statistical methods and not so much by factors.

Third, "similarity" applied to our data basis illustrates that nearly all factors/specific statistical methods are dissimilar where statistical methods are primarily responsible for this lack of similarity and factors play only a minor role.

This chapter makes the following contribution compared to the literature:

First, this chapter introduces a systematic analysis of economic significance of both factors and specific statistical methods. The literature on socially responsible investments only analyses a very limited number of factors (Fama/French (1993) and Carhart (1997) as extensions to the CAPM) and does not consider different regression models as in Allen/Singh/Powell (2011). The accounting literature (see Mölls/Strauß (2007) or Appendix 4 for a rather comprehensive overview) analyzes different regressions, but just with respect to statistical significance and not economic significance. Moreover, the interplay between factors and specific statistical methods is ignored. In this respect the literature on multiples (e.g., Nietert/Otto (2018)) is able to partially step in. Yet multiples can by construction not handle regressions. Hand/Coyne/Green/Zhang (2017) compare price estimates based on discounted cash flow and residual income approaches, i.e., focus on factors, but do not consider regressions. Finally, this chapter compares with Europe, U.S., and BRIC three regions. The literature on socially responsible investments (value relevance of accounting figures) focuses on individual countries to be able to elaborate the diversification disadvantage of socially responsible investments (effects of country-specific financial accounting rules).

Second, this chapter develops a test procedure regarding economic relevance of factors/specific statistical methods. The literature so far has: with Gibbons/Ross/Shanken (1989) a statistical test with respect to statistical significance of factors; with Barillas/Shanken (2018) an empirical nesting approach that allows to identify superior models, i.e., models that are equivalent regarding the intercept (alpha is equal to zero), but more parsimonious and, therefore, better, because they need less factors.

The remainder of this Chapter III is organized as follows: Section 2 outlines the design of the analysis. Section 3 describes data set and data cleaning. The results of the empirical analysis are contained in Section 4. Section 5 concludes this chapter.

2 Design of the Analysis

To be able to analyze the economic significance regarding different numbers of factors/specific statistical methods, a two-step procedure is required.—Note in this connection that we use the phrase "empirical model" as superordinate category that is decomposed into two components: factors and specific statistical methods (e.g., ordinary least squares regression).

In a first step, a (theoretical) evaluation criterion regarding economic significance must be developed. In a second step, factors/specific statistical methods must be selected to compute company prices/returns, thus creating the data basis for the application of the evaluation criterion.

2.1 Developing an Evaluation Criterion Regarding Economic Significance (First Step)

2.1.1 Definition of Economic Significance

Economic significance regarding different factors/specific statistical methods addresses the question how and not just whether (as with statistical significance) the choice of different factors/specific statistical methods influences company prices/returns.

It comprises clearly the magnitude of the price/return differences between different factors/specific statistical methods. Specifically, "magnitude" offers the following deeper insights: Are there few large (outliers) and otherwise small differences or are differences generally large? Which factors/specific statistical methods result in higher, which in lower prices/returns? Can differences be observed irrespective of industries, regions, and years?

While "magnitude" stresses the differences between factors/specific statistical methods, i.e., focuses on dissimilarities, it cannot capture adequately the flip side of differences, namely similarity. E.g., many small differences between two empirical models might lead to the conclusion that these models are not different. Not different does, however, not necessarily mean that these two empirical models are similar. In fact, the two empirical models could be not different, but also not similar. Therefore, “similarity” analyzes: Are there certain combinations of factors/specific statistical methods that are always similar and others that are always dissimilar? Is the degree of similarity between factors/specific statistical methods constant over various industries, regions, and years?

In summary, economic significance is defined to answer the following two questions:

(i) What is the magnitude of the price/return differences between different factors/specific statistical methods?

(ii) What factors/specific statistical methods are similar regarding their price/return differences?

2.1.2 General Requirements for an Evaluation Criterion Regarding Economic Significance

To be able to evaluate economic significance, i.e., to answer the questions regarding “magnitude” and “similarity”, an evaluation criterion must meet the following two requirements:

(i) All differences between factors/specific statistical methods must be judged simultaneously.

(ii) The sign of the differences between factors/specific statistical methods matters.

When analyzing “magnitude” both requirements (i) and (ii) advocate a direct access to differences and no aggregation because this would lead to a loss in information. Aggregated differences, namely, can neither distinguish between many small and few large differences (requirement (i)) nor identify factors/specific statistical methods that produce higher prices/returns than others (requirement (ii)). Therefore, e.g., differences' means or the often used statistical criteria (out-of-sample) R^2 (e.g., Campbell/Thompson (2008)) or generalized least squares R^2 (see Lewellen/Nagel/Shanken (2010), p. 183) will not work as evaluation criterion regarding economic significance.[2]

"Similarity" requires a slightly different treatment than "magnitude". On the one hand, detailed information on differences as captured by requirements (i) and (ii) is also needed for "similarity": Positive and negative differences must not be netted because this would result in a wrong picture of "similarity". Moreover, the distribution of differences—one big and many small versus many medium-sized differences—is important to judge "similarity". For both reasons again neither differences' means nor (out-of-sample) R^2 are good criteria to measure "similarity". On the other hand, differences must be evaluated in total and to that end possibly aggregated to make a statement on "similarity".

2.1.3 Derivation of an Evaluation Criterion Regarding Economic Significance

2.1.3.1 Common Basis for Measuring "Magnitude" and "Similarity"

Core of both "magnitude" and "similarity" is the difference between the estimated company prices/returns. Therefore, a formalization of the difference is the common basis for applying "magnitude" and "similarity" and, hence, for judging economic significance.

In this connection it is recommended normalizing differences since a large difference in combination with a large company price/return appears to be less problematic than in combination with a small company price/return. The current price/return of the company is a good choice as numeraire. Using the current

[2] Note in addition that also a technical aspect argues against the use of (out-of-sample) R^2. (Out-of-sample) R^2 relies on the variance of errors which will not be adequate if, e.g., quantile regression is considered: There, the sum of the absolute values of errors should be used as quality measure and not their variance.

price/return means that a uniform numeraire is used in all calculations and, therefore, comparability across all specific statistical methods is enabled. If the company price/return of one reference statistical method is used as numeraire, only results regarding different factors for this reference method can be compared. Comparisons between results of factors from statistical methods that do not contain the reference method cannot be made because the numeraire differs.

Formalizing these ideas, the following ratio can be defined for each company C_i:

(3.1)

$$ratio_{C_i,j} = \frac{\hat{y}_{C_i,stat.meth._j} - \hat{y}_{C_i,stat.meth_{ref}}}{y_{C_i}}$$

where $\hat{y}_{C_i,stat.meth._j}$ denotes the company price/return estimated based on the statistical method j, $\hat{y}_{C_i,stat.meth_{ref}}$ the company price/return estimated based on the reference statistical method ref, and y_{C_i} the current price/return of the company.

A $ratio_{C_i,j}$ of zero signifies no difference between company prices/returns calculated by statistical method j and the reference statistical method ref. On the other hand, a $ratio_{C_i,j}$ of, e.g., 1 means that the difference between the two estimated company prices/returns is as large as the current company price/return. In this respect, a $ratio_{C_i,j}$ of 1 can be regarded as very large difference.

Finally, note three aspects regarding Ratio (3.1). First, $ratio_{C_i,j}$ possesses a triangular structure, i.e., e.g., $\hat{P}_{C,Quant(q)} - \hat{P}_{C,WLS}$ ($\hat{P}_{C,WLS}$ is the price estimated with the help of the reference regression weighted least squares) returns the same result apart from the sign as $\hat{P}_{C,WLS} - \hat{P}_{C,Quant(q)}$ ($\hat{P}_{C,Quant(q)}$ is the price estimated with the help of the reference regression quantile regression). However, since we are interested in analyzing which factors/specific statistical methods produce high and which low prices/returns, we compute all ratios, not just those of the upper or lower triangle. Second, Ratio (3.1) is a (normalized) price/return difference and, thus, has some similarity to alpha. Yet recognize the completely different intention behind both measures: alpha[3] portrays the difference between estimated and actual prices/returns, whereas Ratio (3.1) depicts the difference between prices/returns estimated with the help of two different empirical models. Third, Ratio (3.1) just describes the difference between prices/returns of different empirical models. It has no connotation in the sense of "better/worse empirical model" since Ratio (3.1) neither measures against a price/return derived from a "true" pricing model nor tries to reproduce current prices/returns. For the same reason, phrases like "over-" or "underestimation" are not used.

2.1.3.2 "Magnitude"

To infer the evaluation criterion "magnitude" of economic significance from Ratio (3.1), the relative frequency distribution of Ratio (3.1) for all companies under consideration is computed. By using Ratio (3.1)'s relative frequency distribution, all information regarding size and sign of differences is provided without a loss of information and requirements (i) and (ii) of Section 2.1.2 are met. In addition, since one particular Ratio (3.1) refers to just one statistical method j/reference statistical method-combination, frequency distributions (for all companies) must be determined for all statistical method j/reference statistical method-combinations.

Therefore, a better overview of results regarding "magnitude" will be achieved if the relative frequency distribution is condensed to few classes:

$$0\% < \text{Ratio (3.1)} \leq 10\% \quad \text{and} \quad -10\% < \text{Ratio (3.1)} \leq 0\%$$

$$10\% < \text{Ratio (3.1)} \leq 50\% \quad \text{and} \quad -50\% < \text{Ratio (3.1)} \leq -10\%$$

[3] To be more precise, there are two different types of alphas in the literature: on the one hand, Jensen's (original) alpha and, as special case, the alpha in the analyses of socially responsible investments; on the other hand, the alpha in factor models/predictably approaches. In the latter type of models alphas should be zero because the model is tested in-sample and should fully explain stock returns. Jensen's (original) alpha as well as the alpha in the socially responsible investment-type of models is determined out-of-sample and used as measure of investment quality: Investors should seek positive alpha stocks.

$$50\% < \text{Ratio (3.1)} \leq 100\% \quad \text{and} \quad -100\% < \text{Ratio (3.1)} \leq -50\%$$

$$100\% < \text{Ratio (3.1)} \leq 200\% \quad \text{and} \quad -200\% < \text{Ratio (3.1)} \leq -100\%$$

$$200\% < \text{Ratio (3.1)} \leq 500\% \quad \text{and} \quad -500\% < \text{Ratio (3.1)} \leq -200\%$$

$$\text{Ratio (3.1)} > 500\% \quad \text{and} \quad \text{Ratio (3.1)} \leq -500\%$$

To finish the design of the criterion "magnitude" of economic significance, a bound must be defined that separates low and, thus, acceptable "magnitudes" from high (inacceptable) "magnitudes". Given that a Ratio (3.1) of 1 means that the difference between the two estimated company prices/returns is as large as the current company price/return, a size of 10% or less might be considered as an acceptable "magnitude". Absolute values of Ratio (3.1) that exceed 10% might be seen as too high.

2.1.3.3 "Similarity"

To infer the evaluation criterion "similarity" of economic significance from Ratio (3.1), start with the observation that two empirical models will be identical, i.e., perfectly similar, if they exhibit zero differences in prices/returns. The more differences occur, the more dissimilar models will be.

To implement this intuition formally, a two-step procedure is applied. In a first step, the relative frequency distribution is transformed into a cumulative relative frequency distribution. To do this, company-specific Ratios (3.1) are sorted by size and added on a percentage-weighted basis, i.e.,

(3.2)

$$Prop(\bar{R}) = \frac{\sum_{i=1}^{n} 1_{ratio_j \leq \bar{R}}\left(ratio_{C_i,j}\right)}{n}$$

where $1_A(z)$ denotes the value of the indicator function on a set A for variable z and n is the number of companies in the sample.

By construction, $Prop(\bar{R})$

(i) is monotonically increasing in $\bar{R}$ and lies in the interval [0%, 100%].

(ii) indicates the percentage of companies which have a Ratio (3.1) less than or equal to the value $\bar{R}$.

In a second step, the cumulative relative frequency distribution (3.2) is transformed into a measure of "similarity" with the help of the following procedure: In the extreme case of perfect similarity there are no differences between two empirical models. Then, $Prop(\bar{R})$ shows a cumulative relative frequency distribution function that is identical to a Dirac distribution, i.e., a function whose value is zero for Ratios (3.1) smaller than zero and one for Ratios (3.1) greater than or equal to zero. This can be read as: 0% of the companies have different prices/returns, and 100% of the companies have identical ones. Consequently, deviations from perfect similarity can be identified as the area between the Dirac distribution (= ideal case) and the cumulative relative frequency distribution (3.2).

Formally, this measure of "similarity" can be computed as follows: For Ratios (3.1) < 0 (company price/return of statistical method j is less than the company price/return of the reference statistical method ref), the cumulative relative frequency distribution function (3.2) is above the Dirac distribution (which has a value of 0).

For Ratios (3.1) > 0 (company price/return of statistical method j is greater than the company price/return of the reference statistical method ref), the cumulative relative frequency distribution function (3.2) is below the Dirac distribution (which has a value of 1). Therefore, both sub-areas are positive and can be added to determine the total area that reflects the (normalized) difference of company prices/returns estimated based on the statistical method j and the company prices/returns estimated based on the reference statistical method ref. As shown in Appendix 2, this area can be calculated as the arithmetic mean of the absolute values of Ratio (3.1) for all companies:

(3.3)

$$\frac{1}{n} \cdot \sum_{i=1}^{n} \left| ratio_{C_i,j} \right|$$

Area (3.3) is the evaluation criterion regarding “similarity” of economic significance. If Area (3.3) is small, company prices/returns estimated based on the statistical method j and the company prices/returns estimated based on the reference statistical method ref will be similar, otherwise dissimilar. In other words, the greater Area (3.3) is, the more dissimilar statistical method j and the reference statistical method ref will be.

To analyze all statistical method j/reference statistical method-combinations, a histogram of Areas (3.3) is created that summarizes the values of Area (3.3) for each statistical method j/reference statistical method-combination.

To finish the design of the criterion “similarity” of economic significance, a bound must be defined that separates small Areas (3.3) and, thus, similar factors/specific statistical methods from dissimilar ones. However, a direct interpretation of (3.3) proves impossible because (3.3) is primarily a formal measure and gives no direct intuition as to how similar two empirical models are: Assume (3.3) takes a value of 1. Does this mean that regression j is similar to the reference statistical method ref?

To develop an intuition regarding acceptable sizes of Area (3.3), an interpretation similar to the one of Ratio (3.1) would be helpful where differences were related to current prices/returns and, thus, gave rise to an intuitive upper bound.

Such a relation does not exist[4] for the arithmetic mean (3.3). However, with the help of the geometric mean such a relation can be established. Note that the geometric mean is a lower bound for the arithmetic mean, i.e., it holds

---

[4] The arithmetic mean $\frac{1}{n} \cdot \sum_{i=1}^{n} \left| ratio_{C_i,j} \right|$ reads in more detail $\frac{1}{n} \cdot \sum_{i=1}^{n} \frac{\left| difference_{C_i} \right|}{y_{C_i}}$.
As long as $y_{C_i} \geq 0$, it holds $\frac{\left| difference_{C_i} \right|}{y_{C_i}} \geq \frac{\left| difference_{C_i} \right|}{\sum_{i=1}^{n} y_{C_i}}$ and, thus,

$$\text{arithmetic mean}\left(\left|ratio_{C_i,j}\right|\right) \geq \text{geometric mean}\left(\left|ratio_{C_i,j}\right|\right)$$

The geometric mean of $\left|ratio_{C_i,j}\right|$, in turn, can be related to the geometric mean of prices/returns:

(3.4)

$$\begin{aligned}\text{geometric mean}\left(\left|ratio_{C_i,j}\right|\right) &= \sqrt[n]{\prod_{i=1}^{n}\left|ratio_{C_i,j}\right|} = \sqrt[n]{\prod_{i=1}^{n}\frac{\left|difference_{C_i}\right|}{y_{C_i}}} \\ &= \frac{\sqrt[n]{\prod_{i=1}^{n}\left|difference_{C_i}\right|}}{\sqrt[n]{\prod_{i=1}^{n} y_{C_i}}} \\ &= \frac{\text{geometric mean}\left(\left|difference_{C_i}\right|\right)}{\text{geometric mean}\left(y_{C_i}\right)}\end{aligned}$$

(3.4) relates average differences to average prices/returns. Hence, it is the desired economic intuition behind Area (3.3). Therefore, similar to the bound for the "magnitude" (Ratio (3.1)) a bound of 10% again seems to be reasonable.

$$\begin{aligned}\frac{1}{n}\cdot\sum_{i=1}^{n}\frac{\left|difference_{C_i}\right|}{y_{C_i}} &\geq \frac{1}{n}\cdot\frac{\sum_{i=1}^{n}\left|difference_{C_i}\right|}{\sum_{i=1}^{n} y_{C_i}} \\ &= \frac{1}{n}\cdot\frac{\frac{\sum_{i=1}^{n}\left|difference_{C_i}\right|}{n}}{\frac{\sum_{i=1}^{n} y_{C_i}}{n}} \\ &= \frac{1}{n}\cdot\frac{\text{arith. mean}\left(\left|difference_{C_i}\right|\right)}{\text{arith. mean}\left(y_{C_i}\right)}\end{aligned}$$

In other words, the arithmetic mean of differences divided by the arithmetic mean of company prices/returns (second term) is meaningful. Unfortunately, this ratio is additionally divided by the number of companies (first term). Hence, the lower bound for $\frac{1}{n}\cdot\sum_{i=1}^{n}\left|ratio_{C_i,j}\right|$ is so low so that it cannot provide a meaningful economic judgement.
One final question must be clarified: When is $y_{C_i} > 0$ given? If prices are considered, i.e., $y_{C_i} = P_{C_i}$, the positivity always holds since prices of companies with limited liability are by definition positive (only in insolvency do they equal zero). Returns, i.e., $y_{C_i} = R_{C_i,t,t+1}$, on the other hand can assume negative values. From that perspective, $\frac{\left|difference_{C_i}\right|}{y_{C_i}} \geq \frac{\left|difference_{C_i}\right|}{\sum_{i=1}^{n} y_{C_i}}$ cannot be guaranteed and will hold only if: $y_{C_i} > 0$ and $\sum_{i=1}^{n} y_{C_i}$ does not contain too many negative or zero returns so that $y_{C_i} < \sum_{i=1}^{n} y_{C_i}$; or: $y_{C_i} < 0$ and $\sum_{i=1}^{n} y_{C_i}$ does not contain too many positive or zero returns so that $y_{C_i} > \sum_{i=1}^{n} y_{C_i}$ (y_{C_i} is less negative than $\sum_{i=1}^{n} y_{C_i}$). Intuitively, these conditions will be met if in a positive (negative) return environment the sum of returns of a portfolio exceeds (is less than) the return of one asset, a condition that is usually met.

In other words, Areas (3.3) up to 10% indicate that two statistical methods are similar, Areas (3.3) greater that 10% can be interpreted as dissimilar models. The greater Area (3.3) is, the more dissimilar two empirical models will be.

2.2 Accounting Characteristics as Factors and Regressions as Specific Statistical Methods (Second Step)

The evaluation criteria "magnitude" and "similarity" regarding economic significance are broadly defined and, thus, can be applied to both financial and accounting data consisting of both prices and returns. Nevertheless, to prepare for an empirical application of our evaluation criteria, it is recommended narrowing the scope of empirical models: Harvey/Liu/Zhu (2016) lists 316 predictors, Harvey/Liu (2019) even more than 400 factors for asset returns (finance papers), Appendix 4 illustrates the different factors in value relevance studies (accounting papers) together with the variety of specific statistical methods used in finance and accounting papers.—It is impossible to analyze all these models.

We, therefore, restrict ourselves to accounting studies because accounting studies use far less factors (see Appendix 4) and, hence, leave room for the analysis of specific statistical methods as well as the interplay between factors and specific statistical methods. With 316 factors on the other hand, the aspect of specific statistical methods would just play a subordinate role.

Moreover, the focus on specific statistical methods as well as the interplay between factors and specific statistical methods advice not to invent any new empirical model (as was done in Chapter II with L_p-norms), but take exclusively well-established empirical models from the literature. That way, we can be sure that non-standard models do not bias our evaluation regarding economic significance. More specifically, accounting characteristics from the value relevance and multiple literature are taken to specify factors. Specific statistical methods are captured with the help of the most frequently used regressions (see Appendix 4).

2.2.1 Accounting Characteristics as Factors

One of the most comprehensive lists of accounting characteristics is contained in Schreiner (2007), p. 39. Therefore, his collection of accounting figures is used with one minor adjustment: The exact label of the accounting characteristics is based on the label used in Thompson Reuters (see Thompson Reuters (2015)) and not on Schreiner (2007).

Important side note: All accounting characteristics are used to explain companies' equity and not entity values.

<u>Single-factor models</u>

The accounting characteristics of Schreiner (2007) can immediately be translated into the following single-factor models.—Translated is the keyword here because Schreiner (2007) uses his factors to compute multiples and does not use them in a regression context.

Model number	**Independent variable**	**Regression equation**
M1	Net Sales Or Revenues (SA)	$\hat{y} = \beta_0 + \beta_1 \cdot SA$
M2	Gross Income (GI)	$\hat{y} = \beta_0 + \beta_1 \cdot GI$
M3	Earnings Before Interest, Taxes & Depreciation (EBITDA)	$\hat{y} = \beta_0 + \beta_1 \cdot EBITDA$
M4	Earnings Before Interest And Taxes (EBIT)	$\hat{y} = \beta_0 + \beta_1 \cdot EBIT$
M5	Earnings Before Taxes (EBT)	$\hat{y} = \beta_0 + \beta_1 \cdot EBT$
M6	Earnings (E)	$\hat{y} = \beta_0 + \beta_1 \cdot E$
M7	Total Assets (TA)	$\hat{y} = \beta_0 + \beta_1 \cdot TA$
M8	Book Value Of Common Equity (B)	$\hat{y} = \beta_0 + \beta_1 \cdot B$

Model number	Independent variable	Regression equation
M9	Invested Capital (IC)	$\hat{y} = \beta_0 + \beta_1 \cdot IC$
M10	Operating Cash Flow (OCF)	$\hat{y} = \beta_0 + \beta_1 \cdot OCF$
M11	Ordinary Cash Dividends (D)	$\hat{y} = \beta_0 + \beta_1 \cdot D$

Table 3.1: List of single-factor models

where y is a symbol that comprises prices/returns and $\widehat{\ }$ denotes estimation.

<u>Multi-factor models</u>

For the multi-factor models, a direct translation of Schreiner's (2007) factors is impossible: Multiples just use one factor, whereas for multi-factor models combinations of factors must be chosen.

To construct multi-factor models, however, value relevance studies in general and, in particular, Ohlson (1995), p. 661 prove helpful: Ohlson (1995) uses earnings (component from the income statement), book values (component from the balance sheet), and dividends (component from the cash flow statement) as explanatory variables of firm's market value. We follow his idea and develop multi-factor models that consist of all three variables as well as multi-factor models that are based on any combination of two out of the three variables. Finally, a model with all 11 independent variables is considered.

Model number	Independent variables	Regression equation
M12	Earnings (E) Book Value Of Common Equity (B)	$\hat{y} = \beta_0 + \beta_1 \cdot E + \beta_2 \cdot B$
M13	Earnings (E) Ordinary Cash Dividends (D)	$\hat{y} = \beta_0 + \beta_1 \cdot E + \beta_2 \cdot D$

Model number	Independent variables	Regression equation
M14	Book Value Of Common Equity (B) Ordinary Cash Dividends (D)	$\hat{y} = \beta_0 + \beta_1 \cdot B + \beta_2 \cdot D$
M15	Earnings (E) Book Value Of Common Equity (B) Ordinary Cash Dividends (D)	$\hat{y} = \beta_0 + \beta_1 \cdot E + \beta_2 \cdot B + \beta_3 \cdot D$
M16	Net Sales Or Revenues (SA) Gross Income (GI) Earnings Before Interest, Taxes & Depreciation (EBITDA) Earnings Before Interest And Taxes (EBIT) Earnings Before Taxes (EBT) Earnings (E) Total Assets (TA) Book Value Of Common Equity (B) Invested Capital (IC) Operating Cash Flow (OCF) Ordinary Cash Dividends (D)	$\hat{y} = \beta_0 + \beta_1 \cdot SA + \beta_2 \cdot GI + \beta_3 \cdot EBITDA + \beta_4 \cdot EBIT + \beta_5 \cdot EBT + \beta_6 \cdot E + \beta_7 \cdot TA + \beta_8 \cdot B + \beta_9 \cdot IC + \beta_{10} \cdot OCF + \beta_{11} \cdot D$

Table 3.2: List of multi-factor models

2.2.2 Regressions as Specific Statistical Methods

Specific statistical methods are captured on the one hand by means of ordinary and weighted least squares regressions because they are the most frequently used statistical methods in empirical asset pricing (see Appendix 4). Moreover, weighted least squares regressions are interesting since they can correct for heteroscedasticity of error terms by eliminating the scale effect (see Easton/ Sommers (2003), Formula (2), p. 42 and Brown/Lo/Lys (1999), Formula (15), p. 105). On the other hand, quantile regressions are considered for two reasons event though they have—with the exception of Allen/Singh/Powell (2011)—not been used in empirical asset pricing. First, quantile regressions possess some nice economic features as Chapter IV will show. Second, they are able to analyze the extreme outcomes in the tail of a distribution by allowing to weight over- and underestimations differently (see Allen/Singh/Powell (2011)).

In summary, the following regressions are used to specify the term "specific statistical methods":

- Ordinary least squares regression
- Weighted least squares regression
- Quantile regression with $\tau = 0.25$
- Quantile regression with $\tau = 0.50$ (median regression)
- Quantile regression with $\tau = 0.75$

2.2.3 Further Restrictions on the Empirical Models Analyzed

Even the restriction to accounting data done in Section 2.2.1 leaves many empirical models to be analyzed: prices and returns in time series or cross section. Therefore, narrowing further down the class of empirical models to be analyzed is strongly recommended.

First, regarding the question of price or return models, we choose price models for two reasons. On the one hand, empirical finance papers tend to focus more on returns, whereas accounting papers are more interested in prices (see Appendix 4). Since we have chosen accounting data, this argument favors price models. On the other hand, Kothari/Zimmermann (1995) and Brown/Lo/Lys (1999) show empirically that returns possess better econometric properties, but prices produce less biased earnings responses.—Since we are interested in economic and not statistical significance, the argument of better econometric properties of returns weights less than the higher economic content of prices.

Second regarding cross-sectional versus time series analysis, we take cross-sectional analysis.—Appendix 4 shows that cross section of prices is far more common in the accounting literature. Hence, choosing cross-sectional analysis prevents our analysis of economic significance from becoming marginal as it would be the case with time series of prices.

2.2.4 Exact Procedure of Determining Company Prices

Company prices are determined out-of-sample, i.e., company i's price (P_i, $i \in \{1, \ldots, n\}$) and its accounting characteristic j ($A_{i,j}$, $j \in \{1, \ldots, m\}$) are not included in the cross-sectional estimation of the regression coefficients. Once the regression coefficients are determined in the cross section, company i's price is calculated as

$$\hat{P}_i = \beta_0 + \beta_1 \cdot A_{i,1} + \cdots + \beta_m \cdot A_{i,m}$$

Each of the 16 models described in Section 2.2.1 is now estimated using each of the 5 regressions described in Section 2.2.2, giving a total of 80 possible different prices for company i.

3 Data Set and Data Cleaning

3.1 Data Set

To apply our evaluation criteria of economic significance in general and, in particular, to elaborate the valuation differences between different factors/specific statistical methods, we use the following data set:

First, companies are taken whose Industry Classification Benchmark (see London Stock Exchange Group plc (2016)) code begins with 2 (“Industrials”) or 3 (“Consumer Goods”) because traditional accounting characteristics should have the highest explanatory power with regard to company prices in these industries. Financials, e.g., do not have meaningful sales that could be compared to companies in other industries or the depreciation of loans would not be comparable to the depreciation of buildings. With the help of the four-digit Industry Classification Benchmark codes it is possible to divide the companies into 10 industries, 19 supersectors, 41 sectors and 114 subsectors.

Second, companies contained in three different regional indices are used: Thompson Reuters Europe, Thompson Reuters United States and Thompson Reuters BRIC (see Thompson Reuters (2016)). The separate consideration of these three different regions makes it possible to detect and not to mix up potential differences in the stock markets and accounting standards and to check the results of this study for their robustness. Europe and the U.S. are used because the majority of existing empirical studies are based on U.S. companies (see, e.g., the survey papers of Harvey/Liu/Zhu (2016), Harvey/Liu (2019), and Dechow/Ge/Schrand (2010)). Europe and emerging markets, however, gain importance (see, e.g., Mölls/Strauß (2007) and Outa/Ozili/Eisenberg (2017)).

Third, the years 2010 to 2014 are considered to examine intertemporal stability.

In summary, the resulting data set consists of 30 partial data sets comprising two industries, three regions, and five years.

Prices and accounting characteristics (see Section 2.2.1) are taken from Thomson Reuters Worldscope.

3.2 Data Cleaning

Data cleaning comprises the following steps:

First, currency-dependent variables are expressed in Euro, i.e., the automatic conversion of Thomson Reuters Worldscope is used to ensure that they are comparable across the companies.

Second, yearly data is used, i.e., market prices and accounting figures are based on the same day, namely the end of the financial year. This means that prices at different days are used for different companies.—Since we use cross-sectional regression, this fact is technically innocuous. Economically, we believe that homogenizing the date of market prices and accounting data is a better idea than explaining companies' end of the year prices (identical date for all companies) with accounting characteristics that come from different dates due to companies' different fiscal years.

Third, all companies are eliminated that contain negative accounting figures between 2010 and 2014. The literature (e.g., the pioneering paper of Collins/Pincus/Xie (1999) and, more recently, Balachandran/Mohanram (2011), Barth/Landsman/Lang/Williams (2012), Givoly/Hayn/Katz (2017), and Baboukardos (2018)) indicates that negative accounting characteristics might lead to a bias in pricing.

After data cleaning, the following numbers of observations remain in the data set:

Data set description	**Data set label**	**Number of companies before data cleaning**	**Number of companies after data cleaning**
European industrials 2010	01_EUROPE_2010_2	321	254
U.S. industrials 2010	02_USA_2010_2	353	160
BRIC industrials 2010	03_BRIC_2010_2	173	136
European industrials 2011	04_EUROPE_2011_2	322	269
U.S. industrials 2011	05_USA_2011_2	364	177
BRIC industrials 2011	06_BRIC_2011_2	170	126
European industrials 2012	07_EUROPE_2012_2	337	240
U.S. industrials 2012	08_USA_2012_2	376	190
BRIC industrials 2012	09_BRIC_2012_2	179	135

Data set description	**Data set label**	**Number of companies before data cleaning**	**Number of companies after data cleaning**
European industrials 2013	10_EUROPE_2013_2	352	232
U.S. industrials 2013	11_USA_2013_2	384	200
BRIC industrials 2013	12_BRIC_2013_2	181	126
European industrials 2014	13_EUROPE_2014_2	365	235
U.S. industrials 2014	14_USA_2014_2	397	222
BRIC industrials 2014	15_BRIC_2014_2	179	123
European consumer goods companies 2010	16_EUROPE_2010_3	159	110
U.S. consumer goods companies 2010	17_USA_2010_3	175	82
BRIC consumer goods companies 2010	18_BRIC_2010_3	117	77
European consumer goods companies 2011	19_EUROPE_2011_3	163	120
U.S. consumer goods companies 2011	20_USA_2011_3	182	92
BRIC consumer goods companies 2011	21_BRIC_2011_3	119	80
European consumer goods companies 2012	22_EUROPE_2012_3	162	110
U.S. consumer goods companies 2012	23_USA_2012_3	183	91
BRIC consumer goods companies 2012	24_BRIC_2012_3	124	90
European consumer goods companies 2013	25_EUROPE_2013_3	173	118
U.S. consumer goods companies 2013	26_USA_2013_3	190	100
BRIC consumer goods companies 2013	27_BRIC_2013_3	125	90

Data set description	Data set label	Number of companies before data cleaning	Number of companies after data cleaning
European consumer goods companies 2014	28_EUROPE_2014_3	180	115
U.S. consumer goods companies 2014	29_USA_2014_3	190	102
BRIC consumer goods companies 2014	30_BRIC_2014_3	122	85

Table 3.3: Number of companies in each data set before and after data cleaning

Table 3.3 shows that the number of companies n is always visibly greater than the number of accounting figures m (with maximum $m = 11$).

4 Results of the Empirical Analysis

To implement our evaluation criterion regarding economic significance of different factors/specific statistical methods developed in Section 2.1, its components "magnitude" and "similarity" are determined for the data set outlined in Section 3.

All computations are performed with RStudio Version 1.1.463 resting upon R version 3.6.0 (see R Core Team (2019)) using the following packages:

- quantreg (version 5.38) for quantile regressions (function rq)

- stats (version 3.6.0) for OLS and WLS regressions (function lm)

4.1 Construction Principle Behind the Ensuing Figures

The ensuing figures of Sections 4.2 and 4.3 are histograms of Ratio (3.1) or Area (3.3) respectively. The exact construction principle behind these histograms, however, deserves some illustration.

Assume that values have been computed for the following companies U with the help of factors$_i$ and regressions:

U_1 factor$_1$ OLS	U_1 factor$_2$ OLS	U_1 factor$_3$ OLS
U_1 factor$_1$ WLS	U_1 factor$_2$ WLS	U_1 factor$_3$ WLS
U_2 factor$_1$ OLS	U_2 factor$_2$ OLS	U_2 factor$_3$ OLS
U_2 factor$_1$ WLS	U_2 factor$_2$ WLS	U_2 factor$_3$ WLS

To analyze the role of regression, e.g., OLS, the following Ratios (3.1) are computed

$$\frac{\mathrm{U_1\ factor_1\ WLS - U_1\ factor_1\ OLS}}{\mathrm{market\ price\ U_1}} \qquad \frac{\mathrm{U_1\ factor_2\ WLS - U_1\ factor_2\ OLS}}{\mathrm{market\ price\ U_1}} \qquad \frac{\mathrm{U_1\ factor_3\ WLS - U_1\ factor_3\ OLS}}{\mathrm{market\ price\ U_1}}$$

$$\frac{\mathrm{U_2\ factor_1\ WLS - U_2\ factor_1\ OLS}}{\mathrm{market\ price\ U_2}} \qquad \frac{\mathrm{U_2\ factor_2\ WLS - U_2\ factor_2\ OLS}}{\mathrm{market\ price\ U_2}} \qquad \frac{\mathrm{U_2\ factor_3\ WLS - U_2\ factor_3\ OLS}}{\mathrm{market\ price\ U_2}}$$

These six ratios form the basis of the histogram computation.

In a similar vein, factors are examined; consider, e.g., factor$_1$:

$$\frac{\mathrm{U_1\ factor_1\ WLS - U_1\ factor_1\ OLS}}{\mathrm{market\ price\ U_1}} \qquad \frac{\mathrm{U_2\ factor_1\ WLS - U_2\ factor_1\ OLS}}{\mathrm{market\ price\ U_2}}$$

$$\frac{\mathrm{U_1\ factor_1\ OLS - U_1\ factor_1\ WLS}}{\mathrm{market\ price\ U_1}} \qquad \frac{\mathrm{U_2\ factor_1\ OLS - U_2\ factor_1\ WLS}}{\mathrm{market\ price\ U_2}}$$

These four (two different) ratios are used to compute histograms.

4.2 “Magnitude” of Price Differences Between Different Factors/Regressions

Before detailed results on “magnitude” are presented, an overview and, that way, a first impression might be helpful. Then, the role of factors as well as the role of regressions are analyzed and, finally, robustness analyses with respect to industry, region, and year are conducted.

4.2.1 Overall Results

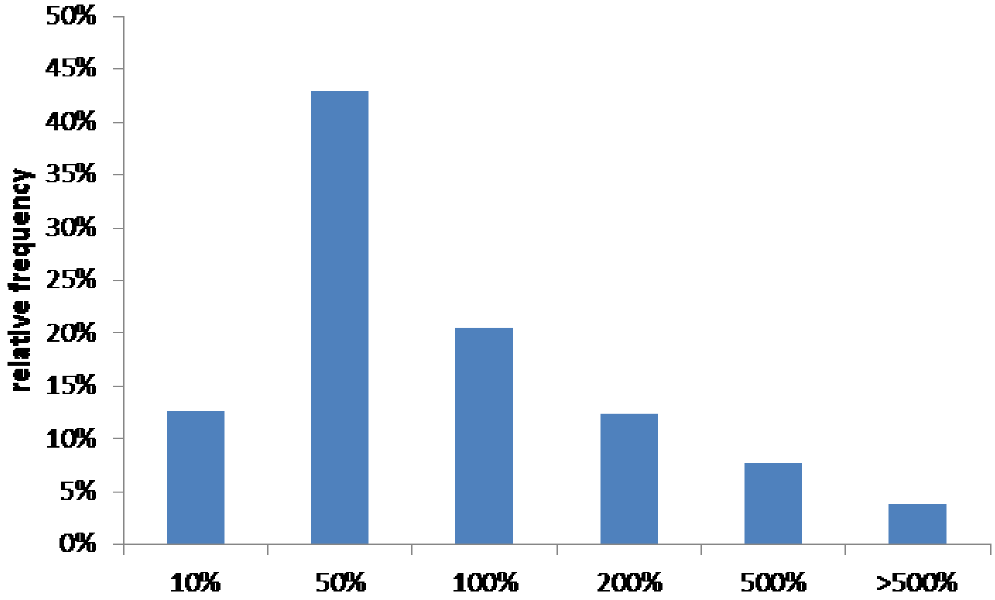


Figure 3.1: Histogram of the absolute values of price differences between different statistical methods measured with the help of |Ratio (3.1)|
Maximum difference: 10,079%

Figure 3.1 shows that only 13% of |Ratio (3.1)| computed across all models belong to the category of "acceptable magnitude" (|Ratio (3.1)| assumes values of 10% or less, see Section 2.1.3.2). On the other hand, 12% of all ratios exhibit a value of more than 200%, i.e., price differences that are more than two times greater than companies' current prices.

In the light of Figure 3.1 it becomes clear that price differences between factors/regressions are not caused by few outliers and are otherwise small. Instead, price differences are generally large.

However, if there are so few differences less than 10% and so many large price differences, the question arises what causes these price differences. In other words, it becomes necessary to analyze in detail the factors of Section 2.2.1 and the specific regressions of Section 2.2.2 to examine whether selected factors/regressions are responsible for these price differences or whether all factors/regressions contribute rather equally.

4.2.2 The Role of Factors

4.2.2.1 Factors in General

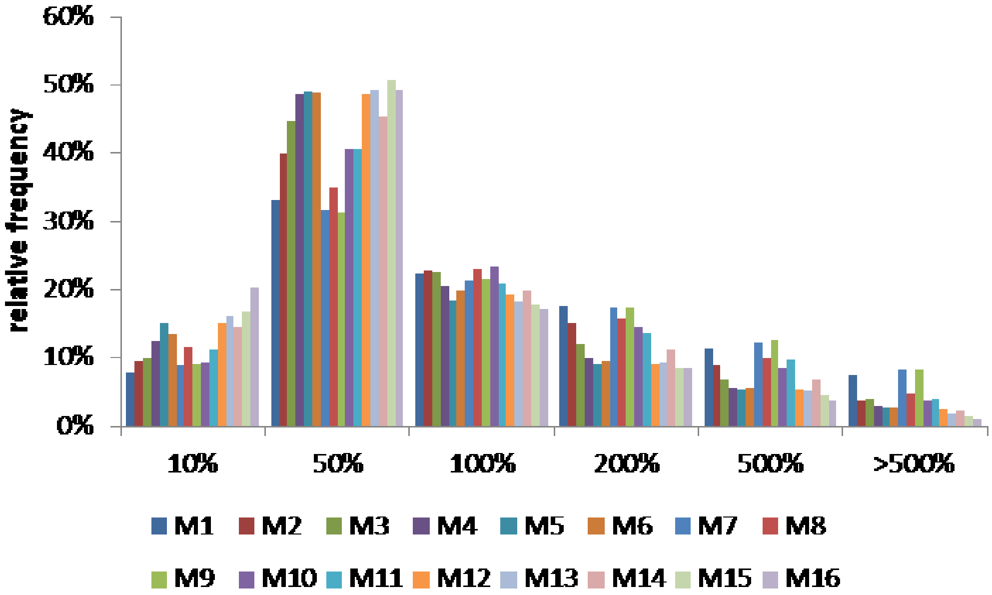


Figure 3.2: Histogram of the absolute values of price differences between different statistical methods measured with the help of |Ratio (3.1)| broken down by factors
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)
Maximum difference: single-factor models: 10,079% (M9); two-factor models: 4,918% (M12); three-factor-model: 2,319%; 11-factor model: 3,979%

Figure 3.2 illustrates that all factors produce large price differences. In detail, the following observations from Figure 3.2 are worth mentioning: Compared to the benchmark of 13% (12%) from the overview Figure 3.1 in the category “less than 10%” (“greater than 200%”), the multi-factor models M12 to M16 fare better with a percentage of M12: 15% (8%), M13: 16% (7%), M14: 15% (9%), M15: 17% (6%), and M16: 20% (5%). From the one-factor models, only M5 15% (8%) and M6 13% (8%) do well compared to the 13% (12%)-benchmark. On the other hand, models M1: 8% (19%), M7: 9% (21%), and M9: 9% (21%) produce the worst values in the category “less than 10%” (“greater than 200%”).

In summary, the factor choice is of high economic significance for company valuation regarding the component “magnitude” of price differences. The fact that all factors produce large price differences can be interpreted as different factors contribute differently to company prices.

4.2.2.2 Factors when Controlled for Regressions

From Figure 3.2 it remains an open question how the interplay between factors and regressions influences the high economic significance of the factor choice. E.g., maybe the factor choice is of economic significance just for ordinary least squares regressions but not for the other regressions.—To answer this question, Figure 3.2 is analyzed for each regression separately.

Regarding the economic significance of factors when controlled for regressions, it is obtained: For the OLS versus WLS, OLS versus Quantile (0.25), OLS versus Quantile (0.75), WLS versus Quantile (0.50), WLS versus Quantile (0.75), Quantile (0.25) versus Quantile (0.50), Quantile (0.25) versus Quantile (0.75), and Quantile (0.50) versus Quantile (0.75) different factors do not matter. The low number in the best category “10% or less” is caused by the regression leaving only a minor influential potential to the factor choice as the following (exemplary) figures illustrates—all figures can be found in Appendix 5.1.1:

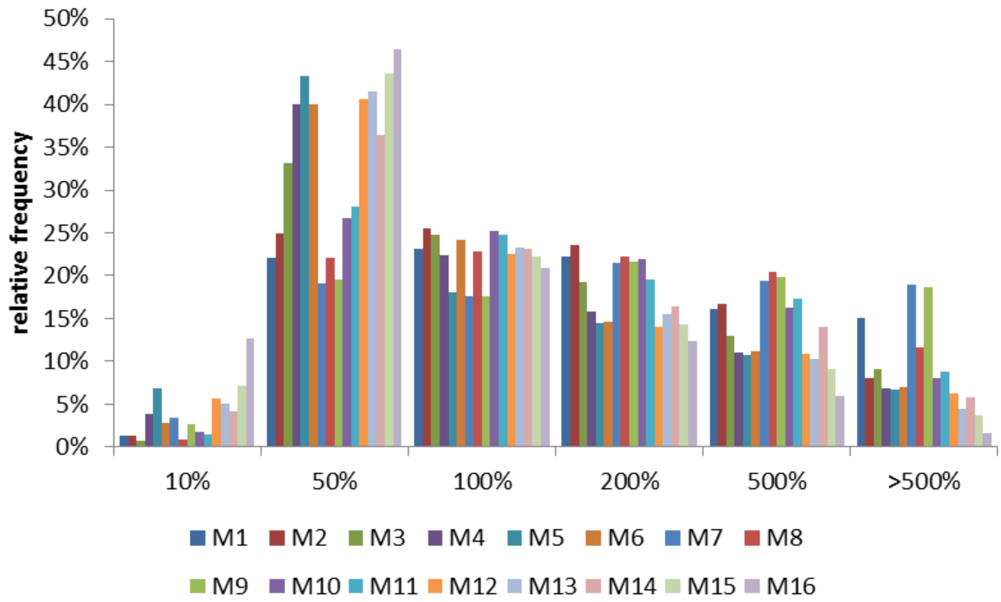


Figure 3.3: Histogram of the absolute values of price differences between method WLS with reference OLS measured with the help of |Ratio (3.1)| broken down by factors M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)

The factor influence is different with OLS versus Quantile (0.50) and WLS versus Quantile (0.25) regressions:

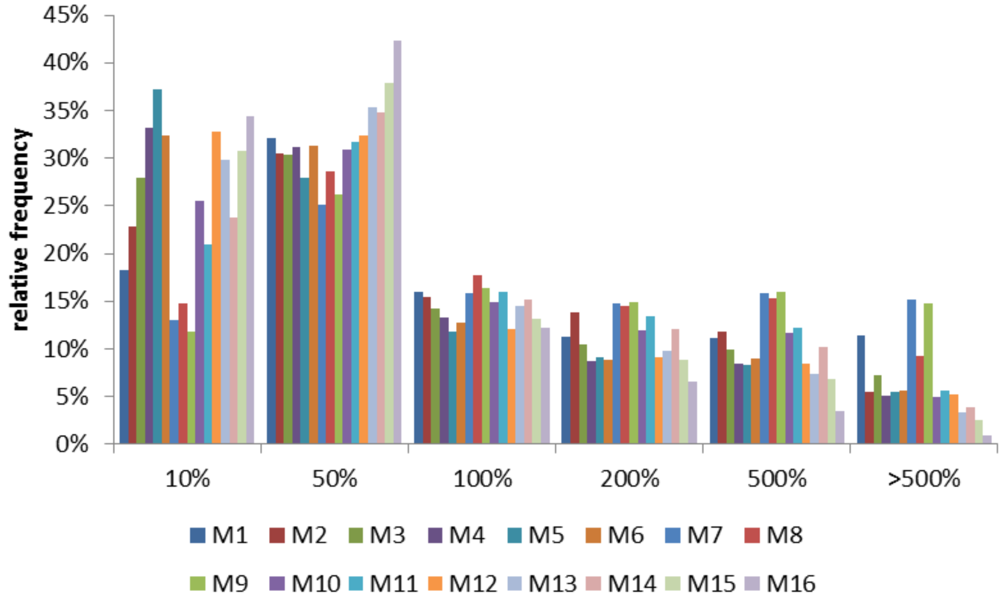


Figure 3.4: **Histogram of the absolute values of price differences between method Quantile (0.50) with reference OLS measured with the help of |Ratio (3.1)| broken down by factors**
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)

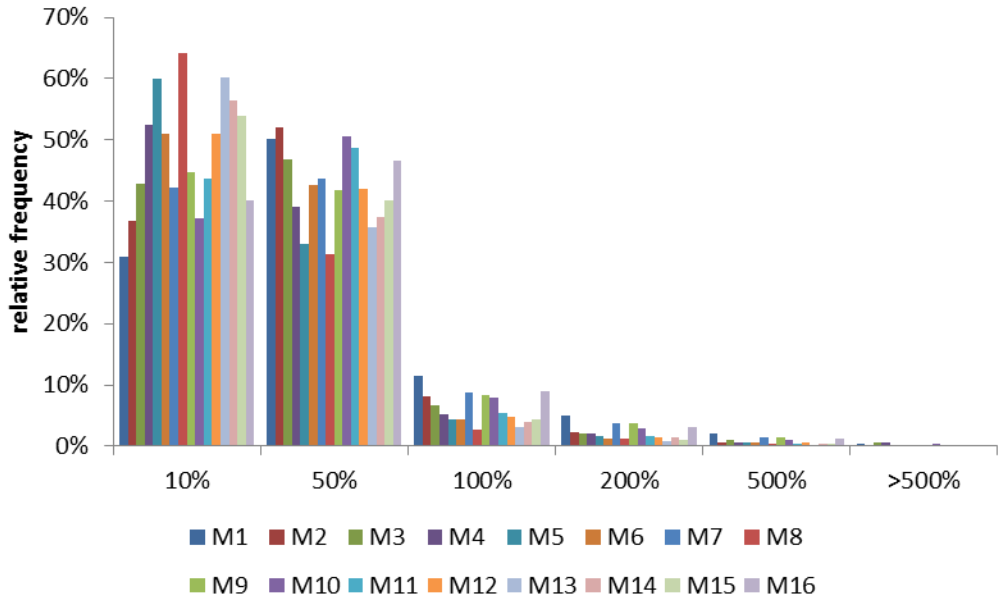


Figure 3.5: **Histogram of the absolute values of price differences between method Quantile (0.25) with reference WLS measured with the help of |Ratio (3.1)| broken down by factors**
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)

In these cases, the statistical methods do not induce big price differences why the factor choice influences economic significance. This can be seen in particular from Figure 3.4: M7, M8, and M9 have a by 50% smaller percentage in the best category “10% or less” than M4, M5, M6 or the multi-factor models M12 to M16.

4.2.3 The Role of Regressions

4.2.3.1 Regressions in General

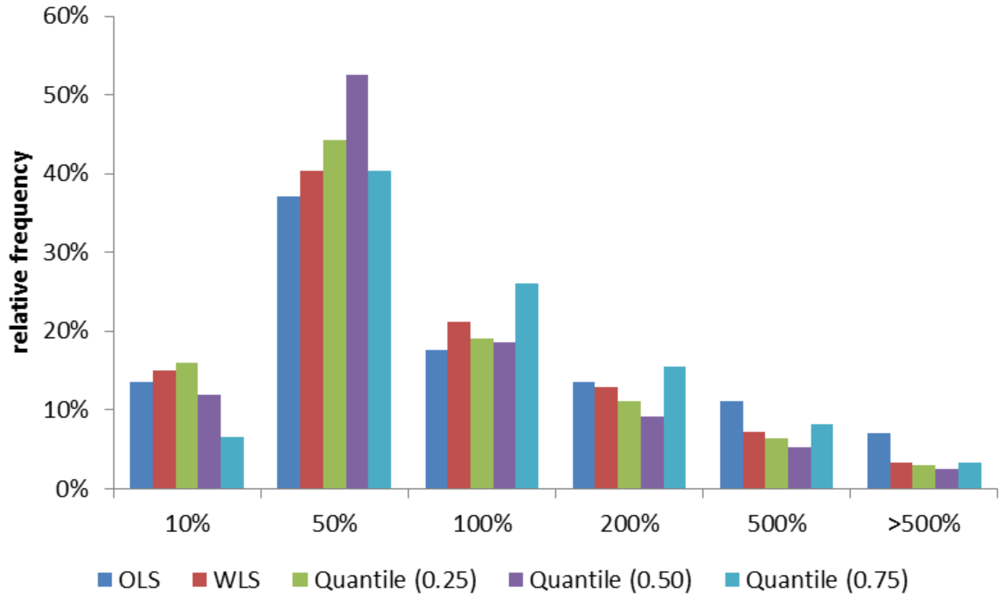


Figure 3.6 Histogram of the absolute values of price differences between different statistical methods measured with the help of |Ratio (3.1)| broken down by the statistical method chosen as reference
Maximum difference: from OLS: 8,318%; from WLS: 10,012%; from Quantile (0.25): 10,079%; from Quantile (0.50): 9,279%; from Quantile (0.75): 10,079%

According to Figure 3.6 all regressions are responsible for large price differences. In detail, the following observations from Figure 3.6 are worth mentioning: Compared to the benchmark of 13% from the overview Figure 3.1 in the category “less than 10%” OLS: 14%, WLS: 15%, and Quantile (0.25): 16% fare best, Quantile (0.50): 12% and Quantile (0.75): 7% worst. Regarding the category “greater than 200%”, it is obtained: WLS: 11%, Quantile (0.25): 10%, Quantile (0.50): 8% are good, OLS: 18% and Quantile (0.75): 12% are bad because they exceed the benchmark percentage of 12%.

In summary, the regression choice is of high economic significance for company valuation regarding the component “magnitude” of price differences.

4.2.3.2 Regressions when Controlled for Factors

Section 4.2.2.2 seems to indicate that regressions are of higher economic significance than factors. This guess must be checked in this section by analyzing the interplay between regressions and factors. To that end, Figure 3.6 is analyzed for each factor separately.

Regarding the economic significance of regressions when controlled for factors, it is obtained: For the OLS versus WLS, OLS versus Quantile (0.25), OLS versus Quantile (0.75), WLS versus Quantile (0.50), WLS versus Quantile (0.75), Quantile (0.25) versus Quantile (0.50), Quantile (0.25) versus Quantile (0.75), and Quantile (0.50) versus Quantile (0.75) different factors do not matter, i.e., they do change the big price differences that these regressions produce. For OLS versus Quantile (0.50) and WLS versus Quantile (0.25) moderate differences are observed that are again not altered by factors as the following (exemplary) figure illustrates—all figures can be found in Appendix 5.1.2:

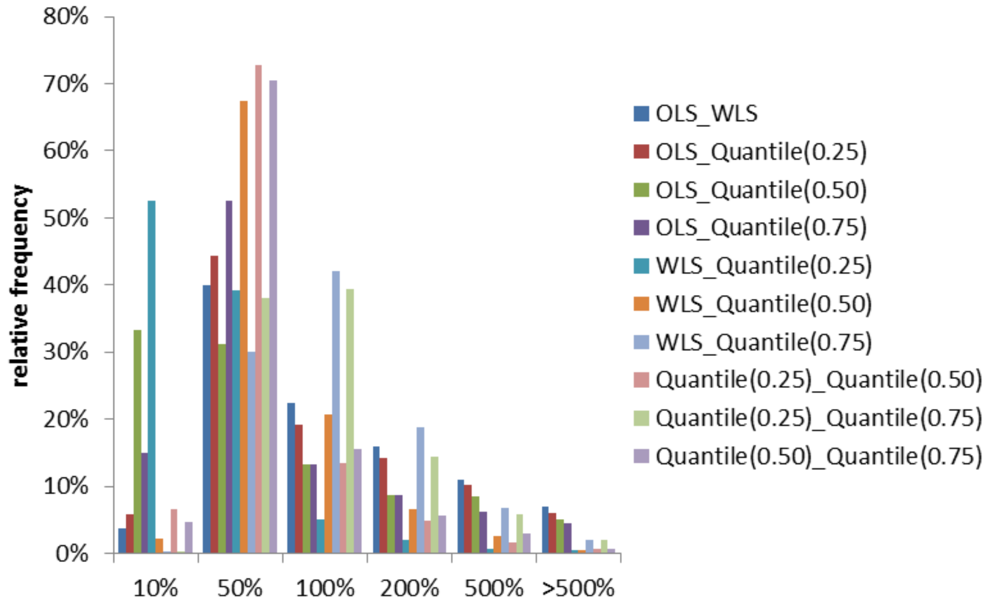


Figure 3.7: Histogram of the absolute values of price differences for M4 measured with the help of |Ratio (3.1)|

Figure 3.7 complements and confirms nicely the findings of Section 4.2.2.2 „Factors when Controlled for Regressions".

4.2.3.3 Regressions that Generate High or Low Prices

As last step to understand the influence of regressions we analyze whether some regressions generate generally higher or lower prices than other regressions. The economic background of this analysis is that, e.g., buyers and sellers of companies have different interest in pricing (buyers: low price; sellers: high price) and, thus, might want to know what regression supports their views.

The following relations between prices and regressions are observable—see Appendix 5.1.3 for a detailed graphical analysis:

		Regression used as reference				
		OLS	WLS	Quantile (0.25)	Quantile (0.50)	Quantile (0.75)
Regression analyzed	OLS	–	higher	higher	higher/ ambiguous	higher/ ambiguous
	WLS	lower	–	ambiguous	lower	lower
	Quantile (0.25)	lower	ambiguous	–	lower	lower
	Quantile (0.50)	lower/ ambiguous	higher	higher	–	lower
	Quantile (0.75)	lower/ ambiguous	higher	higher	higher	–

Table 3.4: Comparison of the sign of Ratio (3.1) to identify regressions that translate into high or low prices
Hints how to read Table 3.4: Table 3.4 computes the difference between prices of regressions analyzed and prices of regressions used as reference. E.g., "lower" means that the regression analyzed produces lower prices than the regression used as reference, e.g., WLS-prices are lower than OLS-prices.

According to Table 3.4, OLS and Quantile (0.75) regressions deliver higher prices than WLS und Quantile (0.25) regressions. In fact, OLS regressions tend to deliver the highest prices of all approaches, whereas WLS and Quantile (0.25) regressions tend to produce the lowest prices of all regressions. This statement is true irrespective of the factors considered (see Appendix 5.1.3). Ambiguity on the other hand might be seen as an indication that these statistical methods are somewhat similar.

Finally, the results regarding the sign of Ratio (3.1) for the three quantile regressions—prices do not necessarily rise with quantiles—deserve a clarifying comment. High (low) quantile regressions determine betas in a way so that estimated prices rarely fall below (increase above) current prices. In other words, in-sample high (low) quantile regressions produce high (low) prices by construction. However, out-of-sample this might not be true due to quantile crossing in the outlying regions of the design space (see, e.g., Koenker (2005), pp. 55 f.). Therefore, the intuition that higher quantiles lead to higher prices can be regarded as good working hypothesis, but cannot be taken as always given.

4.2.4 Robustness Check: Role of Industries, Regions, and Years

Note that the results regarding "magnitude" obtained so far are not driven by industry, region, or year.

– Industry

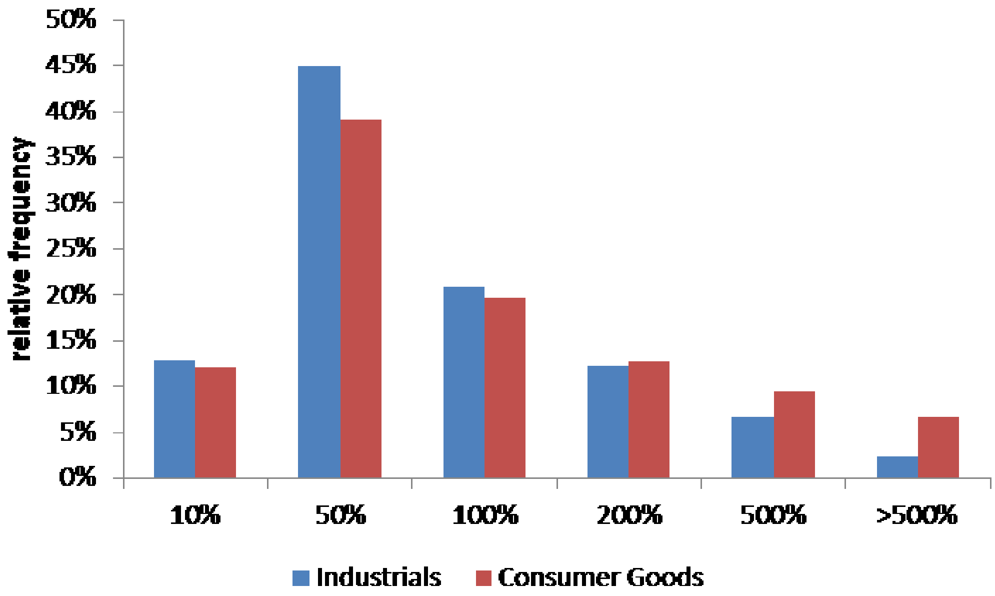


Figure 3.8: **Histogram of the absolute values of price differences between different statistical methods measured with the help of |Ratio (3.1)| broken down by industry Maximum difference: industrials: 7,821%; consumer goods: 10,079%**

Figure 3.8 clarifies that both industrials and consumer goods exhibit a low percentage of differences in the best category “10% or less” and, hence, a similar pattern as in the overview Figure 3.1.

In other words, economic significance of factors and regressions regarding “magnitude” remains valid even if controlled for industries.

– Region

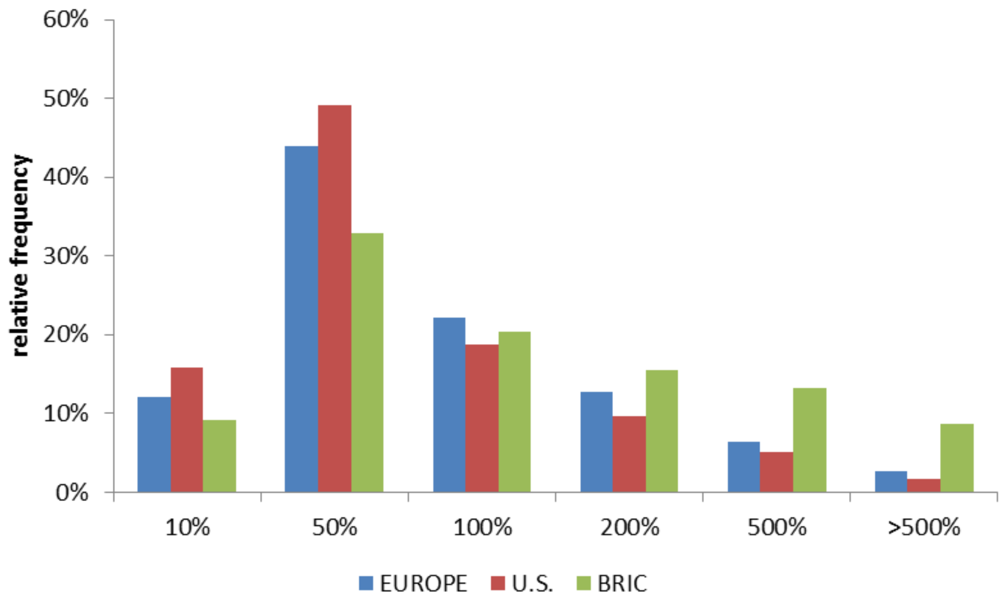


Figure 3.9: Histogram of the absolute values of price differences between different statistical methods measured with the help of |Ratio (3.1)| broken down by region Maximum difference: Europe: 8,278%; U.S.: 6,054%; BRIC: 10,079%

According to Figure 3.9, low percentage of differences in the best category “10% or less” remains small even for the U.S. (16%) where BRIC produces generally lower values in the best category. Therefore, potential differences in the stock market efficiency and accounting standards in different regions do not change the patterns as in the overview Figure 3.1.

In other words, economic significance of factors and regressions regarding “magnitude” remains valid even if controlled for regions.

– Year

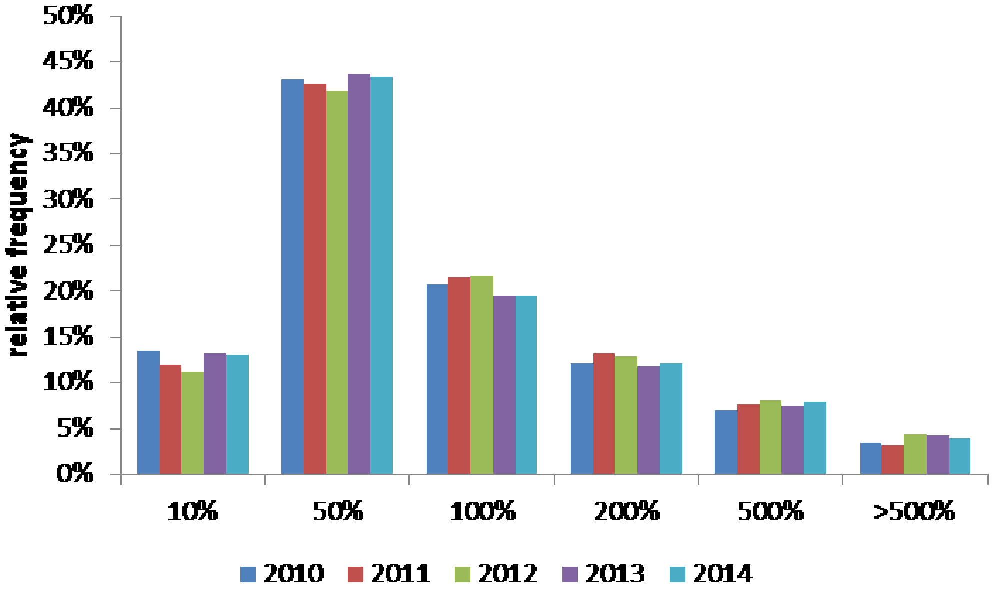


Figure 3.10: Histogram of the absolute values of price differences between different statistical methods measured with the help of |Ratio (3.1)| broken down by year Maximum difference: 2010: 7,821%; 2011: 9,847%; 2012: 10,079%; 2013: 8,728% ; 2014: 8,189%

Figure 3.10 illustrates that potentially different stock market conditions in different years do not change the patterns as in the overview Figure 3.1 regarding model generated price differences: The percentage of differences in the best category “10% or less” does not exceed 15% in any year between 2010 and 2014.

In other words, economic significance of factors and regressions regarding “magnitude” remains valid even if controlled for years.

4.3 “Similarity” of Different Factors/Regressions

While “magnitude” stresses the differences between factors/regressions, i.e., focuses on dissimilarities, “similarity” focuses on the common aspects of factors/regressions. The corresponding nature of “magnitude” and “similarity”, i.e., both components of economic significance, can be illustrated best with the help of some important results regarding “magnitude” from Section 4.2:

The analysis of differences with the help of “magnitude” in the foregoing Section 4.2 suggests as intuition regarding “similarity”: (i) Models that exhibit a high percentage in the best category (“10% or less”), e.g., WLS versus Quantile (0.25) for M4 in Figure 3.7, are similar; (ii) models where the sign of Ratio (3.1) is ambiguous in Table 3.4, e.g., WLS versus Quantile (0.25), are similar.

However, this intuition is too crude since it cannot evaluate the higher/lower cases of Table 3.4 and, of course, cannot systematically judge similarities between factors/regressions. E.g., what exact percentage in the best category (“10% or less”) is required so that empirical models are classified as similar?

4.3.1 Overall Results

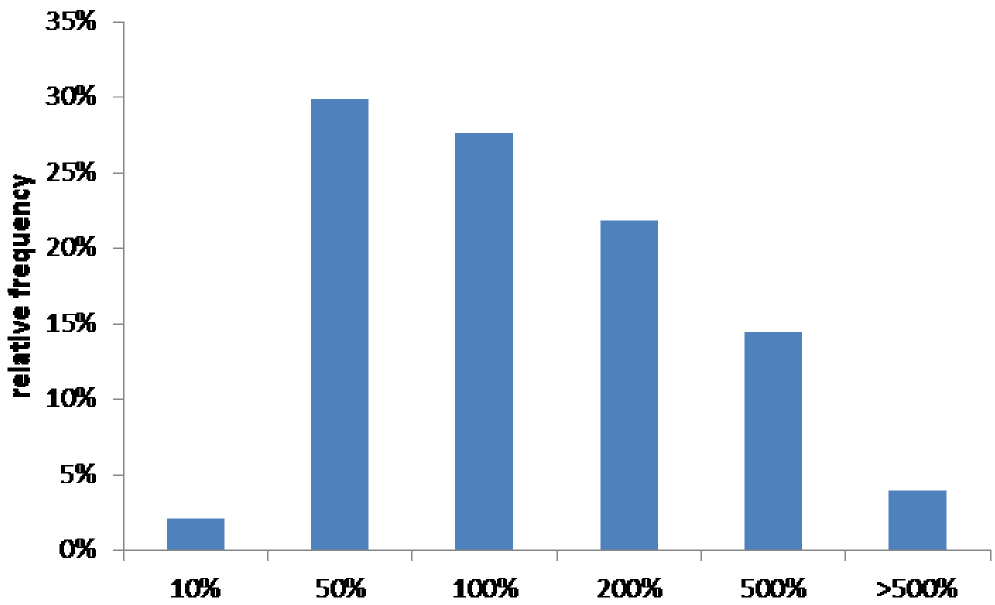


Figure 3.11: Histogram of dissimilarities between different statistical methods measured with the help of Area (3.3)
Maximum size of the dissimilarity area: 1,079%

Figure 3.11 shows that only 2% of Area (3.3) computed across all models belong to the category with the highest similarity (“Area (3.3) of size 10% or less”). Therefore, it becomes clear that dissimilarities between different factors/regressions are not caused by few outliers and are otherwise small. Instead, dissimilarities are generally large. In particular, Figure 3.11 illustrates that rules of thumb like “OLS and median regressions should be similar because they both use mean values" or "models where the sign of Ratio (3.1) is ambiguous in Table 3.4, e.g., WLS versus Quantile (0.25), are similar" are not true.

However, if there are so few similarities in the best category ("10% or less"), the question arises what causes these dissimilarities. In other words, it becomes necessary to analyze in detail the factors of Section 2.2.1 and the regressions of Section 2.2.2 to examine whether selected factors/regressions are responsible for these dissimilarities or whether all factors/regressions contribute rather equally. In particular, are there certain factors/regressions that are always similar and others that are always dissimilar? Is the degree of similarity between factors/regressions constant over various industries, regions, and years?

4.3.2 The Role of Factors

4.3.2.1 Factors in General

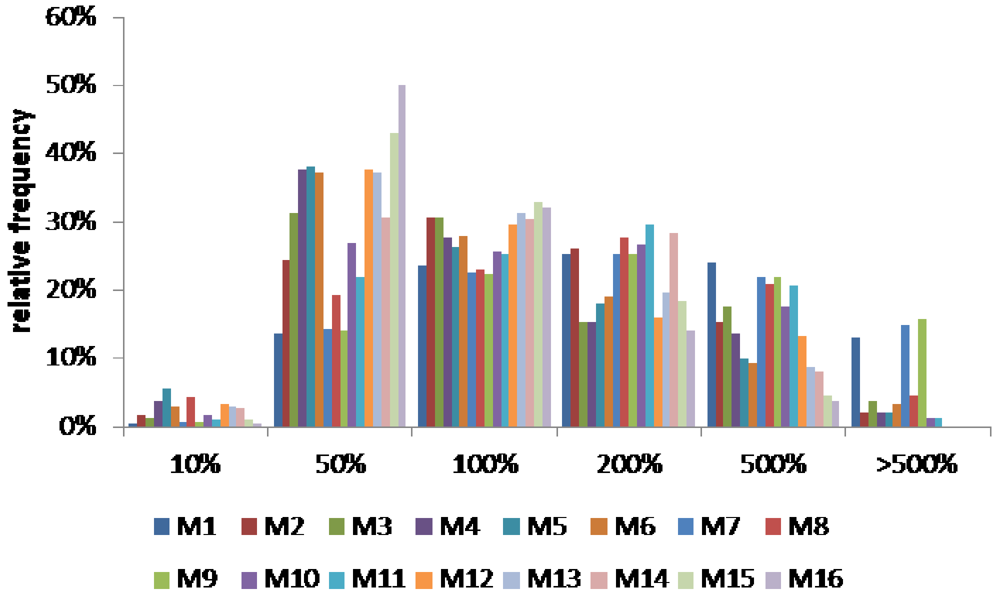


Figure 3.12: Histogram of dissimilarities between different statistical methods measured with the help of Area (3.3) broken down by factors
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)
Maximum size of the dissimilarity area: single-factor models: 1,079% (M9); two-factor models: 494% (M12); three-factor-model: 384%; 11-factor model: 423%

Figure 3.12 illustrates that all factors produce large dissimilarities. The fact that some models do not have large differences in Ratio (3.1) does obviously not mean that these models are similar. Notably, multi-factor models that produce moderate differences measured by means of Ratio (3.1) are as dissimilar as one-factor models which had large differences in the form of high Ratios (3.1).

In summary, the factor choice is of high economic significance for company valuation regarding the component “similarity” of price differences.

4.3.2.2 Factors when Controlled for Regressions

The problem with Figure 3.12 is, however, that it commingles the influence of factors/regressions on “similarity”. To be able to answer the question how the interplay between factors and regressions influences “similarity”, Figure 3.12 is examined for each regression separately.

Regarding the economic significance of factors when controlled for regressions, it is obtained:

On the one hand, the analysis regarding “similarity” reproduces the results of “magnitude” done in Section 4.2.2.2: OLS versus WLS, OLS versus Quantile (0.25), OLS versus Quantile (0.75), WLS versus Quantile (0.50), WLS versus Quantile (0.75), Quantile (0.25) versus Quantile (0.50), Quantile (0.25) versus Quantile (0.75), and Quantile (0.50) versus Quantile (0.75) are irrespective of the specific factors considered as dissimilar as the following (exemplary) figures illustrate—all figures can be found in Appendix 5.1.4:

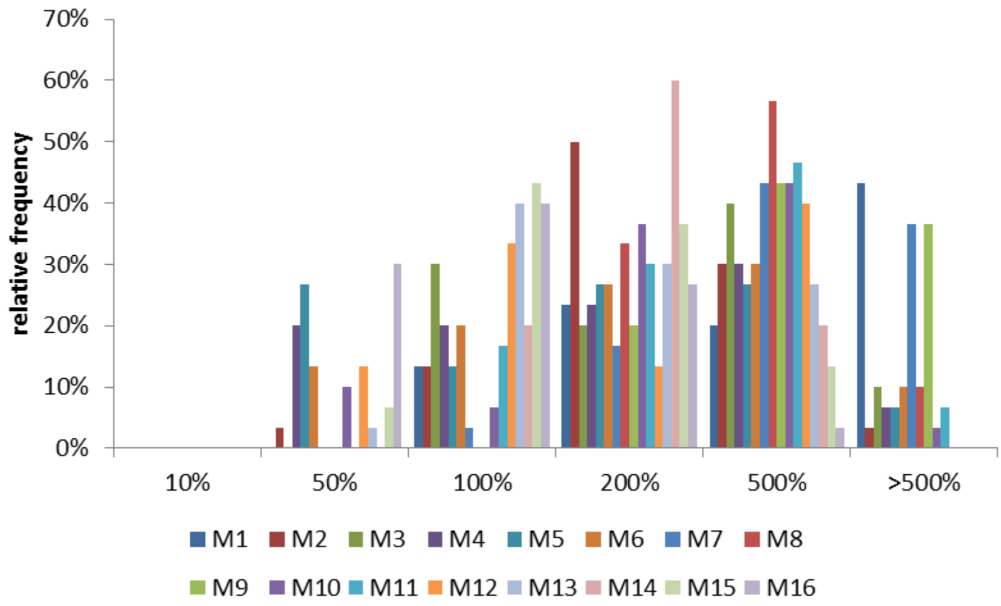


Figure 3.13: Histogram of dissimilarities between method WLS with reference OLS measured with the help of Area (3.3) broken down by factors
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)

On the other hand, the comparison “magnitude” between OLS und Quantile (0.50) regressions in Figure 3.4 underestimates the dissimilarity between both regressions even if it is controlled for factors as the ensuing Figure 3.14 illustrates:

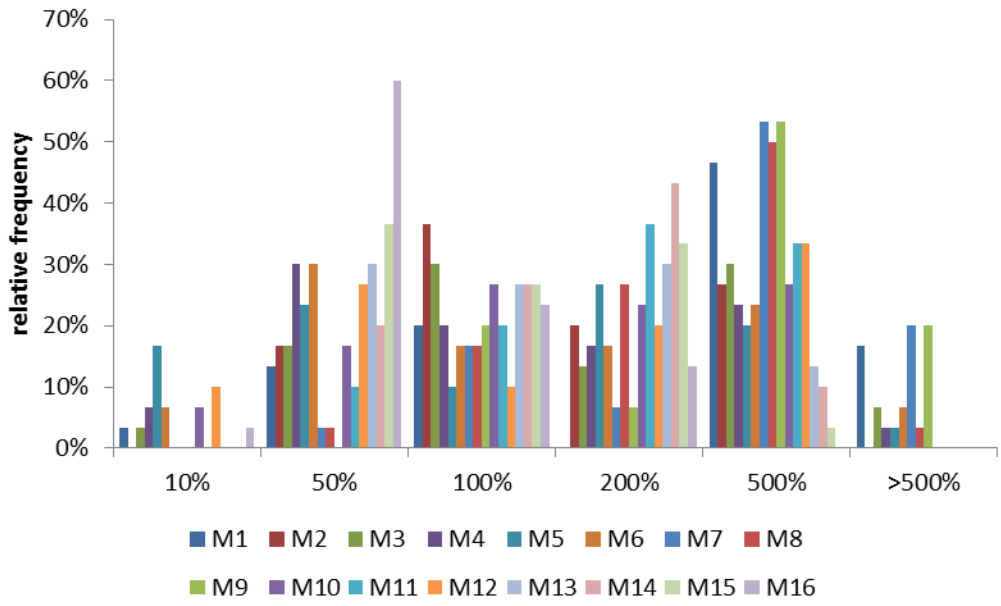


Figure 3.14: Histogram of dissimilarities between method Quantile (0.50) with reference OLS measured with the help of Area (3.3) broken down by factors
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)

Finally, factors exert an effect on “similarity” of WLS versus Quantile (0.25). For 7 of 16 factors a percentage of more than 20% of the best “similarity” category “10% or less” is achieved, for the other 9 factors such a percentage cannot be achieved as Figure 3.15 demonstrates:

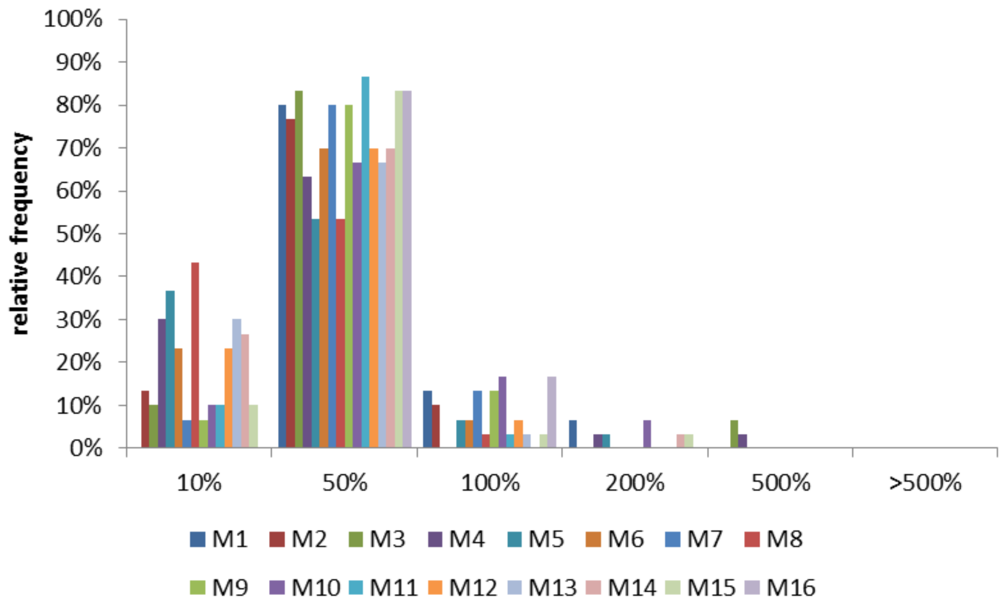


Figure 3.15: Histogram of dissimilarities between method Quantile (0.25) with reference WLS measured with the help of Area (3.3) broken down by factors
M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept)

The overall verdict of dissimilarity in Figure 3.12 regarding WLS versus Quantile (0.25), thus, hides the influence of factors.

4.3.3 The Role of Regressions

4.3.3.1 Regressions in General

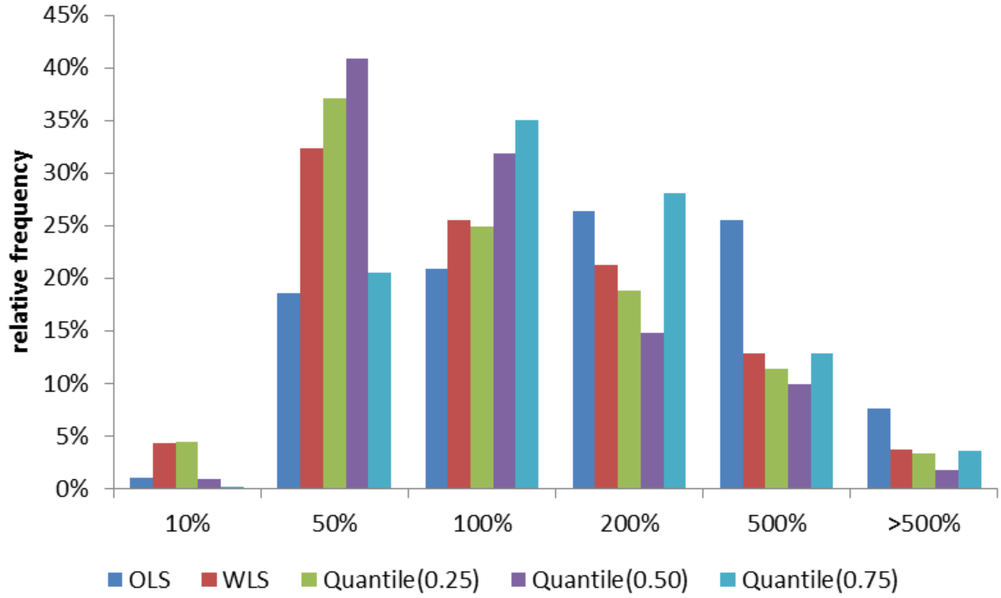


Figure 3.16: Histogram of dissimilarities between different statistical methods measured with the help of Area (3.3) broken down by the statistical method chosen as reference model
Maximum difference: from OLS: 949%; from WLS: 1,079%; from Quantile (0.25): 1,036%; from Quantile (0.50): 908%; from Quantile (0.75): 1,079%

According to Figure 3.16 all regressions are responsible for dissimilarities meaning that the regression choice is of high economic significance for company valuation regarding the component "similarity".

4.3.3.2 Regressions when Controlled for Factors

Section 4.3.3.1 seems to indicate that regressions are of higher economic significance regarding the explanation of dissimilarities than factors. This guess must be checked in this section by analyzing the interplay between regressions and factors. To that end, Figure 3.16 is analyzed for each factor separately.

Regarding the economic significance of regressions when controlled for factors, it is obtained: OLS versus WLS, OLS versus Quantile (0.25), OLS versus Quantile (0.75), WLS versus Quantile (0.50), WLS versus Quantile (0.75), Quantile (0.25) versus Quantile (0.50), Quantile (0.25) versus Quantile (0.75), and Quantile (0.50) versus Quantile (0.75) produce high dissimilarities irrespective of factors, i.e., factors do not neutralize the economic significance of regression as the following (exemplary) figure illustrates—all figures can be found in Appendix 5.1.5:

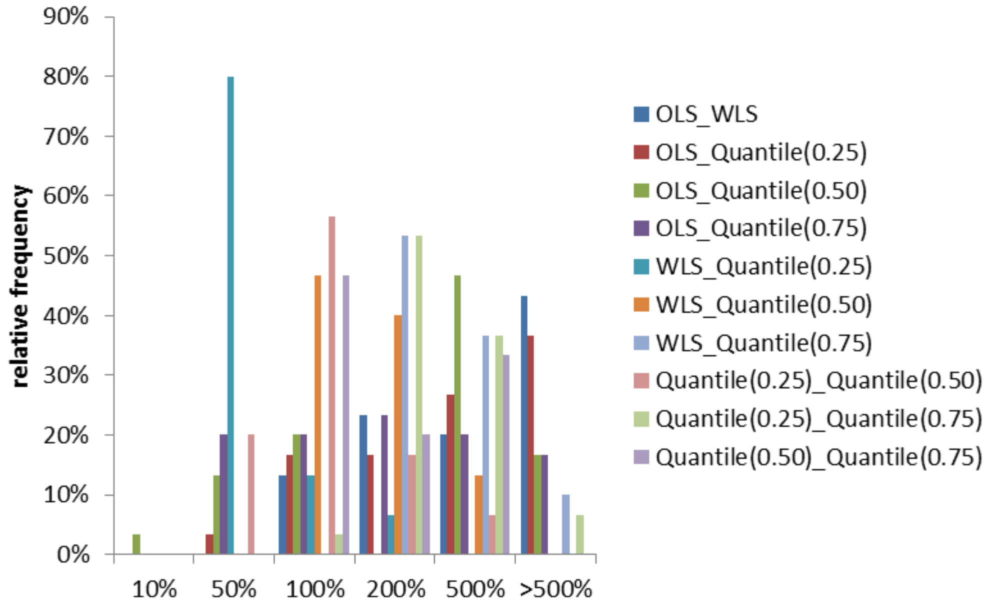


Figure 3.17: Histogram of dissimilarities for M1 measured with the help of Area (3.3)

Factors do not alter the dissimilarity results regarding OLS versus Quantile (0.50) regressions. However, factors exert influence on the "similarity" of WLS versus Quantile (0.25) regressions as the following (exemplary) figure illustrates—all figures can be found in Appendix 5.1.5:

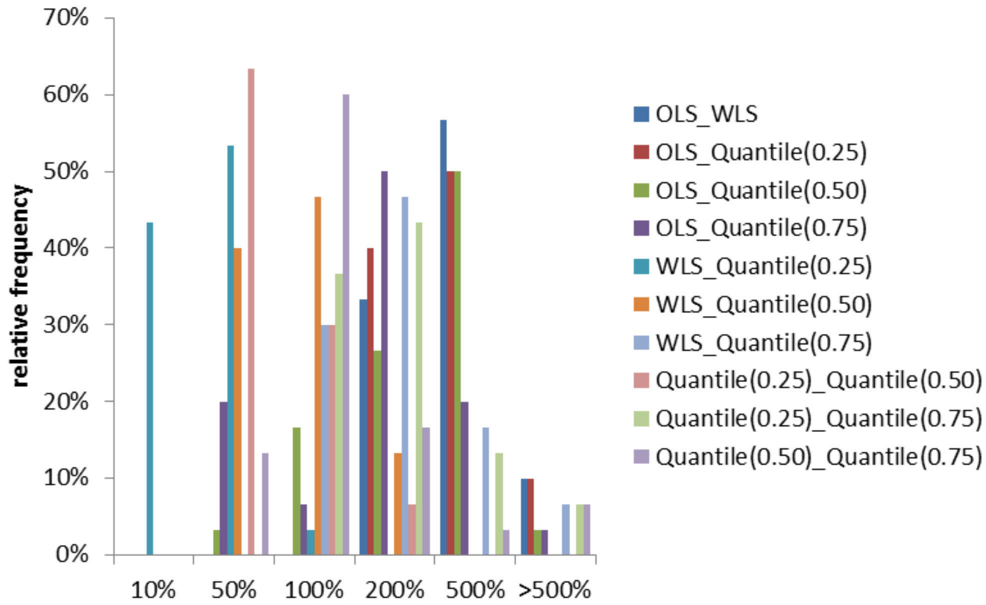


Figure 3.18: Histogram of dissimilarities for M8 measured with the help of Area (3.3)

Figure 3.18 indicates that in model M8 WLS and Quantile (0.25) are rather similar, which is not the case in model M1.

To be more precise regarding the phrase “rather similar”, high dissimilarity (percentage of the best category “10% or less” is equal to less than 20%) for WLS versus Quantile (0.25) is given for M1, M2, M3, M7, M9, M10, M11, M15, M16, whereas moderate (percentage of the best category “10% or less” is equal to between 20% and 50%) dissimilarity holds for M4, M5, M6, M8, M12, M13, M14.

The overall verdict of dissimilarity in Figure 3.16 regarding WLS versus Quantile (0.25), thus, hides the influence of factors.

4.3.4 Robustness Check: Role of Industries, Regions, and Years

Note that the results regarding “similarity” obtained so far are not driven by industry, region, or year.

– Industry

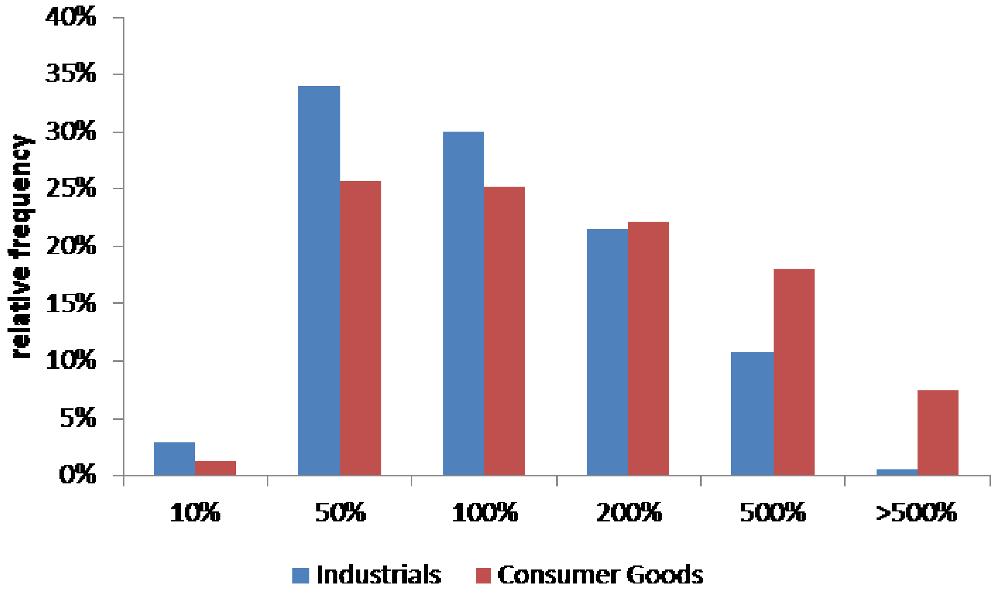


Figure 3.19: Histogram of dissimilarities between different statistical methods measured with the help of Area (3.3) broken down by industry
Maximum size of the dissimilarity area: industrials: 650%; consumer goods: 1,079%

Figure 3.19 clarifies that both industrials and consumer goods exhibit a low percentage of the best category “10% or less” and, hence, a similar pattern as in the overview Figure 3.11.

In other words, economic significance of factors and regressions regarding “similarity” remains valid even if controlled for industries.

– Region

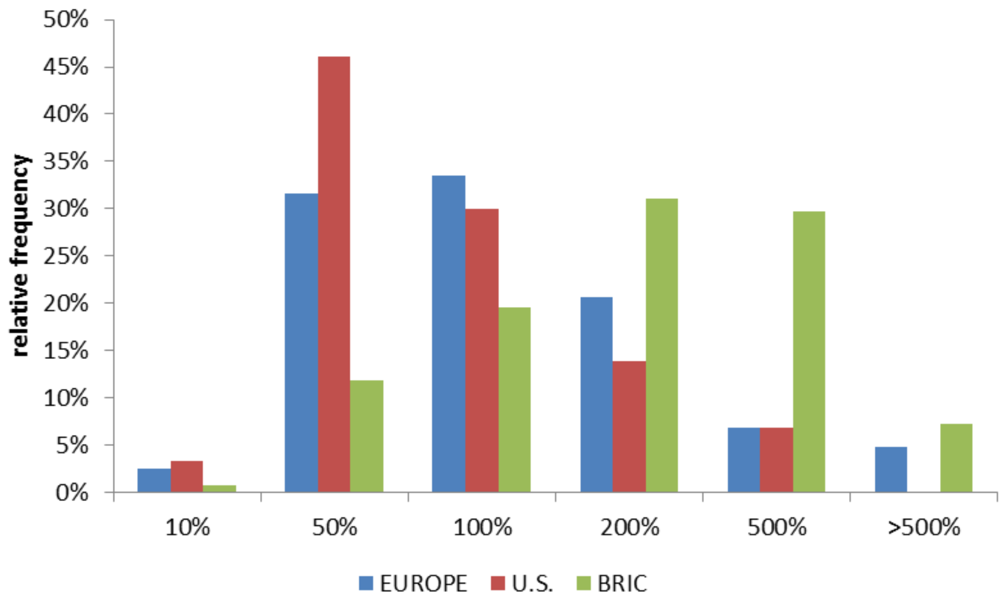


Figure 3.20: Histogram of dissimilarities between different statistical methods measured with the help of Area (3.3) broken down by region
Maximum size of the dissimilarity area: Europe: 929%; U.S.: 468%; BRIC: 1,079%

According to Figure 3.20, the percentage in the best category “10% or less” remains small even for the U.S. (3% compared to 2% of overview Figure 3.11) where BRIC produces generally higher dissimilarities (1% compared to 2% of overview Figure 3.11). Therefore, potential differences in the stock market efficiency and accounting standards in different regions do not change the patterns as in the overview Figure 3.11.

In other words, economic significance of factors and regressions regarding “similarity” remains valid even if controlled for regions.

– Year

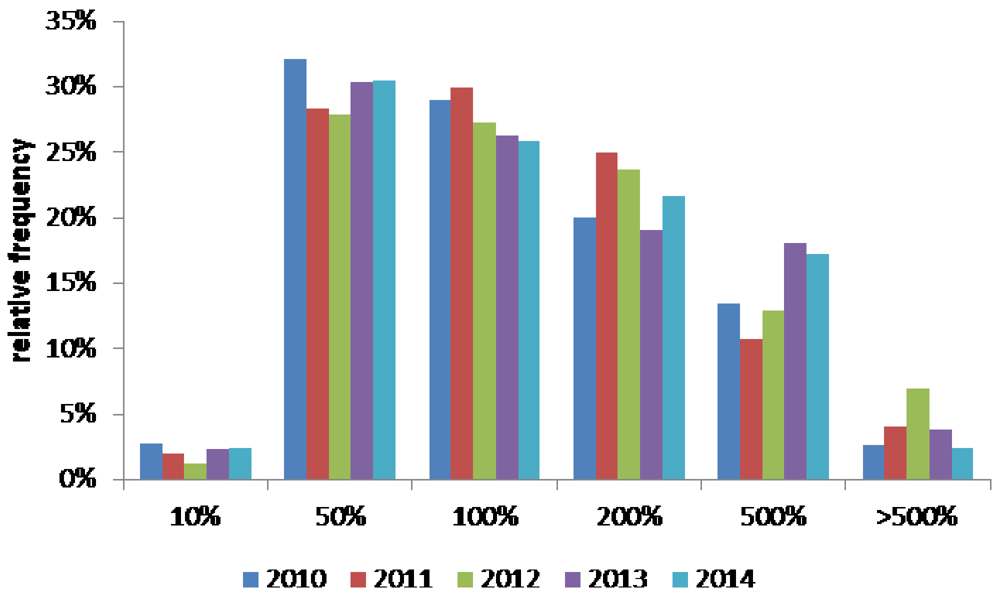


Figure 3.21: Histogram of dissimilarities between different statistical methods measured with the help of Area (3.3) broken down by year
Maximum size of the dissimilarity area: 2010: 717%; 2011: 789%; 2012: 1,079%; 2013: 949%; 2014: 656%

Figure 3.21 illustrates that potentially different stock market conditions in different years do not change the patterns as in the overview Figure 3.11 regarding model generated dissimilarities: The size of Area (3.3) in the best category "10% or less" does not exceed 3% in any year between 2010 and 2014.

In other words, economic significance of factors and regressions regarding "similarity" remains valid even if controlled for years.

5 Conclusion

The explanatory power of each empirical analysis depends on the chosen factors (numbers and specific selection of explanatory variables) as well as the specific statistical method used (e.g., ordinary least squares regression, quantile regression). The literature is aware of the importance of factors/specific statistical methods and, hence, analyzes the statistical significance of factors/specific statistical methods. Economic relevance of factors/specific statistical methods, on the other hand, is far less analyzed and, hence, understood.

Therefore, it is the objective of this Chapter III to analyze the economic significance of different factors/specific statistical methods. To achieve this objective, cross-sectional regression models with accounting figures as explanatory variables are used.

The results of this chapter can be summarized as follows:

First, economic significance regarding different factors/specific statistical methods addresses the question how and not just whether (as with statistical significance) the choice of different factors/specific statistical methods influences company prices/returns and consists of two components: “magnitude” and “similarity”. “Magnitude” focuses on the size of differences between prices/returns that different factors/specific statistical methods produce. “Similarity” condenses the cumulative relative frequency distribution of price/return differences into one number and addresses the problem that moderate price/return differences do not necessarily mean similar empirical models.

Second, “magnitude” shows for our data basis, i.e., company prices in the cross section, that price differences are generally large. Only 13% of all factors/specific statistical methods belong to the best category (absolute values of price differences of 10% or less). These price differences are primarily caused by specific statistical methods and not so much by factors.

Third, “similarity” applied to our data basis illustrates that nearly all factors/specific statistical methods are dissimilar where statistical methods are primarily responsible for this lack of similarity and factors play only a minor role.

Given that the specific statistical method is the primary reason for both high “magnitude” values and low „similarity” degrees of empirical models, the specific statistical method should be chosen carefully. This means, that an economic model selection criterion would be helpful because economic valuation problems should be tackled using economic and not statistical criteria.—Such an economic model evaluation criterion is developed in the ensuing Chapter IV.

Chapter IV: Developing an Economic Model Evaluation Criterion and Applying it to Selected Empirical Asset Pricing Models

1 Introduction

Working with empirical models in general means that two fundamental questions must be answered: (i) What and how many factors, i.e., explanatory variables, should be used? (ii) What empirical model, i.e., ordinary least squares regression, quantile regression etc., is to be applied?—Both questions are nowadays seen as critical. Harvey (2017), pp. 1413 f. argues that trying different empirical models can be regarded as one form of p-hacking. The American Statistical Association (2016) points out that business decisions should not be based only on whether a p-value passes a specific threshold and that statistical significance (p-value) cannot measure the size of an effect or the importance of a result. In other words, for economic problems an economic model evaluation criterion is desirable.—This chapter considers a sub-class of economic models: empirical asset pricing models.

No arbitrage provides a general guideline for economic model evaluation for theoretical asset pricing models in that prices must be a linear function of their future cash flows. Empirical asset pricing models, however, do not rely on present values of cash flows, but on assumed relations between accounting characteristics/factor returns and company prices/returns (see Chapter II for an overview). For that reason, no theoretical guidelines regarding the components of the model exist. In particular, there are neither hints regarding the number and type of explanatory variables nor the specific empirical model (ordinary least squares regression, quantile regression etc.). To make things worse, (i) Chapter III shows that there are huge differences in corporate values when different factors and statistical methods are applied so that virtually arbitrary corporate values can be justified. Nietert/Otto (2018) demonstrate that the same is true if the method of multiples is used to compute company prices. (ii) Moreover, there is a recent trend in the literature (see Appendix 4), to use more complex and diverse statistical methods. Initially, the literature used ordinary and partially weighted least squares regressions. Now generalized least squares regressions gain importance (see the explicit recommendation of Lewellen/Nagel/Shanken (2010), p. 183 to use generalized least squares R^2) together with sophisticated machine learning (Gu/Kelly/Xiu (2018) and Barth/Li/McClure (2018)).

Given this high need for an economic model evaluation criterion, the objective of this Chapter IV is twofold: (i) first develop an economic model evaluation criterion; (ii) come up with an economic ranking of different empirical models.

To achieve this objective, the optimization problems of the empirical asset pricing approaches (of Chapter II) are transformed with the help of Lagrange duality to their corresponding dual programs. The dual program contains the price of the company in the objective function and, hence, possesses a clear economic interpretation that can be related to arbitrage theory of theoretical asset pricing. Based on the dual program a ranking of models can be derived in a sense that the best models are those that use the most innocuous assumptions.

The results of this chapter can be summarized as follows:

First, the economic model evaluation criterion judges the implicit economic assumptions revealed by computing the dual program along the two dimensions compliance with the economic principle and institutional circumstances.

Second, applying the economic evaluation criterion to empirical models reveals that regressions on cross section of prices can be regarded as acceptable from an economic perspective, whereas regressions on cross section of returns and time series models as well as the method of multiples do not comply with the economic principle.

Third, within the group of cross-sectional price models quantile regression proves to be the best model because it is able to offer a good approximation to the economic principle and mimics best the institutional circumstances, in particular, if the regression is run without a constant. On the other hand, statistically more advanced models like generalized least squares regressions deteriorate the implied economic content of models: They work with weighted prices; however assets can only be purchased and sold at (unweighted) prices.

This chapter makes the following contribution compared to the literature.

First, both Harvey (2017), pp. 1413 f. and The American Statistical Association (2016) advocate economic model evaluation criteria. This chapter makes a first attempt at deriving such an economic model evaluation criterion by evaluating empirical models based on economic/theoretic criteria like no arbitrage/economic principle. As opposed to our approach, model evaluation in the literature still rests primarily on statistical criteria. (i) Black/Jensen/Scholes (1972), p. 6 develop the standard quality assessment for empirical models: The intercept of a regression (alpha) should not be significantly different from zero, Gibbons/Ross/Shanken (1989) design the corresponding statistical test, Cochrane (2005), p. 230 extends this test to heteroskedastic and autocorrelated errors. A further development of alpha towards a better economic interpretation is the squared Sharpe ratio of MacKinlay (1995), p. 6 in connection with Barillas/ Shanken (2017), pp. 1317 f.—However, Sharpe ratios are limited to μ-σ-preferences which are known to miss arbitrage opportunities. Therefore, their role as economic model evaluation criterion is doubtful. (ii) As purely statistical criteria the out-of-sample R^2 (e.g., Campbell/Thompson (2008)) or the size of the estimated slope coefficients (e.g., Lev/Zarowin (1999), p. 356) are used. Lewellen/Nagel/Shanken (2010) give prescriptions how to improve empirical models statistically and explicitly recommend using generalized least squares R^2 (see Lewellen/Nagel/Shanken (2010), p. 183). However, R^2 is not naturally applicable to non-quadratic objective functions as in quantile regressions (see Allen/ Singh/Powell (2011)) or Classification and Regression Trees (CART) estimation functions (see, e.g., Barth/Li/McClure (2018)) where CART is a non-parametric estimation approach that does not require the researcher to specify the relation's functional form. The most advanced paper regarding statistical method evaluation, Barillas/Shanken (2018), employs an empirical nesting approach. If, e.g., the CAPM and the Fama/French (1993) three factor model were equivalent regarding the intercept (alpha is equal to zero), the CAPM would be favored because it was the more parsimonious model.—All these approaches do not consider an economic model evaluation criterion. (iii) Some papers use intuitive arguments to justify particular statistical methods: Brown/Lo/Lys (1999) raise concerns about the use of the coefficient of determination as a measure of value relevance in price regressions because it might be biased due to a scale effect. Allen/Singh/Powell (2011) deliberately use quantile regressions instead of ordinary least squares regressions because quantile regressions are able to better analyze the extreme outcomes in the tail of a distribution. Easton/Sommers (2003) prefer weighted least squares regressions to deal with heteroscedasticity.—Again, these arguments are intuitive, but lack an economic/theoretical reasoning, a gap that we fill with our economic model evaluation criterion.

Second, we provide an economic/theoretical argument for choosing prices over returns as dependent variables in empirical asset pricing. Kothari/Zimmermann (1995) were among the first to raise the question whether prices or returns should be selected as dependent variables: They show empirically that returns possess better econometric properties, prices produce less biased earnings responses. Brown/Lo/Lys (1999) reach similar results and Barth/Beaver/Landsman (2001) conclude that price studies are interested in determining what is reflected in firm value while return studies (price changes) are interested in determining what is reflected in change in value over a specific period of time.—We provide a theoretical analysis of price versus return as dependent variable and a model-based justification of Kothari/Zimmermann's (1995) statement.

Third, we apply duality theory to model evaluation in the field of empirical asset pricing. That duality theory can be used to get a better understanding of linear programming, in particular, production planning (see, e.g., Boyd/Vandenberghe (2009), p. 240) is common knowledge in the literature. An application in empirical finance is rare, however. We are only aware of Wilhelm/Brüning (1992) who apply duality theory to identify the implicit economic assumptions in the field of empirical term structure estimation.—This chapter is inspired by them.

The remainder of this Chapter IV is organized as follows: Section 2 develops an economic model evaluation criterion. Section 3 applies this criterion to various models and identifies economically convincing empirical models. Section 4 concludes this chapter.

2 Developing an Economic Model Evaluation Criterion

Developing an economic model evaluation criterion means (i) revealing the implicit economic assumptions of empirical asset pricing models and (ii) judging them from an economic perspective.—Revealing the implicit economic assumption is achieved using Lagrange duality. Judging the implicit economic assumptions is done by applying what we call economic dominance of models.

2.1 Components of the Economic Model Evaluation Criterion

2.1.1 Lagrange Duality

The original economic application of (Lagrange) duality has been production planning of companies (see, e.g., Boyd/Vandenberghe (2009), p. 240). Companies determine their optimal production program, i.e., the optimal numbers of each product, by maximizing their contribution margin subject to resource constraints (= primal program). The corresponding dual program then identifies the optimal costs for the company (subject to constraints). In other words, primal and dual program approach the production planning problem from two different directions (production numbers and costs) and thereby offer different economic insights.—Exactly this different view of the dual program is what can be used to reveal implicit economic assumptions. Hodges/Schaefer (1977) apply this view to determine discount factors on bond markets, Wilhelm/Brüning (1992) to term structure models, and we to empirical asset pricing models.

Using Lagrange duality in the context of empirical asset pricing models allows us to circumvent one typical problem with empirical asset pricing models: The primal program—minimization of errors between estimated and actual values of a variable—cannot be interpreted from an economic perspective because error minimization is not an economic concept. The dual program, however, minimizes assets' acquisition costs subject to some constraints. At this point, an economic interpretation becomes apparent since cost minimization is related to the economic principle, i.e., an economic core concept.

Note that the dual program is not needed for pricing.—It just serves as a means to reveal the implicit economic assumptions of the primal program. This is along the lines of Wilhelm/Brüning (1992), where the term structure can only be determined from the primal but not from the dual program (see Wilhelm/Brüning (1992), Formula (26), p. 270). However, only with the help of the dual program economic interpretations become possible.

2.1.2 Economic Dominance of Models

"Economic dominance of models" is our criterion to judge the implicit economic assumptions revealed by computing the dual program and it is developed as follows:

All components of an empirical model (objective functions and constraints) must be considered simultaneously and never the individual components separately because objective function and constraint$_1$ to constraint$_z$ together form $z + 1$ goals to be judged when evaluating models. Transferring results from multi-goal decision theory to economic model evaluation, a model is regarded to dominante another model economically if it is better with respect to at least one goal, but never worse with respect to all other goals. A model that is never better with respect to one goal, but worse with respect to other goals is said to be economically dominated.—Models that are not economically dominated are economically efficient.

To finish the definition of economic dominance of models, the terms "better/worse with respect to goals" must be clarified. "Better/worse" are developed along two lines: (i) the economic principle; (ii) institutional circumstances.

The economic principle tries to achieve a given output with minimum input or obtain with a given input maximum output. The economic principle is very general because it does not depend on investor preferences and wealth situations. In a financial environment the economic principle is reflected in arbitrage theory: Investors strive at acquiring a given cash flow at the lowest possible price, again independent of investor preferences and wealth situations.

In other words, the closer a model follows the idea of the economic principle, the better the model is judged, i.e., the objective function should minimize input and a subset of the constraints should characterize a given output.

"Better/worse" regarding institutional circumstances means that the model's constraints match the actual legal environment and market usages. E.g., if a model has no short sale constraints, such a model is regarded as worse than a model that has such constraints: Uncovered short sales are forbidden for stocks (see Regulation (EU) No 236/2012, Article 12) and, hence, the model is not able to reflect the legal environment properly. Covered short sales are allowed by the EU Regulation. Nevertheless market usages require a certain amount of collateral for securities lending. In other words, there is an implicit upper bound for short sales meaning that a model without short sale constraints also fails to portray market usages adequately.

2.2 Computing Dual Programs and Identifying Their Components

In a first step dual programs are computed for the superordinate categories of models that have been identified in Chapter II. In a second step, their specific components are discussed depending on whether cross section/time series of prices/returns are considered.

2.2.1 Computing Dual Programs

2.2.1.1 Dual Program of the Superordinate Category Regression Approaches

According to Appendix 1.2.4.2, the dual program reads

(4.1)

$$\min_{\lambda_1^+,\lambda_1^-,\dots,\lambda_n^+,\lambda_n^-} \sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^*$$

s.t.

(4.2)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$\begin{aligned} \lambda_1^+ - \lambda_1^- + \cdots + \lambda_n^+ - \lambda_n^- &= 0 \\ \lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* &= 0 \\ &\vdots \\ \lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* &= 0 \end{aligned}$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \dots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$

2.2.1.2 Dual Program of the Superordinate Category Method of Multiples

According to Appendix 1.3.4, the dual program reads

(4.3)

$$\min_{\substack{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-,\\ \beta_1,\ldots,\beta_m}} \sum_{i=1}^{n}(\lambda_i^+ - \lambda_i^-)\cdot y_i^* + \frac{1}{2}\cdot\sum_{j=1}^{m}\beta_j^2$$

s.t.

(4.4)

$$\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$\beta_1 - f(y, A_1) = 0$$

$$\vdots$$

$$\beta_m - f(y, A_m) = 0$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \ldots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$

2.2.2 Components of Dual Programs

To characterize the components of dual programs it is necessary to distinguish between cross-sectional and time series as well as price and return models. All these types of models will result in slightly different interpretations of objective functions and constraints that in turn might influence the economic evaluation of the specific model.

Dependent and explanatory variables can directly be taken from the primal program and, hence, fitted to the cross-sectional/time series/price/return framework. The interpretation of the dual variables in the cross-sectional/time series/price/return framework is less simple because it cannot be taken from the primal program. Hence, it is specified in a way so that the dual program gets the best possible economic interpretation.

2.2.2.1 Prices of Companies in the Cross Section

When prices of companies in the cross section are considered, it is known from Chapter II, Section 2.1.1 that the variable y_i^* in the primal (2.52) and dual (4.1) program is equal to the price of company i at time t. The variables $A_{i,j}^*$ are accounting figures, e.g., $EBIT_{i,t}$ (but not their growth rates) at time t.

Then, the dual variable λ_i^+ (λ_i^-) can be interpreted as the number of purchases (sales) of asset i at time t, i.e., portfolio holdings. The interpretation of the dual variable as portfolio holdings is motivated by the fact that portfolio holdings fit well to a price/accounting figure framework.—Portfolio weights are better suited to a return/growth rate setting.

2.2.2.1.1 Superordinate Category: Regression Approaches

Objective function

With dual variables specified as portfolio holdings, the objective function of the dual program

(4.1)

$$\min_{\lambda_1^+,\lambda_1^-,\dots,\lambda_n^+,\lambda_n^-} \sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^*$$

reads:

(4.5)

$$\min_{N_{1,t}^+,N_{1,t}^-,\dots,N_{n,t}^+,N_{n,t}^-} \sum_{i=1}^{n} \left(N_{i,t}^+ - N_{i,t}^-\right) \cdot P_{i,t}^*$$

The price of a portfolio is to be minimized. The portfolio itself is specified closer with the help of the constraints (4.2).

Constraints on accounting characteristics

The constraint

(4.2)

$$\lambda_1^+ - \lambda_1^- + \cdots + \lambda_n^+ - \lambda_n^- = 0$$

is specified as

(4.6)

$$N_{1,t}^+ - N_{1,t}^- + \cdots + N_{n,t}^+ - N_{n,t}^- = 0$$

and signifies that portfolio holdings must add to zero. In other words, purchases must always be accompanied by short sales to achieve zero investment. Constraint (4.6) should, however, not be confused with a self-financing constraint which states that the amount purchased is equal to the amount sold so that total wealth (and not total portfolio holdings) is equal to zero.

The constraints on portfolio characteristics in the narrower sense

(4.2)

$$\begin{aligned}\lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* &= 0\\ &\vdots\\ \lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* &= 0\end{aligned}$$

reads

(4.7)

$$\begin{aligned}N_{1,t}^+ A_{1,1,t}^* - N_{1,t}^- A_{1,1,t}^* + \cdots + N_{n,t}^+ A_{n,1,t}^* - N_{n,t}^- A_{n,1,t}^* &= 0\\ &\vdots\\ N_{1,t}^+ A_{1,m,t}^* - N_{1,t}^- A_{1,m,t}^* + \cdots + N_{n,t}^+ A_{n,m,t}^* - N_{n,t}^- A_{n,m,t}^* &= 0\end{aligned}$$

It states that in the portfolio each accounting figure $A_{i,t}^*$ must be equal to zero. Note that (4.7) captures the secondary objectives of the decision maker, whereas (4.5) incorporates the primary objective.

<u>Constraint on dual variables</u>

The constraint (part of (4.2)) on dual variables can be specified as

(4.8)

$$\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}\leq x$$

i.e., a constraint on portfolio holdings. Since x is an arbitrary positive scaling factor (see (2.52)), it influences the tightness of the portfolio holdings constraint: The greater x is, the less is (4.8) binding.

2.2.2.1.2 Superordinate Category: Method of Multiples

Objective function

(4.9)

$$\min_{\substack{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-,\\ \beta_1,\ldots,\beta_m}} \sum_{i=1}^{n}(\lambda_i^+-\lambda_i^-)\cdot y_i^*+\frac{1}{2}\cdot\sum_{j=1}^{m}\beta_j^2$$

means that the price of a portfolio including the artefact $\frac{1}{2}\cdot\sum_{j=1}^{m}\beta_j^2$ is minimized.

Constraints on accounting characteristics

Such a constraint does not exist because multiples are determined directly from accounting figures (see Chapter II, Equations (2.23) to (2.27) and Equation (A1.20)) and then are integrated by means of price deviations into the objective function.

Constraint on dual variables

Constraint (4.4) on dual variables

(4.10)

$$\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}\leq x$$

can now be interpreted as a constraint on portfolio holdings. Since x is an arbitrary positive scaling factor (see (2.52)), it influences the tightness of the portfolio holdings constraint: The greater x is, the less is (4.10) binding.

2.2.2.2 Other Cases

2.2.2.2.1 Returns of Companies in the Cross Section

When returns of companies in the cross section are considered, it is known from Chapter II, Section 2.1.1 that the variable y_i^* in the primal (2.52) and dual (4.1) program is equal to the return or the return differential of company i at time t to the riskless rate. The variables $A_{i,j}^*$ are either relative accounting figures or beta factors determined from time series regressions (first step of the two-pass regression).

Then, the dual variable λ_i^+ (λ_i^-) can be interpreted as the portfolio weight of purchases (sales) of asset i. However, the portfolio weights' time subscript is yet to be clarified. To that end, start from investors' (terminal) wealth equation

$$W_{t+1} = N_{1,t} \cdot P_{1,t+1} + N_{2,t} \cdot P_{2,t+1} + \cdots$$

or

$$\frac{W_{t+1}}{W_t} = \frac{N_{1,t} \cdot P_{1,t}}{W_t} \cdot \frac{P_{1,t+1}}{P_{1,t}} + \frac{N_{2,t} \cdot P_{2,t}}{W_t} \cdot \frac{P_{2,t+1}}{P_{2,t}} + \cdots$$

This translates finally to

$$1 + R_{W,t,t+1} = \underbrace{\frac{N_{1,t} \cdot P_{1,t}}{W_t}}_{=w_{1,t}} \cdot \left(1 + R_{1,t,t+1}\right) + \underbrace{\frac{N_{2,t} \cdot P_{2,t}}{W_t}}_{=w_{2,t}} \cdot \left(1 + R_{2,t,t+1}\right) + \cdots$$

In other words, portfolio weights have a lag of one, i.e., the return between times t and $t+1$ is associated with portfolio weights of time t. This in turn means that λ_i^+ (λ_i^-) are the portfolio weights at time $t-1$ if return $R_{i,t}$ is to be explained.

With the now identified variables of the cross-sectional return model its components can be specified.

<u>Objective function</u>

The dual program

(4.1)

$$\min_{\lambda_1^+,\lambda_1^-,\dots,\lambda_n^+,\lambda_n^-} \sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^*$$

reads

(4.11)

$$\min_{w_{1,t-1}^+,w_{1,t-1}^-,\dots,w_{n,t-1}^+,w_{n,t-1}^-} \sum_{i=1}^{n} \left(w_{i,t-1}^+ - w_{i,t-1}^-\right) \cdot R_{i,t-1,t}^*$$

The return of a portfolio of assets i to n is to be minimized.

<u>Constraints on accounting characteristics</u>

The constraint

(4.12)

$$w_{1,t-1}^+ - w_{1,t-1}^- + \cdots + w_{n,t-1}^+ - w_{n,t-1}^- = 0$$

signifies that portfolio weights must add to zero.

The constraints on portfolio characteristics in the narrower sense

(4.13)

$$\begin{aligned} w_{1,t-1}^+ A_{1,1,t}^* - w_{1,t-1}^- A_{1,1,t}^* + \cdots + w_{n,t-1}^+ A_{n,1,t}^* - w_{n,t-1}^- A_{n,1,t}^* &= 0 \\ &\vdots \\ w_{1,t-1}^+ A_{1,m,t}^* - w_{1,t-1}^- A_{1,m,t}^* + \cdots + w_{n,t-1}^+ A_{n,m,t}^* - w_{n,t-1}^- A_{n,m,t}^* &= 0 \end{aligned}$$

state that in the portfolio each accounting figure $A_{i,t}^*$ must be equal to zero.— Again these constraints on accounting figures capture decision makers' secondary objectives.

<u>Constraint on dual variables</u>

The constraint (part of (4.2)) on dual variables

(4.14)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} \left(w_{i,t-1}^{+}\right)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} \left(w_{i,t-1}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

can now be interpreted as a constraint on portfolio weights. Since x is an arbitrary positive scaling factor (see (2.52)), it influences the tightness of the portfolio weights constraint: The greater x is, the less is (4.14) binding.

2.2.2.2.2 Prices of Companies in Time Series

When prices of companies in time series are considered, it is known from Chapter II, Section 2.1.1 that the variable y_i^* in the primal (2.52) and dual (4.1) program is equal to the price of a company i at time τ. The variable $A_{i,j,\tau}^*$ denote accounting figure j of company i at different points in time τ.

Then, the dual variable λ_i^+ (λ_i^-) can be interpreted as the portfolio holdings of purchases (sales) of company i at time τ.

With the now identified variables of the time series price model its components can be specified.

<u>Objective function</u>

The objective function of the dual program (4.1) reads

(4.15)

$$\min_{N_{i,1}^{+}, N_{i,1}^{-}, \ldots, N_{i,t}^{+}, N_{i,t}^{-}} \sum_{\tau=1}^{t} \left(N_{i,\tau}^{+} - N_{i,\tau}^{-}\right) \cdot P_{i,\tau}$$

i.e., the price of company i at time τ multiplied by portfolio holdings at time τ is minimized where time runs from 1 to t.

<u>Constraints on accounting characteristics</u>

The constraint

(4.16)

$$N_{i,1}^{+} - N_{i,1}^{-} + \cdots + N_{i,t}^{+} - N_{i,t}^{-} = 0$$

signifies that portfolio holdings over time (from time 1 to time t) must add to zero.

The constraints on portfolio characteristics in the narrower sense

(4.17)

$$\begin{gathered} N_{i,1}^{+}A_{i,1,1}^{*} - N_{i,1}^{-}A_{i,1,1}^{*} + \cdots + N_{i,t}^{+}A_{i,1,t}^{*} - N_{i,t}^{-}A_{i,1,t}^{*} = 0 \\ \vdots \\ N_{i,1}^{+}A_{i,m,1}^{*} - N_{i,1}^{-}A_{i,m,1}^{*} + \cdots + N_{i,t}^{+}A_{i,m,t}^{*} - N_{i,t}^{-}A_{i,m,t}^{*} = 0 \end{gathered}$$

states that the accounting figure j of company i at time τ multiplied by portfolio holdings at time τ must be equal to zero over time where time runs from 1 to t.

<u>Constraint on dual variables</u>

The constraint (part of (4.2)) on dual variables

(4.18)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{\tau=1}^{t} \left(N_{i,\tau}^{+}\right)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{\tau=1}^{t} \left(N_{i,\tau}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

can now be interpreted as a constraint on portfolio holdings: The sum of portfolio holdings over time is not allowed to exceed x where time runs from 1 to t.

2.2.2.2.3 Returns of Companies in Time Series

When returns of companies in time series are considered, it is known from Chapter II, Section 2.1.1 that the variable y_i^* in the primal (2.52) and dual (4.1) program is equal to the return or return differential to the riskless rate of company i at time τ. The variable $A_{i,j,\tau}^{*}$ denotes returns of factor j at different points in time τ.

Then, the dual variable λ_i^+ (λ_i^-) can be interpreted as the portfolio weight of purchases (sales) of asset i at different point in time τ with a lag of one, i.e., $w_{i,\tau-1}^+$ and $w_{i,\tau-1}^-$.

With the now specified variables of the time series return model its components can be specified.

Objective function

The objective function of the dual program (4.1) reads

(4.19)

$$\min_{w_{i,0}^+, w_{i,0}^-, \ldots, w_{i,t-1}^+, w_{i,t-1}^-} \sum_{\tau=1}^{t} \left(w_{i,\tau-1}^+ - w_{i,\tau-1}^-\right) \cdot \left(R_{i,\tau} - r_\tau\right)$$

i.e., the return differential to the riskless rate of company i at time τ multiplied by portfolio weights at time $\tau - 1$ is minimized where time runs from 0 to $t - 1$ for portfolio weights and from 1 to t for returns.

Constraints on accounting characteristics

The constraint

(4.20)

$$w_{i,0}^+ - w_{i,0}^- + \cdots + w_{i,t-1}^+ - w_{i,t-1}^- = 0$$

signifies that portfolio weights must add to zero over time.

The constraints on portfolio characteristics in the narrower sense

(4.21)

$$\begin{aligned} w_{i,0}^+ A_{1,1}^* - w_{i,0}^- A_{1,1}^* + \cdots + w_{i,t-1}^+ A_{1,t}^* - w_{i,t-1}^- A_{1,t}^* &= 0 \\ &\vdots \\ w_{i,0}^+ A_{m,1}^* - w_{i,0}^- A_{m,1}^* + \cdots + w_{i,t-1}^+ A_{m,t}^* - w_{i,t-1}^- A_{m,t}^* &= 0 \end{aligned}$$

states that the weighted factor return j must be equal to zero over time.

Note in this connection that it does not matter whether discrete or log return/growth rates (as an example, consider the variable MP(t) in Chen/Roll/Ross (1986), p. 394 that is defined as the logarithm of the quotient of industrial production) are used. Each explanatory variable gets its own equation in (4.21) and, hence, different definitions of growth rates are not mixed in one equation.

<u>Constraint on dual variables</u>

The constraint (part of (4.2)) on dual variables

(4.22)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{\tau=1}^{t} \left(w_{i,\tau-1}^{+}\right)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{\tau=1}^{t} \left(w_{i,\tau-1}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

can now be interpreted as a constraint on portfolio weights: The sum of portfolio weights over time is not allowed to exceed x where time runs from 0 to $t-1$.

2.3 Specification of the Economic Model Evaluation Criterion

To be able to apply the economic model evaluation criterion as a benchmark, i.e., as a collection of features that models should possess, it becomes necessary to specify it in more detail.—Such a specification was impossible in Section 2.1 because there it has been unclear that, e.g., a constraint on portfolio holdings in the dual program exists that must then be judged from an economic point of view.

The specification of the economic model evaluation criterion develops along two lines. First, economic principle and institutional circumstances must be specified and connections with the components of the dual program (objective function, constraints on, e.g., accounting figures, constraints on, e.g., portfolio holdings) identified. That way, an economically convincing model can be established. Second, economic dominance of models can actually be tested and a ranking of models can be carried out. In this connection, a relative (how good are models compared to each other?) and an absolute ranking (are models acceptable at all?) is provided.

2.3.1 Specification of the Economic Principle

The economic principle simply means that a given output should be obtained by means of a minimum input or with a given input a maximum output should be generated. In other words, goods should be bought as cheap as possible or sold as expensive as possible. Such an approach is sometimes called “arbitration” (see Munn (1983)). It is different from the free lunch of Harrison/Kreps (1979) where investors are not interested in acquiring or selling a physical position. Instead, investors form a difference arbitrage to obtain a positive cash flow in at least one point in time and state without requiring a negative cash flow in all other time and state combinations.

If the economic principle is specified to the context of buying/selling companies, it reads in more detail:

- For the buyer of a company

  A buyer invests cash flow at time t to acquire the company and gets a future cash flow stream in exchange. Alternatively, the cash flow stream can be expressed with the help of multi-period returns or approximated by accounting characteristics at time t in a sense that accounting figures are observable and are used as proxy for the unobservable future cash flow stream.

  From that perspective there are two objectives for the buyer: (i) wealth at time t and (ii) accounting characteristics at time t or cash flow at $t+1$. Both objectives are conflicting as a rule: A low investment at t leads to low cash flows at time $t+1$ or is accompanied by low accounting characteristics at time t (otherwise the investment would not be low); a high investment at time t leads to high cash flows at time $t+1$ or is associated with high accounting characteristics at time t. To deal with this conflict of interests, the maximum principle of efficiency is applied, i.e., one objective (primary objective) is maximized subject to lower bounds on the other objectives (secondary objectives). Maximizing the primary objective subject to lower bounds on the secondary objectives guarantees that only efficient alternatives will be selected. The primary objective is associated with wealth at time t, the secondary objectives with the different accounting characteristics/cash flows. A buyer consequently maximizes wealth at time t, which signifies minimizing the (purchase) price of the company, subject to accounting characteristics/cash flows greater than or equal to a lower bound.

– For the seller of a company

  Identically to the buyer, the seller of a company is interested in wealth at time t and accounting characteristics at time t/a cash flow stream. However, maximizing wealth at time t means that the (sales) price of the company should be as high as possible; the accounting characteristics/cash flow stream constraint must be formulated in a way so that the seller loses as few accounting characteristics/cash flows as possible, i.e., accounting characteristics/cash flows should be less than or equal to an upper bound.

From this specification of the economic principle several consequences regarding the formulation of an economic model evaluation criterion follow. When judging models from an economic perspective,

(i) objective function and constraints must be considered simultaneously and not separately because they both together constitute decision makers' objectives.

(ii) an actual company must be modelled (that is either be purchased or sold), i.e., arbitration and not free lunch must be pursued.[5]

(iii) buyers' and sellers' point of views must be reflected. Both points of view cannot be transformed into each other. Setting $\text{purchase} = -\text{sale}$ leads to

[5] By the way, whenever arbitration is implemented using all available assets on the market, it will automatically be tested for a free lunch at time t since an accounting characteristic/cash flow bound of zero of the free lunch is a special case of the lower bound on accounting characteristic/cash flow in the case of arbitration.

identical objective functions. However, this variable transformation cannot adequately reflect the accounting characteristics/cash flow constraint, namely that the buyer wants to obtain accounting characteristics/cash flows greater than or equal to and the seller wants to abandon less than or equal to a certain bound.[6]

2.3.2 Institutional Circumstances

Institutional circumstances refer to the constraints on portfolio holdings/weights and can be decomposed into two groups: (i) legal environment and (ii) market usages.

The legal environment forbids uncovered short sales of stocks (see Regulation (EU) No 236/2012, Article 12). Of course, any assets already in possession can be sold and covered short sales are allowed by the EU Regulation. In addition, there are no limits on purchases of stocks assuming that companies are not subject to capital adequacy regulation.

The legal environment, thus, imposes the following constraint on portfolio holdings: The sales of stock i must be less than or equal to a lower bound, i.e.,

$$N_{i,t}^{-} \leq x_{i,t}^{-} \text{ for } i = 1, \dots n$$

[6] To see this, note:

- Decision problem from the buyer's perspective:
  $\min_{N_U} N_{U,t} \cdot P_{U,t}$
  s.t. $N_{U,t} \cdot A_{U,t} \geq \bar{A}_t$
  where $N_{U,t}$ denotes the numbers of company U purchased at time t, $P_{U,t}$ the price of company U at time t, $A_{U,t}$ company U's accounting characteristic at time t, and $\bar{A}_t$ the lower bound on the accounting characteristic at time t.
- Decision problem from the seller's perspective:
  $\max_{N_{S,U}} N_{S,U,t} \cdot P_{U,t}$
  s.t. $N_{S,U,t} \cdot A_{U,t} \leq \bar{A}_t$
  where $N_{S,U,t}$ denotes the numbers of company U sold at time t.

Setting $N_{U,t} = N_{S,U,t}$ makes the objective functions coincide. However, the constraint will be different.

There is no explicit legal limit on portfolio weights because portfolio weights depend on investor-specific wealth. However, limits on portfolio holdings can be re-expressed as limits on weights as follows:

$$\underbrace{\frac{N_{i,t}^{-}\cdot P_{i,t}}{W_t}}_{=w_{i,t}^{-}} \leq \frac{x_{i,t}^{-}\cdot P_{i,t}}{W_t} \text{ for } i=1,\dots n$$

Market usages might impose an upper limit on covered short sales because they require a certain amount of collateral for securities lending. Such a limit is again imposed on each asset i. Depending on the overall amount of short sales of individual investors, individual investors might be confronted with different limits, but it seems to be safe to argue that limits nevertheless will be imposed on individual assets.

Moreover, market usages imply that there are market impact costs, i.e., large amounts of purchases and/or (short) sales influence market prices. Market impact costs might be linear (e.g., Kyle (1985)) or nonlinear in the amount traded (e.g., Almgren/Thum/Hauptmann/Li (2005), Grinold (2006), and Gatheral (2010)). The models, however, agree that market impact costs are asset-specific and depend on the sign of the transaction (purchase or sale).

Finally, constraints on portfolio holdings are needed for technical reasons, namely to find an optimal solution to the dual program. To satisfy the m constraints on accounting characteristics, m assets are needed. The remaining $n-m$ assets can be used to obtain the m accounting characteristics at a more and more negative price, i.e., to obtain an arbitrage profit. Constraints on portfolio holdings exactly limit these arbitrage profits.

2.3.3 Relative and Absolutes Ranking of Models

When evaluating investment projects using net present value, usually a two-step procedure is applied. In a first step, all investment project are eliminated whose net present value is negative, i.e., that are from an absolute perspective disadvantageous. In a second step, the remaining investment projects are ranked based on the size of their positive net present value.

When ranking models we employ a similar procedure. Absolute ranking means judging whether the models are acceptable when compared to the model evaluation criterion (1) economic principle (Section 2.3.1 (i) to (iii)) and (2) institutional circumstances (Section 2.3.2). In the context of absolute ranking it only matters whether the criterion is met or not, the exact extent does not matter. Relative ranking answers the question how good models are compared to each other. Here, the exact extent matters of how far the model evaluation criterion (1) economic principle (Section 2.3.1 (i) to (iii)) and (2) institutional circumstances (Section 2.3.2) is met.

3 Applying the Economic Model Evaluation Criterion

Intuitively, applying the economic model evaluation criterion means that the best models are those that use the most innocuous assumptions.

3.1 Absolute Ranking of Empirical Asset Pricing Models: Cross Section of Prices

3.1.1 Model Evaluation Criterion Economic Principle: Section 2.3.1 (i) to (iii)

This model evaluation criterion comprises objective function ((4.5) for regression approaches and (4.9) for the method of multiples) and accounting constraints ((4.7) for regression approaches and none for the method of multiples) of the dual program.

First, both regression approaches and the method of multiples address the pricing problem in an indirect way: They determine regression coefficients or multiples from a sample of companies by minimizing prices and apply them to the company to be valued. From that perspective they do not minimize directly the price of the company to be valued, as the criterion in Section 2.3.1 (i) suggests. Note that such a behavior can be interpreted in parallel to Law-of-One-Price-oriented pricing. There, a price functional is determined from a subset of assets and applied to the cash flow to be valued. Here, regression coefficients/multiples are determined from a subset of assets and applied to the company to be valued.

Second, minimizing prices subject to accounting constraints take all (primary and secondary) objectives into account and not just the primary objective. However, this indirect pricing is also responsible for the fact that the idea of a free lunch and not an arbitration is followed: "accounting characteristics of the portfolio = 0" is used together with price minimization of the portfolio because the investor is not interested in investing in a company, but only in generating an arbitrage profit. Therefore, this pricing approach is not completely compelling in the light of the criterion in Section 2.3.1 (ii) because it overlooks that additional gains might be possible form arbitration. Multiples have an artefact $\frac{1}{2} \cdot \sum_{j=1}^{m} \beta_j^2$ in the objective function that is incompatible with the economic principle because it does not solely minimize prices.

Third, by minimizing prices to determine regression coefficients, regression approaches take the buyers' perspective. They cannot handle the sellers' perspective and, hence, cannot deal with the criterion in Section 2.3.1 (iii). Note in this connection that different weights τ on over- and underestimations in the primal program cannot capture buyers' and sellers' perspective because they do not enter the objective function of the dual program. The method of multiples determines multiples by averaging over a group of companies. Therefore, it does not explicitly take the buyers' perspective. However, it cannot take the sellers' perspective either and, thus, cannot handle the criterion in Section 2.3.1 (iii), too.

3.1.2 Model Evaluation Criterion Institutional Circumstances: Section 2.3.2

This model evaluation criterion comprises constraints on portfolio holdings, namely (4.6) and (4.8) for regression approaches and (4.10) for the method of multiples.

Constraint (4.6) demands that the sum of portfolio holdings over all assets of the portfolio must be equal to zero.—This should not be confused with a self-financing constraint where investments are funded by sales. Here, just holdings not holdings multiplied by price (= investments) are considered. This constraint stems from the constant β_0 of the regression. It cannot, however, be justified from either (i) legal environment or (ii) market usages and, hence, clearly violates the model evaluation criterion institutional circumstances in Section 2.3.2.

The constraints on portfolio holdings (4.8) and (4.10) are difficult to justify with the help of the legal environment, i.e., short sale constraints, because these constraints are not based on individual assets. In addition, market usages cannot rationalize such a constraint. It is true that only moderate orders will be executed at a given price. But market impact refers to transactions in individual asset i and usually not to transactions in all assets. Therefore, these constraints on portfolio holdings can be justified merely because they limit arbitrage profits.—In reality infinite arbitrage profits are not observable.

In this connection, the question arises as to what type of constraint on portfolio holdings results in the lowest price, i.e., allows the highest arbitrage profit? If we define $q \equiv \frac{p}{p-1}$, then monotonicity of L_p-norms implies that $\|\cdot\|_{q_2} \leq \|\cdot\|_{q_1}$ for $q_1 < q_2$. In other words, for q_2 the constraint on portfolio holdings (4.8) is less restrictive since the left-hand side is less and the right-hand side is constant. Moreover, q falls with increasing p (for $p > 1$, as can be seen from its first derivative). This means, an increase in p leads to a fall in q making $\|\cdot\|_q$ greater and the constraint on portfolio holdings (4.8) more binding. A more binding constraint translates into higher prices, i.e., lower arbitrage profits. E.g., ordinary least squares regression $p = 2$ results in lower prices/higher arbitrage profits than minimizing the maximum error ($p = \infty$).

3.1.3 Absolute Ranking of Empirical Asset Pricing Models: Cross Section of Prices

Having the results of the model evaluation criterion (i) economic principle (Section 3.1.1) and (ii) institutional circumstances (Section 3.1.2) in mind, regression models using cross section of prices can be regarded as barely acceptable from an absolute ranking perspective: Indirect pricing, only consideration of free lunches, but not arbitration, and focus on buyers' perspective mean that the economic principle is not fully implemented. On the other hand, at least the idea of minimizing prices (subject to accounting constraints) is captured. The constraints on portfolio holdings cannot be justified fully by means of short shelling constraints, but partially from the perspective of limited arbitrage profits.

Only constraint (4.6), the sum of portfolio holdings over all assets must be equal to zero, cannot at all be justified by means of institutional circumstances.—However, as Cochrane (2005), p. 236 points out cross-sectional regression can also run without a constant meaning that this detrimental constraint can be removed.

The method of multiples is not acceptable from an absolute ranking perspective since there is an artefact in its objective function signifying that not solely prices are minimized.

3.2 Absolute Ranking of Other Model Categories

3.2.1 Cross Section of Returns

Principally the results derived for cross section of prices carry over for cross section of returns. However, the objective function

(4.11)

$$\min_{w_{1,t-1}^{+},w_{1,t-1}^{-},\dots,w_{n,t-1}^{+},w_{n,t-1}^{-}} \sum_{i=1}^{n} \left(w_{i,t-1}^{+} - w_{i,t-1}^{-}\right) \cdot R_{i,t-1,t}^{*}$$

is new and must be analyzed in more detail.

Plugging in for portfolio weights and returns, (4.11) modifies to

$$\min_{N_{1,t-1}^{+},N_{1,t-1}^{-},\dots,N_{n,t-1}^{+},N_{n,t-1}^{-}} \sum_{i=1}^{n} \left(\frac{N_{i,t-1}^{+} \cdot P_{i,t-1}^{*}}{W_{t-1}} - \frac{N_{i,t-1}^{-} \cdot P_{i,t-1}^{*}}{W_{t-1}}\right) \cdot \left(\frac{P_{i,t}^{*}}{P_{i,t-1}^{*}} - 1\right)$$

i.e.,

(4.23)

$$\min_{N_{1,t-1}^{+},N_{1,t-1}^{-},\dots,N_{n,t-1}^{+},N_{n,t-1}^{-}} \sum_{i=1}^{n} \frac{N_{i,t-1}^{+} - N_{i,t-1}^{-}}{W_{t-1}} \cdot P_{i,t}^{*} - \sum_{i=1}^{n} \left(\frac{N_{i,t-1}^{+} \cdot P_{i,t-1}}{W_{t-1}} - \frac{N_{i,t-1}^{-} \cdot P_{i,t-1}}{W_{t-1}}\right) \cdot 1$$

According to (4.23) the price of a portfolio is minimized where there is a certain time lag: Portfolio holdings at $t-1$ are multiplied—at least in the first term—with prices at time t. However, prices at time t are not observable at time $t-1$. Hence, (4.23) implies that prices do not change between $t-1$ and t if it should possess a reasonable economic interpretation.

This time lag or rather the assumption of constant prices between $t-1$ and t is unrealistic. For that reason, cross section of returns is not acceptable from an absolute ranking perspective. Put differently, (4.23) gives a theoretical justification of Barth/Beaver/Landsman's (2001) explanation that price studies are interested in determining what is reflected in firm value while return studies (price changes) are interested in determining what is reflected in change in value over a specific period of time.

Side note: This result does not come at a surprise if minimization of returns is considered from a no-arbitrage perspective. Approaches using Law-of-One-Price-oriented pricing, like Black/Scholes (1973) argue as follows: A riskless portfolio that consists of an option and a risky asset must earn the same return as a riskless asset.—No return optimization is involved, just the application of the Law-of-One-Price. Utility-oriented pricing like Cox/Ingersoll/Ross (1985) makes statements regarding required return in a maximizing expected utility framework. Again, no return optimization or constructing arbitration is employed.

3.2.2 Time Series Models

In our analysis of time series models we do not distinguish between price and return models as will soon become clear.

In time series models prices/returns of one asset i ((4.15) or (4.19)) are minimized over time subject to constraints on accounting characteristics of one asset i ((4.17) or (4.21)) over time. The constraint on portfolio holdings/weights ((4.18) or (4.22)) also refers to sums of portfolio holdings/weights over time.

Analyzing prices and accounting characteristics/returns of one asset i over time is, however, incompatible with the economic principle. The economic principle constructs arbitrages at one point in time using several assets and not one asset over time. Moreover, limits on portfolio holdings/weights of one asset i over time do not coincide with the restriction that institutional circumstances impose.

For that reason, time series price and return models are not acceptable from an absolute ranking perspective.

3.3 Relative Ranking of Empirical Asset Pricing Models: Cross Section of Prices

Given the results of the absolute ranking, only cross-sectional price regressions are (barely) acceptable using the economic model evaluation criterion. Therefore, only this approach will be analyzed from a relative ranking perspective. The method of multiples as well as return and time series regression approaches will be left out since they did not pass the absolute ranking.

To establish a relative ranking, it is proceeded as follows. In a first step, models are identified that differ just between their respective objective functions, but possess identical constraints. Since these models are different only with one component (objective function), they can easily be examined regarding economic model-dominance: A model whose objective function is better than the one of other models regarding the economic principle, is model-dominant, the other models are model-dominated and, hence, model-inefficient. In a second step, the set of efficient models is further analyzed with respect to their constraints. Since the objective functions of each class of efficient models coincide by construction, only one component (constraint) must be analyzed to check for model-dominance in the following sense: A model whose constraint on portfolio holdings is better than the one of other models regarding institutional circumstances, is model-dominant, the other models are model-dominated and, hence, model-inefficient.

3.3.1 First Step: Testing Models with Transformed and Untransformed Dependent Variables

Models with different objective functions but identical constraints can be found by distinguishing between models with transformed ($P^*_{i,t}$) and non-transformed ($P_{i,t}$) dependent variables. Models with transformed dependent variables($P^*_{i,t}$) contain weighted least squares regression/percentage error, error measures with logarithmic error, and generalized least squares regression. Models with untransformed dependent variables ($P_{i,t}$) comprise all other regression models. With respect to constraints, all models are subject to (4.6) to (4.8), i.e., underlie the same set of constraints. In this connection note that the constraints on accounting characteristics (4.7) sometimes depend on transformed ($A^*_{i,j,t}$) and sometimes depend on untransformed ($A_{i,j,t}$) variables. This does not make a difference from the perspective of the economic principle: All that matters is that secondary objectives are captured by means of accounting constraints that demand that each accounting figure in the portfolio is equal to zero (idea of a free lunch). The exact form of this accounting constraint, i.e., transformed or untransformed accounting figures, does not matter.

Table 4.1 indicates that the objective function of untransformed models is superior to the objective function of transformed models. Since all models are identical with respect to constraints (4.6) to (4.8), untransformed models dominate according to economic model-dominance transformed models.

Model	Objective function	Economic evaluation
Non-transformed L_p-norm	$\min\limits_{N_{1,t}^+,N_{1,t}^-,\ldots,N_{n,t}^+,N_{n,t}^-} \sum_{i=1}^{n} \left(N_{i,t}^+ - N_{i,t}^-\right) \cdot P_{i,t}$ (4.5) specified with respect to non-transformed dependent variables	Compatible with the economic principle.
Weighted least squares regres-sion and per-centage error	$\min\limits_{N_{1,t}^+,N_{1,t}^-,\ldots,N_{n,t}^+,N_{n,t}^-} \sum_{i=1}^{n} \left(N_{i,t}^+ - N_{i,t}^-\right) \cdot P_{i,t} \cdot \underbrace{\omega_i}_{=\frac{1}{P_{i,t}}}$ (4.5) in connection with Chapter II Section 3.1.2	Incompatible with the economic principle: Prices of portfolios and not portfolio holdings should be minimized.
Error measures with logarithmic error	$\min\limits_{N_{1,t}^+,N_{1,t}^-,\ldots,N_{n,t}^+,N_{n,t}^-} \sum_{i=1}^{n} \left(N_{i,t}^+ - N_{i,t}^-\right) \cdot \ln\left(P_{i,t}\right)$ (4.5) in connection with Chapter II Section 3.1.2	Incompatible with the economic principle: Prices of portfolios and not portfolios' logarithmic prices should be minimized.
Generalized least squares regres-sion	$\min\limits_{N_{1,t}^+,N_{1,t}^-,\ldots,N_{n,t}^+,N_{n,t}^-} \sum_{i=1}^{n} \left(N_{i,t}^+ - N_{i,t}^-\right) \cdot P_{i,t} \cdot \omega_i$ (4.5) in connection with Chapter II Section 3.1.2	Incompatible with the economic principle: Prices of portfolios and not weighted prices should be minimized.—One cannot acquire assets at weighted prices but just at (market) prices.

Table 4.1: Objective functions of several models and their evaluation according to the economic principle

Put differently, using statistically more advanced models deteriorates the implied economic content of models although they might improve the statistical quality. Therefore, the recommendation to use generalized least squares R^2 (see Lewellen/Nagel/Shanken (2010), p. 183) goes into the wrong direction from an economic point of view.

3.3.2 Second Step: Testing the Subset of Efficient Models by Specifying the L_p-norm

The only model-efficient class is the class of models with untransformed dependent variables. Since all models of this class are based on the same objectives, i.e., objective function and accounting constraints, they differ only with respect to the constraint on portfolio holdings. This constraint in turn depends on the L_p-norm chosen (see (4.8)). In particular, quantile regression, ($p = 1$), ordinary least squares regression ($p = 2$), and L_p-regression ($p = \text{unspecified}$) are analyzed.

Table 4.2 indicates that the constraint on portfolio holdings of quantile regression is superior to the one of the other two models. Since all models are identical with respect to the objective function (primary objective) and the accounting constraints (secondary objective), quantile regression model-dominates ordinary least squares and L_p-norm regressions.

Finally note that the economic content of quantile regression can be further improved if it is run without a constant to avoid the problematic constraint

(4.6)

$$N_{1,t}^{+} - N_{1,t}^{-} + \cdots + N_{n,t}^{+} - N_{n,t}^{-} = 0$$

Cochrane (2005), p. 236 points out cross-sectional regressions can be run without such a constant.

Model	Constraint on portfolio holdings	Economic evaluation
Quantile regression	$N_{1,t}^{+} \leq x \cdot (1-\tau)$ $N_{1,t}^{-} \leq x \cdot \tau$ $\vdots$ $N_{n,t}^{+} \leq x \cdot (1-\tau)$ $N_{n,t}^{-} \leq x \cdot \tau$ (A1.19)	(Partially) fits institutional circumstances because such constraints are compatible with legal environment and market usages: Short sale and market impact constraints refer to individual assets. It is, however, not completely clear that the upper bound on purchases adequately reflects market impact cost.
Ordinary least squares regression	$\left[\sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{2}+\sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{2}\right]^{\frac{1}{2}} \leq x$ (4.8) in connection with Chapter II Section 3.1.2.1	Does not fit institutional circumstances because such constraints are incompatible with legal environment and market usages: Short sale and market impact constraints refer to individual assets. Moreover, short sale constraints are not quadratic.
L_p-norm	$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$ (4.8)	Does not fit institutional circumstances because such constraints are incompatible with legal environment and market usages: Short sale and market impact constraints refer to individual assets. Moreover, short sale constraints are not based on the $\frac{p}{p-1}$ th root.

Table 4.2: Constraints on portfolio holdings of several models and their evaluation according to institutional circumstances

4 Conclusion

Harvey (2017), p. 1413 argues that trying different empirical models can be regarded as one form of p-hacking. The American Statistical Association (2016) points out that business decisions should not be based only on whether a p-value passes a specific threshold. Moreover, Chapter III shows that there are large differences in corporate values when different empirical models are applied so that virtually arbitrary corporate values can be justified. Nietert/Otto (2018) demonstrate that the same is true if the method of multiples is used to compute company values.—There is a clear need for an economic model evaluation criterion.

Given this high need for an economic model evaluation criterion, the objective of this Chapter IV is twofold: (i) first develop an economic model evaluation criterion; (ii) come up with an economic ranking of different empirical models.

The results of this chapter can be summarized as follows:

First, the economic model evaluation criterion judges the implicit economic assumptions revealed by computing the dual program along the two dimensions compliance with the economic principle and institutional circumstances.

Second, applying the economic evaluation criterion to empirical models reveals that regressions on cross section of prices can be regarded as acceptable from an economic perspective, whereas regressions on cross section of returns and time series models as well as the method of multiples do not comply with the economic principle.

Third, within the group of cross-sectional price models quantile regression proves to be the best model because it is able to offer a good approximation to the economic principle and mimics best the institutional circumstances, in particular, if the regression is run without a constant. On the other hand, statistically more advanced models like generalized least squares regression deteriorates the implied economic content of models: They work with weighted prices; however assets can only be purchased and sold at (unweighted) prices.

Still, even the best empirical model, quantile regression, is not fully convincing regarding the economic principle. In order to find an economically fully convincing empirical model, it might be wiser to start from economic principle and institutional circumstances and develop a new model instead of trying to adjust existing empirical models to economic principle and institutional circumstances.—This will be done in Chapter V.

Chapter V: An Accounting-Based Empirical Business Valuation Model

1 Introduction

From the perspective of asset pricing theory the correct business valuation model is clear: the present value of future cash flows must be used for valuation because it follows from no arbitrage, i.e., the intuition that a positive investment is needed to generate a positive return. Practically, forecasts of the future are difficult and, in particular, the determination of discount factors proves problematic as the literature overview in Aggarwal/Mishra/Wilson (2018) illustrates. Therefore, the industry introduced two main simplifications into present value computations as can been inferred from textbook formulas (e.g., Brealey/Myers/Allen (2016), p. 497, Damodaran (2006), p. 383, Damodaran (2012), p. 386, Berk/ DeMarzo (2017), p. 323, Ross/Westerfield/Jaffe/Roberts (2015), p. 390, and Ross/Westerfield/Jordan (2015), p. 411): time-constant discount rates (WACC) and horizon values to capture the far distant future.—However, time-constant discount rates cannot be transferred seamlessly to multi-period discounting. First, Fama (1977) shows that they neglect non-flat term structures and, second, they overlook stochastically changing investment opportunity sets. Third, Fama (1996) finds that one-period returns behave differently than multi-period returns meaning that they exhibit a different risk and, thus require different discounting. Fourth, Fama/French (1997) conclude that empirical cost of capital estimates are imprecise for three reasons: (i) difficulties in identifying the right asset pricing model; (ii) imprecisions in estimating factor loadings; (iii) imprecisions in estimating factor risk premia. Finally, the horizon model, by definition, creates some imprecision because a stable growth after n periods is assumed together with a flat term structure.—Therefore, the practical implementation of present value approaches is less theoretically stringent as desired.

If, however, a superior theoretical model—present value—cannot be implemented adequately, it might be better to use a theoretically less convincing model—e.g., use of accounting characteristics—that creates less problems with its application. This impression is supported by the following observations from valuation practice: Imam/Barker/Clubb (2008) find that discounted cash flow models have become significantly more important in valuation practice than prior survey evidence suggests, e.g., Demirakos/Strong/Walker (2004). But still valuation multiples, notably the price/earnings ratio, are used. Peasnell/Yin (2014) stress the still important role of multiples in investment research reports of U.S. firms issued by analysts of leading brokerage firms in 2011–2012. Tan/Yu (2018) support a trend in the intensified use of discounted cash flow models, but also find that discounted cash flows are only used in 21% of all valuation cases. Finally, residual income valuations are still rarely used in analysts' reports (Hand/Coyne/Green/Zhang (2017)). The reasons for the use of accounting-based valuation methods are, on the one hand, the complexity of discounted cash flow approaches (Damodaran (2006), Imam/Barker/Clubb (2008), and Tan/Yu (2018)). On the other hand, forecasting arguments matter: Imam/Barker/Clubb (2008) point out (p. 515) that valuation models are seen as complementary to each other since analysts need to use subjective methods that deliver prices that feel right (p. 503). Tan/Yu (2018) come to the conclusion that analysts are more likely to use discounted cash flow models if earnings quality is low due to earnings management or earnings are negative.

The superior practicability of existing accounting-based valuations, however, is bought with a relatively weak foundation in asset pricing theory:

(i) Multiples

Multiples essentially argue that similar accounting characteristics should result in similar prices.

Problems from the perspective of asset pricing theory: While such a valuation statement is intuitive, it is not backed up by asset pricing/arbitrage theory that states: Identical cash flow streams must possess identical prices. In other words, there are three differences between multiples and arbitrage theory. First, accounting characteristics are considered instead of cash flow streams. Second, similar instead of identical positions are examined. Third, one accounting characteristic is regarded as enough to characterize a company completely. Only the third problem has been addressed to some degree by the literature by averaging valuation results for several accounting characteristics (e.g., EBIT and sales) because different accounting characteristics translate into different company prices, see, e.g., Beatty/Riffe/Thompson (1999), Cheng/McNamara (2000), and Schreiner (2007). However, business valuations using simultaneously several accounting characteristics do not exist.

(ii) Implementing discounted cash flow models with the help of accounting characteristics

Berk/DeMarzo (2017) or Brealey/Myers/Allen (2016) use multiples to estimate the horizon value of discounted cash flow models. Residual income valuation models (Feltham/Ohlson (1995), Ohlson (1995), and Ohlson (2005)) express cash flows by means of earnings where a function of earnings is discounted using a riskless rate. The most integrated approaches of discounted cash flow and accounting-based models (Claus/Thomas (2001), Easton (2004), Gebhardt/Lee/Swaminathan (2001), and Ohlson/Juettner-Nauroth (2005)) also express the discount rate as a function of earnings or their growth rates (and not just cash flows).

Problems from the perspective of asset pricing theory: Residual income models focus on the numerator of discounted cash flow models, i.e., they strive at expressing expected cash flows with the help of accounting characteristics. The denominator, the discount rate, is still characterized by constant cost of capital. Therefore, Easton (2004), Gebhardt/Lee/Swaminathan (2001), and Ohlson/Juettner-Nauroth (2005) cannot address Fama's (1977) and Fama's (1996) criticism of constant cost of capital. Claus/Thomas (2001) at least use a non-flat term structure, but have to assume constant risk premia, an assumption that does not hold in reality. Finally, Hand/Coyne/Green/Zhang (2017) find less drastic, but still remarkable empirical valuation differences between discounted cash flow and accounting-based residual income valuations.

(iii) Empirical accounting-based approaches

Empirical accounting-based approaches explain stock prices with the help of accounting characteristics (see, e.g., Appendix 4 for an overview).

Problems from the perspective of asset pricing theory: These empirical accounting-based approaches belong to the field of value relevance studies and, thus, are only interested in statistical significance of accounting characteristics, but not economic significance, i.e., they do not derive pricing statements. In principle, the regression coefficients of value relevance studies can also be used to obtain business values. However, Chapter III showed that valuation differences between different regression methods are huge. Chapter IV demonstrated that regression models have a weak economic backing when contrasted with the economic principle.

In summary, there seems to be a trade-off between asset pricing rigor and practicability of models. Present value models are theoretically superior, but their practical implementation in form of constant discount rates and horizon models is far from economically convincing. Accounting-based models are characterized by less asset pricing theory rigor, however, can be implemented without sacrificing much of their theoretical basis, in particular empirical accounting-based asset pricing models. Obtaining better asset pricing models, hence, means either improve the implementation of present value models or the theoretical foundations of accounting-based models. Given the sheer amount of valuation models or heuristics that use accounting data as input (see Cascino/Clatworthy/García Osma/Gassen/Imam/Thomas (2014), p. 191), we would like to improve the asset pricing foundation of accounting-based models, in particular, empirical accounting-based models for two reasons. On the one hand, the accounting literature so far has not fully exploited the asset pricing potential of accounting-based valuation models: It can be increased visibly without sacrificing practicability. On the other hand, purely empirical models always create a justification problem: Who would pay a higher price for a company because sales multiples result in higher prices than earnings multiples? Who would pay a higher price for a company because a lower discount rate for earnings is used? Who would pay a higher price for a company because an empirical estimation procedure, which possesses a higher R^2, recommends a higher price than other empirical estimation procedures?

Therefore, it is the objective of this Chapter V to connect the practicability of accounting-based valuation models with the theoretical rigor of asset pricing theory.

To achieve this objective, two steps are applied. First, the valuation approach of arbitrage theory/economic principle is transferred to the problem of business valuation: optimize the price of the company subject to constraints on accounting characteristics. Second, the optimize-the-price approach is compared to regression approaches to elaborate the economic significance of value differences both theoretically and empirically.

The results of this chapter can be summarized as follows: From a theoretical perspective, the optimize-the-price approach is based on the economic principle and is able to integrate constraints on portfolio holdings that are in line with the institutional environment and market usages. Moreover, the optimize-the-price approach can distinguish between buyers' and sellers' position, use the mispricing potential of the company to be valued (arbitration, Munn (1983)) instead of focusing only on mispricing of other companies (free lunch), and can integrate synergies, multi-period valuations as well as risk. From an empirical perspective, the price differences between the integrated (optimize-the-price approaches) and the separated approaches (regressions) as well as price differences between buyers and sellers are of very high economic significance.

This chapter makes the following contribution compared to the literature:

First, it offers a completely different approach on accounting-based valuation, a direct optimization of the buyer's/seller's price. The theoretical accounting literature so far, in particular residual income valuation models, adapted the discounted cash flow approach by expressing cash flows and/or discount rates with the help of accounting characteristics (Feltham/Ohlson (1995), Ohlson (1995), Ohlson (2005) as well as Claus/Thomas (2001), Easton (2004), Gebhardt/Lee/ Swaminathan (2001), and Ohlson/Juettner-Nauroth (2005)). Since the optimize-the-price approach rests upon the economic principle, it holds for a much broader spectrum of preferences and thus, decision makers, than the discounted cash flow approaches that usually rely on CAPM derivatives, i.e., μ-σ-preferences. One interesting side aspect of the optimize-the-price approach deserves mentioning. Aggarwal/Mishra/Wilson (2018) illustrate that the determination of discount factors is the most critical part when applying discounted cash flow models. The optimize-the-price approach does not need to determine discount factors because prices are determined directly, a procedure that is typical for the determination of price functionals in no arbitrage theory (see, e.g., Ingersoll (1987), p. 29).

Second, we integrate the several factors from value relevance studies into an accounting-based valuation formula. That way, value relevance studies, which are not interested in valuation but use several accounting characteristics as explanatory variables, are combined with multiples, which can price, but can deal only with one accounting characteristic at the same time. In this connection it is important to note that a good business valuation approach does not mean reproducing market prices best possible. If a company already possesses a market price the valuation problem will already be solved. Instead, business valuation must be able to identify under- or overvalued companies thereby taking needs of the particular buyer/seller into account.

Third, we take peculiarities of valuation into account, like buyers'/sellers' position, lifecycle of the firm (including negative earnings), synergies, value of corporate control etc. and show for buyer's/seller's position empirically valuation differences. Principally discounted cash flow models are able to deal with these peculiarities. However, textbook formulas (e.g., Brealey/Myers/Allen (2016), Damodaran (2006), Damodaran (2012), Berk/DeMarzo (2017), Ross/Westerfield/ Jaffe/Roberts (2015), and Ross/Westerfield/Jordan (2015)) ignore at least buyers'/sellers' position. The multiples and the value relevance literature cannot address buyers'/sellers' position either. In addition, the value relevance literature barely analyzes synergies with Henning/Lewis/Shaw (2000) as sole exception. Not surprisingly, the different importance of synergies for buyers and sellers is not addressed at all.

The remainder of this Chapter V is organized as follows: Section 2 develops the optimize-the-price approach theoretically. Section 3 analyzes the optimize-the-price approach empirically. Section 4 concludes this chapter.

2 Optimize-the-Price Approach

Chapter IV, Sections 2.3.1 and 2.3.2 develop an economic model evaluation criterion to evaluate the economic content of empirical asset pricing approaches. In Chapter IV this economic model evaluation criterion is used to detect the problems of empirical asset pricing models.

In this section this economic model evaluation criterion is applied in a constructive way, namely to design an accounting-based business valuation model.

2.1 Requirements for an Economically Convincing Business Valuation Model

The economic model evaluation criterion comprises two components: economic principle and institutional circumstances.

2.1.1 Economic Principle (see Chapter IV, Section 2.3.1)

It follows from the economic principle (see Chapter IV, Section 2.3.1):

(i) Objective function and constraints must be considered simultaneously and not separately because they both together constitute decision makers' objectives.

(ii) An actual company must be modelled (that is either be purchased or sold), i.e., arbitration and not free lunch must be pursued.

(iii) Buyers' and sellers' point of views must be reflected.

Finally, to implement the economic principle formally, it is recommended to define all accounting characteristics in a way so that a higher value of the accounting characteristic is unequivocally associated with a more desirable outcome. Higher earnings are clearly better than lower earnings. However, a debt-to-equity ratio defined as debt divided by equity shows the inverse relation: A lower debt-to-equity-ratio is preferable. For that reason, the ratio is re-defined as equity divided by debt or equity divided by total assets since higher equity means lower insolvency risk and, thus, is regarded as better.

Taking together bullet points (i) to (iii) and the positive definition of accounting characteristics, the economic principle translates into the following two decision problems:

(5.1) Decision problem buyer:
A buyer minimizes the (purchase) price of the company subject to accounting characteristics greater than or equal to a lower bound (accounting characteristics represent what the buyer gets in return for the investment).

(5.2) Decision problem seller:
A seller maximizes the (sales) price of the company subject to accounting characteristics less than or equal to an upper bound (accounting characteristics represent what the seller loses in return for the sales price).

2.1.2 Institutional Circumstances (see Chapter IV, Section 2.3.2)

Institutional circumstances refer to the constraints on portfolio holdings:

(i) Legal environment

Prohibition of uncovered short sales (see Regulation (EU) No 236/2012, Article 12) indicates that the sales of stock i must be less than or equal to a lower bound. There are no limits on purchases of stocks assuming that companies are not subject to capital adequacy regulations.

(ii) Market usages

Market usages might impose an upper limit on covered short sales because they require a certain amount of collateral for securities lending. Such a limit is again imposed on each asset i. Moreover, there are market impact costs, i.e., large amounts of purchases and/or (short) sales influence market prices; a fact that might give rise to an upper limit on portfolio holdings.

In summary, institutional circumstances, hence, impose constraints on portfolio holdings in the decision problem buyer (5.1) and the decision problem seller (5.2).

2.2 One-Period Model

To learn about the formalization of the decision problems (5.1) and (5.2), we start with an (unrealistic) one-period model and extend it in Subsections 2.3 to 2.5.

2.2.1 Model

Based on the verbal description of the decision problem buyer (5.1) and the decision problem seller (5.2), these two decision problems can be formulated as follows:

– Buyer

(5.3)

$$\min_{N_{1,t}^{+},N_{1,t}^{-},\dots,N_{n,t}^{+},N_{n,t}^{-}} \sum_{i=1}^{n} \left(N_{i,t}^{+} - N_{i,t}^{-}\right) \cdot P_{i,t}$$

s.t.

(5.4)

$$\begin{aligned} N_{1,t}^{+}A_{1,1,t} - N_{1,t}^{-}A_{1,1,t} + \cdots + N_{n,t}^{+}A_{n,1,t} - N_{n,t}^{-}A_{n,1,t} &\geq a_1 \\ &\vdots \\ N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} &\geq a_m \end{aligned}$$

(5.5)

$$\begin{aligned} f\left(N_{1,t}^{+}, \dots, N_{i,t}^{+}, \dots, N_{n,t}^{+}, N_{1,t}^{-}, \dots, N_{i,t}^{-}, \dots, N_{n,t}^{-}\right) &\leq g_1(x) \\ &\vdots \\ f\left(N_{1,t}^{+}, \dots, N_{i,t}^{+}, \dots, N_{n,t}^{+}, N_{1,t}^{-}, \dots, N_{i,t}^{-}, \dots, N_{n,t}^{-}\right) &\leq g_n(x) \end{aligned}$$

$$\begin{gathered} N_{1,t}^{+} \geq 0 \\ N_{1,t}^{-} \geq 0 \\ \vdots \\ N_{n,t}^{+} \geq 0 \\ N_{n,t}^{-} \geq 0 \end{gathered}$$

where $N_{i,t}^{+}$ ($N_{i,t}^{-}$) denotes the numbers of asset i purchased (sold) at time t, $P_{i,t}$ the price of company i at time t, $A_{i,j,t}$ accounting characteristics j of company i at time t, and a_j accounting characteristics j of the company to be valued. $g(x)$ and $f(.)$ are functions that determine the portfolio holdings constraints.

– Seller

(5.6)

$$\max_{N_{1,t}^{+},N_{1,t}^{-},\dots,N_{n,t}^{+},N_{n,t}^{-}} \sum_{i=1}^{n} \left(N_{i,t}^{+} - N_{i,t}^{-}\right) \cdot P_{i,t}$$

s.t.

(5.7)

$$N_{1,t}^{+}A_{1,1,t} - N_{1,t}^{-}A_{1,1,t} + \cdots + N_{n,t}^{+}A_{n,1,t} - N_{n,t}^{-}A_{n,1,t} \leq a_1$$
$$\vdots$$
$$N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} \leq a_m$$

(5.8)

$$f\left(N_{1,t}^{+}, \dots, N_{i,t}^{+}, \dots, N_{n,t}^{+}, N_{1,t}^{-}, \dots, N_{i,t}^{-}, \dots, N_{n,t}^{-}\right) \leq g_1(x)$$
$$\vdots$$
$$f\left(N_{1,t}^{+}, \dots, N_{i,t}^{+}, \dots, N_{n,t}^{+}, N_{1,t}^{-}, \dots, N_{i,t}^{-}, \dots, N_{n,t}^{-}\right) \leq g_n(x)$$

$$N_{1,t}^{+} \geq 0$$
$$N_{1,t}^{-} \geq 0$$
$$\vdots$$
$$N_{n,t}^{+} \geq 0$$
$$N_{n,t}^{-} \geq 0$$

2.2.2 Economic Analysis of the Decision Problem

The objective function in combination with the constraints on accounting characteristics (buyer: (5.3) in combination with (5.4); seller: (5.6) in combination with (5.7)) implement bullet point (i) of the economic principle. That the constraints on accounting characteristics refer to the accounting characteristics of the company to be valued, i.e., accounting characteristic$_j$ (5.4) $\geq a_j$ or (5.7) $\leq a_j$ implements bullet point (ii) of the economic principle. A free lunch would have been on (5.4) ≥ 0 or (5.7) ≤ 0. The differentiation between buyer and seller (buyer: (5.3) in combination with (5.4); seller: (5.6) in combination with (5.7)) addresses bullet point (iii) of the economic principle. Finally, the constraints on portfolio holdings (5.5) or rather (5.8) take legal environment and market usages into account.

Three remarks are in order to finalize the analysis of the buyer's and seller's decision problems. First, the constraints on portfolio holdings are not specified to a particular model at this stage of the analysis. Second, buyers' and sellers' decision problems do not contain a budget constraint. This is due to the fact that pricing is done and the actual purchase/sale, where funding comes into play, is subsequent to the pricing problem. Third, (5.3) and (5.6) are empirical models because they try to explain today's prices using accounting characteristics. They do not compute today's prices as present value of future cash flows as theoretical asset pricing models would do.

To analyze both decision problems further, we compare prices for buyers and sellers and explore the effect of the constraints on portfolio holdings.

Intuitively, one would argue that prices of sellers can never be below prices of buyers, at worst, both prices coincide.—This intuition is, however, wrong. In a model with just one accounting characteristic the accounting constraint can always be met exactly, i.e., duplication holds. In models with several accounting characteristics, it could happen that some accounting characteristics constraints can only be met as inequalities, i.e., higher (lower) values than the desired a_j are obtained for the buyer (seller), because only then the other accounting characteristics constraints will be met. In other words super-replication (supra-replication) holds. The price of a super-replication (supra-replication) portfolio is higher than/equal to (lower than/equal to) the price of a duplication portfolio. Thus, under admittedly rare circumstances, the super-replicating price of a buyer can be higher than the supra-replicating price of a seller in models with several accounting characteristics, in particular if there are by far more companies than constraints (see Section 3.3.2.1.3, Figure 5.6 for an example).

The constraints on portfolio holdings (5.5) or rather (5.8) could be formulated as constraints for the portfolio holdings on individual assets or on a (weighted) sum of assets. An example for a constraint on an individual asset is a short sale constraint on asset$_i$ or an upper limit on the portfolio holding of asset$_j$. An example for a constraint on portfolio holdings of a (weighted) sum of assets is the constraint obtained in regression approaches.

$$\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}\leq x$$

The effect, i.e., the tightness of the portfolio holdings constraints (5.5) or rather (5.8) is driven by two aspects. On the one hand, the shape of the constraint captured by the specific form of the of function $f(.)$: It is known from Chapter IV, Section 3.1.2 that an increase in $q\equiv\frac{p}{p-1}$ makes the constraints on portfolio holdings (5.5) or rather (5.8) less binding. A less binding constraint, however, results in lower (higher) prices for buyers (sellers) because buyers can sell more expensive companies short and, that way, bring prices further down and sellers can hold an higher number of expensive companies bringing prices up. On the other hand, the right-hand side $g_i(x)$—the size—of portfolio holdings constraints (5.5) or rather (5.8) exerts influence. A higher right-hand side $g_i(x)$ means a less binding constraint (enlarged short sales for buyers, enlarged purchases for sellers) and, thus, lower (higher) prices for buyers (sellers).

2.3 Extension to the Optimize-the-Price Approach: Synergies

2.3.1 Modelling Synergies

When purchasing or selling an actual company and not trying to explain, e.g., drivers of stock returns in general, synergies should not be overlooked as they arise when two firms are combined (see, e.g., Damodoran (2006), p. 1013). It can be distinguished (see, e.g., Damodaran (2006), pp. 1014 f.) between operating synergies that allow firms to increase operating income from existing assets and/or growth, and financial synergies that arise from higher cash flows and/or lower cost of capital.

Intuitively, synergies mean that a buyer of a company gains more and a seller loses more than the accounting characteristics of the company. More precisely, modelling synergies comprises: (i) Changes in (all of the) m accounting characteristics of the company to be valued must be captured. That way, both operating and financial synergies can be integrated, e.g., an increase in earnings and the equity-to-debt ratio as well as positive and negative synergies, e.g., an increase in earnings and a decrease in the equity-to-debt ratio. (ii) Synergies might not arise only in connection with one company, but with several companies if the acquiring and/or acquired company is a conglomerate.

Formally, the base case of bullet point (i) means: Company U possesses accounting characteristic $a_{U,j}$ and acquires company C_i thereby obtaining accounting characteristic $a_{C_i,j}$. Positive synergies signify $s\left(a_{U,j}, a_{C_i,j}\right) > a_{C_i,j}$, negative synergies signify $s\left(a_{U,j}, a_{C_i,j}\right) < a_{C_i,j}$. In that connection it does not matter whether company U or C_i or both are responsible for the synergies. All that matters is that for valuation purposes company C_i has accounting characteristics $s\left(a_{U,j}, a_{C_i,j}\right) \neq a_{C_i,j}$. The sale of company C_i results in a situation where the seller of company C_i loses $s\left(a_{U,j}, a_{C_i,j}\right) > a_{C_i,j}$ instead of just $a_{C_i,j}$ with positive synergies and $s\left(a_{U,j}, a_{C_i,j}\right) < a_{C_i,j}$ with negative synergies.

If the acquiring (selling) company U is a conglomerate (bullet point (ii)), its accounting characteristic j is equal to the weighted sum of accounting characteristics j of the n_U components of the conglomerate, i.e., $a_{U,j} = \sum_{k=1}^{n_U} a_{U_k,j}$. The company to be acquired/sold possesses as accounting characteristic j $a_{C_i,j} = \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}$ if it is a conglomerate. Therefore, synergies can be formalized as

(5.9)

$$s\left(\sum_{k=1}^{n_U} a_{U_k,j}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}\right)$$

Looking at the synergy formalization (5.9) reveals, however, that accounting characteristic a_j is considered in isolation of a_k, i.e., there is no effect like $s\left(a_{U,j}, a_{C_i,j}\right) + s\left(a_{U,k}, a_{C_i,k}\right) < s\left(a_{U,j}, a_{C_i,j}, a_{U,k}, a_{C_i,k}\right)$. We have decided against modelling such an effect for two reasons. First, such an effect would consider accounting characteristics as source and not just as result of synergies. E.g., economies of scale (= source) can influence both sales and earnings (= results) positively. However, $s\left(a_{U,j}, a_{C_i,j}, a_{U,k}, a_{C_i,k}\right)$ would imply that economies of scale and sales together cause higher earnings. While such a causal relation between accounting characteristics is not impossible[7], it will probably be rare and, hence, not be a good starting point for the valuation effect of synergies. Second, using $s\left(a_{U,j}, a_{C_i,j}, a_{U,k}, a_{C_i,k}\right)$ creates an allocation problem: Which accounting characteristic gets the benefit of the synergy, accounting characteristic j or k?—Recall an allocation is required because each accounting characteristic needs its own constraint (5.4) (buyer) or (5.7) (seller).

2.3.2 Valuation Model

Integrating the formalization of synergies (5.9) into the buyer's/seller's decision problem delivers—recall core of interest is the valuation of a company with synergies, but not the valuation of the synergy:

- Buyer

  (5.3)

$$\min_{N_{1,t}^+, N_{1,t}^-, \ldots, N_{n,t}^+, N_{n,t}^-} \sum_{i=1}^{n} \left(N_{i,t}^+ - N_{i,t}^-\right) \cdot P_{i,t}$$

  s.t.

  (5.10)

$$N_{1,t}^+ A_{1,1,t} - N_{1,t}^- A_{1,1,t} + \cdots + N_{n,t}^+ A_{n,1,t} - N_{n,t}^- A_{n,1,t} \geq s\left(\sum_{k=1}^{n_U} a_{U_k,1}, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},1}\right)$$
$$\vdots$$

7 E.g., a company could be close to insolvency due to lack of cash flows why customers do no longer want to purchase products from this company. A stock-based acquisition of a cash-rich company will result in a better liquidity situation. Since insolvency risk is now reduced, customers will begin buying company's products. In other words, a better liquidity situation caused the increase in sales.

$$N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} \geq s\left(\sum_{k=1}^{n_U} a_{U_k,m}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},m}\right)$$

(5.5)

Constraints on portfolio holdings

– Seller

(5.6)

$$\max_{N_{1,t}^{+},N_{1,t}^{-},\ldots,N_{n,t}^{+},N_{n,t}^{-}} \sum_{i=1}^{n} \left(N_{i,t}^{+} - N_{i,t}^{-}\right) \cdot P_{i,t}$$

s.t.

(5.11)

$$N_{1,t}^{+}A_{1,1,t} - N_{1,t}^{-}A_{1,1,t} + \cdots + N_{n,t}^{+}A_{n,1,t} - N_{n,t}^{-}A_{n,1,t} \leq s\left(\sum_{k=1}^{n_U} a_{U_k,1}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},1}\right)$$

$$\vdots$$

$$N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} \leq s\left(\sum_{k=1}^{n_U} a_{U_k,m}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},m}\right)$$

(5.8)

Constraints on portfolio holdings

2.3.3 Analysis of the Effects of Synergies

The exact effect of synergies on company prices depends on the specific form of $s\left(\sum_{k=1}^{n_U} a_{U_k,j}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}\right)$ and, hence, can only be evaluated numerically. Nevertheless, a theoretical analysis is able to gain several insights.

(i) Positive versus negative versus mixed synergies

It holds with positive synergies: $s\left(\sum_{k=1}^{n_U} a_{U_k,j}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}\right) > \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}$; with negative synergies: $s\left(\sum_{k=1}^{n_U} a_{U_k,j}\,, \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}\right) < \sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}$. Consequently, positive (negative) synergies call for higher (lower) portfolio holding in the (super-/supra-) replication portfolio that result in higher (lower) prices.

Mixed synergies mean that some accounting characteristics increase, e.g., earnings, and other decrease, e.g., equity-to-debt ratio. They have a mixed effect on portfolio holdings of the (super-/supra-) replication portfolio which translates into an ambiguous effect on prices.

(ii) Buyer versus seller

In the case of positive synergies buyers gain more than $\sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}$, sellers lose more than $\sum_{k=1}^{n_{C_i}} a_{U_{C_i},j}$. Therefore, the prices for both buyers and sellers increase.—This result fits nicely to the economic intuition: A buyer receives in the case of positive synergies more than accounting characteristic j and, hence, should be willing to pay more. A seller loses more than accounting characteristic j. For that reason, compensation in the form of a higher sales price is demanded.—These results transfer to negative and mixed synergies: Negative synergies call for a lower price for both buyers and sellers, whereas the effect of mixed synergies is ambiguous.

(iii) Which companies are used for (super-/supra-) replication in the presence of synergies?

Creating a (super-/supra-) replication portfolio means that several companies are combined hypothetically, whereas synergies arise from the combination of actual companies. Therefore, only companies outside of the conglomerate to be valued should be used for replication purposes. Since these companies are economically and often legally independent, no synergies[8] will arise when their accounting characteristics are (hypothetically) combined in a (super-/supra-) replication portfolio.

[8] Things might be different with financial synergies since a portfolio always offers diversification benefits. For this reason, we discuss aspects of risk in a subsection of its own (see Subsection 2.5).

(iv) Value of corporate control

The value of corporate control rests upon the idea that a controlling owner would operate the firm differently from the way it is operated currently (see, e.g., Damodaran (2006), p. 845) and, hence, allows to achieve more desirable accounting characteristics compared to a situation when the decision maker is not the controlling owner.

Technically, getting more desirable accounting characteristics means that the right-hand side of the accounting characteristics constraints (5.4) (buyer) or (5.7) (seller) increases. In other words, the value of corporate control can be captured (technically) by means of the accounting characteristics constraints with synergies (5.10) (buyer) and (5.11) (seller).

2.4 Extension to the Optimize-the-Price Approach: Multi-Period Features

2.4.1 Modelling Multi-Period Features

Market prices comprise future cash flow streams in one value. However, accounting characteristics do not possess a similar reference to the future because, e.g., today's earnings cannot be interpreted as the present value of future earnings. Put more precisely, Damodaran (2012), pp. 611, 633, 644-645 gives some economic justifications for the use of multi-period models (or rather the inadequateness of a steady state assumption):

(i) Companies in general are subject to a life cycle with different values for revenues and earnings in each phase. In particular, negative or abnormally low earnings create valuation problems. This fact becomes particularly visible with start-up firms that often lose money but at the same time are characterized by high values.

(ii) If companies have a significant likelihood of distress or default, a going concern assumption cannot be applied. Instead, accounting characteristics must be adjusted, e.g., low earnings and equity-to-debt-ratios are used during such crisis times.

(iii) Translating corporate control into a different corporate policy might take time. Therefore, accounting characteristics might not change immediately but until some time has passed so that the new corporate policy can become effective.

Principally, there are two possibilities to model multi-period aspects.

(i) Accounting characteristics are considered at different points in time. This means, instead of arguing with a_j as in the one-period model, $a_{j,t}, a_{j,t+1}, \dots$ are applied. Such a complete specification of accounting characteristics at all points in time is able to capture the life cycle of a company and, as such, can also be applied to start-ups. However, a complete specification of accounting characteristics at all points in time might be impossible given the (ex ante) infinite life of a company. Therefore, a simplification in the form $a_{j,t}, a_{j,t+1}, \dots, a_{j,T}$ might be needed, where $a_{j,T}$ could be regarded as a steady state accounting characteristic that is associated with a horizon value. In other words, only years t to $T-1$ are modelled in detail and starting from year T standardization is applied.—Note in this connecting that such a model implies identically constraints on accounting characteristics from time T on.

(ii) Growth rates for accounting characteristics are specified. The idea behind this approach can be found in Barth/Beaver/Landsman (2001), p. 95 who conclude that "price studies are interested in determining what is reflected in firm value while return studies (price changes) are interested in determining what is reflected in change in value over a specific period of time". In this connection, different growth rates can be specified for each period to reflect the life cycle of companies or just one growth rate if companies have reached the steady state. Nevertheless, growth rates have two disadvantages: First, the temporary nature of negative or abnormally low earnings cannot be captured by earnings growth rates (see Damodaran (2012),

p. 611). Second, the theoretical analyses in Chapter III, Section 2.2.3 showed that pure returns/growth rates are incompatible with the economic principle.

2.4.2 Valuation Model

Integrating the formalization of multi-period models into the buyers'/sellers' decision problem delivers (only the version is depicted that specifies $a_{j,t}, a_{j,t+1}, \dots, a_{j,T}$):

– Buyer

(5.3)

$$\min_{N_{1,t}^{+}, N_{1,t}^{-}, \dots, N_{n,t}^{+}, N_{n,t}^{-}} \sum_{i=1}^{n} \left(N_{i,t}^{+} - N_{i,t}^{-}\right) \cdot P_{i,t}$$

s.t.

(5.12)

$$\begin{gathered} N_{1,t}^{+} A_{1,1,t} - N_{1,t}^{-} A_{1,1,t} + \cdots + N_{n,t}^{+} A_{n,1,t} - N_{n,t}^{-} A_{n,1,t} \geq a_{1,t} \\ \vdots \\ N_{1,t}^{+} A_{1,m,t} - N_{1,t}^{-} A_{1,m,t} + \cdots + N_{n,t}^{+} A_{n,m,t} - N_{n,t}^{-} A_{n,m,t} \geq a_{m,t} \\ \vdots \\ N_{1,t}^{+} A_{1,1,T} - N_{1,t}^{-} A_{1,1,T} + \cdots + N_{n,t}^{+} A_{n,1,T} - N_{n,t}^{-} A_{n,1,T} \geq a_{1,T} \\ \vdots \\ N_{1,t}^{+} A_{1,m,T} - N_{1,t}^{-} A_{1,m,T} + \cdots + N_{n,t}^{+} A_{n,m,T} - N_{n,t}^{-} A_{n,m,T} \geq a_{m,T} \end{gathered}$$

(5.5)

Constraints on portfolio holdings

– Seller

(5.6)

$$\max_{N_{1,t}^{+}, N_{1,t}^{-}, \dots, N_{n,t}^{+}, N_{n,t}^{-}} \sum_{i=1}^{n} \left(N_{i,t}^{+} - N_{i,t}^{-}\right) \cdot P_{i,t}$$

s.t.

(5.13)

$$N_{1,t}^{+}A_{1,1,t} - N_{1,t}^{-}A_{1,1,t} + \cdots + N_{n,t}^{+}A_{n,1,t} - N_{n,t}^{-}A_{n,1,t} \leq a_{1,t}$$
$$\vdots$$
$$N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} \leq a_{m,t}$$
$$\vdots$$
$$N_{1,t}^{+}A_{1,1,T} - N_{1,t}^{-}A_{1,1,T} + \cdots + N_{n,t}^{+}A_{n,1,T} - N_{n,t}^{-}A_{n,1,T} \leq a_{1,T}$$
$$\vdots$$
$$N_{1,t}^{+}A_{1,m,T} - N_{1,t}^{-}A_{1,m,T} + \cdots + N_{n,t}^{+}A_{n,m,T} - N_{n,t}^{-}A_{n,m,T} \leq a_{m,T}$$

(5.8)

Constraints on portfolio holdings

2.4.3 Analysis of the Effects of Multi-Period Features

Multi-period models cause two differences compared to the one-period model of Section 2.2.

First, they contain more accounting characteristics constraints (5.12) and (5.13). More constraints, however, reduce the scope for portfolio optimization meaning that buyers' prices will be higher and sellers' prices will be lower in the multi-period compared to the one-period model.

Second, the right-hand sides of the accounting characteristics constraints exert influence on the pricing problem as well. Assume accounting characteristic$_j$ of the company to be valued increases faster (slower) over time as for the companies of the (super-/supra-) replication portfolio, e.g., due to a different phase of the company life cycle. Then, the portfolio holdings in the (super-/supra-) replication portfolio increase (decrease) meaning higher (lower) prices for buyers and sellers.

Effects (i) and (ii) have contradictory effects on company prices and might give rise to a situation where the multi-period model delivers results similar to the one-period model of Section 2.2. In other words, the steady state assumption of the one-period model might not be completely unrealistic.

2.5 Extension to the Optimize-the-Price Approach: Risk/Uncertainty

2.5.1 Modelling Risk/Uncertainty

The optimize-the-price approach developed so far just works under certainty. Multiples and value relevance studies also work under certainty. However, residual income models express expected cash flows with the help of accounting characteristics and discount them with constant cost of capital (Easton (2004), Gebhardt/Lee/Swaminathan (2001), and Ohlson/Juettner-Nauroth (2005)) or a constant risk premium (Claus/Thomas (2001)). From that perspective they can handle risk, albeit in a stylized way. Since it is the objective of the optimize-the-price approach to improve the theoretical foundation of accounting-based asset pricing, it needs an extension towards risk/uncertainty to become an economically dominant approach.

Principally, there are two possibilities of integrating risk/uncertainty into the optimize-the-price approach.

First, the stochasticity of future accounting characteristics is addressed by means of different states. To that end, a forecast for accounting characteristic$_j$ at each point in time τ and in each state s is developed, i.e., $a_{j,\tau,s}$. Since time- and state-dependent accounting characteristics are modelled, no state probabilities are needed, (Knightian) uncertainty is given. States can be derived from scenarios that themselves can be independent in each point of time (e.g., at time $t+1$ there is a good, normal, and bad state and at time $t+2$ there are other good, normal, and bad states; the states at time $t+1$ are not intertemporally connected with the states at time $t+2$) or can be intertemporally connected (e.g., from the good state at time $t+1$ there are three possible consecutive states at time $t + 2$ so that a sequence is obtained: good-good, good-normal, and good-bad).

Buyers construct the (super-) replication portfolio to obtain at least the state-dependent accounting characteristics of the business to be valued at a minimum price; sellers setup the (supra-) replication portfolio to lose not more than the state-dependent accounting characteristics of the business to be valued and to achieve a maximum price. Note with respect to the optimization potential: If there are more companies than states, markets will be complete. In such a situation optimization will be nontrivial because accounting characteristic$_j$ can be achieved by different combinations of companies. Only if the number of companies is equal to the number of states, a unique solution (under some regulatory conditions) will be obtained thus eliminating any optimization potential.

Second, expected values and risk measures (e.g., Lower Partial Moments, Value at Risk etc.) based on the distribution of accounting characteristics are employed. Since distributions are considered, state probabilities must be known and, hence, a situation under risk is obtained.

Buyers construct the (super-) replication portfolio to obtain the expected value of the accounting characteristics of the business to be valued at a minimum price, sellers setup the (supra-) replication portfolio to lose not more than the expected value of the accounting characteristics of the business to be valued and to achieve a maximum price. Since risk measures take the fluctuation aspect into account, they are regarded as negative. Thus, both buyers and sellers would like to keep fluctuations in check. This can be achieved by adding a constraint on the risk measure of each accounting characteristic in the optimization problem.

To implement the risk measure approach in detail, two different paths may be chosen. First, expected values and risk measures are computed for each point in time. On the one hand, this can be achieved with the help of scenarios. E.g., if there are three equally probable realizations of accounting characteristics between time τ and $\tau + 1$, the (conditional) expected values simply is the weighted average of the three realizations.—However: If scenarios have already been derived, it would not be useful to create an information loss by condensing them into expected values and risk measures. On the other hand, expected values and risk measures are estimated from past realizations using time series estimation. If the time series of accounting characteristics are (weakly) stationary, time series estimators can be applied to obtain estimators of expected values and risk measures. Otherwise, the time series must be made (weakly) stationary, e.g., by means of using growth rates instead of values of accounting characteristics.—Caveat: Growth rates have problems in adequately dealing with negative accounting characteristics. Second, the development of expected values and risk measures of accounting characteristics over time is modelled. Expected values and/or risk measures of accounting characteristics might exhibit constant growth or a triangular form over time similar to the expected value forecasts in dividend discount models.

2.5.2 Valuation Model

Integrating the formalization of risk/uncertainty into the buyers'/sellers' decision problem delivers—objective functions and constraints on portfolio holdings remain the same as in Sections 2.3.2 or 2.4.2—[9]:

– Accounting characteristics constraints using time- and state-dependent forecasts for accounting characteristics

  (5.14)

$$N_{1,\tau}^{+} A_{1,j,\tau,s} - N_{1,\tau}^{-} A_{1,j,\tau,s} + \cdots + N_{n,\tau}^{+} A_{n,\tau,s} - N_{n,\tau}^{-} A_{n,j,\tau,s} \;\begin{array}{c} \geq \\ \scriptstyle buyer \\ \leq \\ \scriptstyle seller \end{array}\; a_{j,\tau,s}$$

  for point in time $\tau \in \{t, \dots, T\}$ and state $s \in \{1, \dots, S_\tau\}$.

[9] Principally constraints on portfolio holdings could be time- and state-dependent due to, e.g., time- and state-variable market impact costs.—However, such an approach seems to be artificially complicated and will hide the structure of the decision problem behind technicalities. Therefore, we will not consider this case.

– Accounting characteristics constraints using expected values and risk measures for accounting characteristics

Expected value

(5.15)

$$N_{1,\tau}^{+}E_t\{A_{1,j,\tau}\} - N_{1,\tau}^{-}E_t\{A_{1,j,\tau}\} + \cdots + N_{n,\tau}^{+}E_t\{A_{n,\mathrm{j},\tau}\} - N_{n,\tau}^{-}E_t\{A_{n,j,\tau}\} \underset{\substack{\overset{!}{\leq}\\ seller}}{\overset{\substack{\overset{!}{\geq}\\ buyer}}{}} E_t\{a_{j,\tau}\}$$

Risk measure

(5.16)

$$RM_k\left(N_{1,\tau}^{+}A_{1,j,\tau} - N_{1,\tau}^{-}A_{1,j,\tau} + \cdots + N_{n,\tau}^{+}A_{n,\mathrm{j},\tau} - N_{n,\tau}^{-}A_{n,j,\tau}\right) \underset{\substack{\overset{!}{\leq}\\ seller}}{\overset{\substack{\overset{!}{\leq}\\ buyer}}{}} RM_k\left(a_{j,\tau}\right)$$

where RM_k denote risk measure$_k$.

The left-hand sides of the accounting characteristics constraints (5.16) comprise diversification effects, i.e., financial synergies. This means the accounting characteristics of the (super-/supra-) replication portfolio exhibit less fluctuations than the accounting characteristics of an isolated position, the company to be valued. If financial synergies are unwanted, risk measures of the isolated positions of the (super-/supra-) replication portfolio must be taken instead of the risk measure of the (super-/supra-) replication portfolio itself. Note, however, that financial synergies cannot be excluded from time- and state-dependent accounting characteristics: There it is impossible to separate the influence of an accounting characteristic of one company from the one of other companies because risk measures that indicate risk connections, e.g., covariances, do not exist and, thus, they cannot be suppressed.

2.5.3 Analysis of the Effects of Risk/Uncertainty

Integrating risk/uncertainty into the optimize-the-price approach introduces effects to the one-period model of Section 2.2 that are very similar to the multi-period model of Section 2.4.

First, more constraints signify less flexibility in determining the (super-/supra-) replication portfolios. Hence, buyers' prices will be higher and sellers' prices lower.

Second, the right-hand sides of the accounting characteristics constraints (5.14) and the constraints on expected values of accounting characteristics (5.15) exert influence as well. Assume (the expected value of) accounting characteristic$_j$ of the company to be valued increases faster (slower) over time as for the companies of the (super-/supra-) replication portfolio, e.g., due a different phase of the company life cycle. Then, the portfolio holdings in the (super-/supra-) replication portfolio increase (decrease) meaning higher (lower) prices for buyers and sellers. The analyses are slightly more difficult with the risk measure constraints (5.16) because of diversification effects. A higher risk measure of the company to be valued (right-hand side of (5.16)) does not necessarily translate into higher portfolio holdings and, thus, prices.—Recall the right-hand side of (5.16) constitutes just an upper bound.

Effects (i) and (ii) have contradictory effects on company prices resulting in an ambiguous total effect. This ambiguous total effect, however, is similar to the one observed with multi-period but deterministic models in Section 2.4.3.

2.6 Comparison of the Optimize-the-Price Approach with Regression Approaches

2.6.1 Idea Behind and Implementation of the Comparison

Regressions determine regression coefficients for accounting characteristics from a group of companies and apply these regression coefficients to the accounting characteristics of the company to be valued. In other words, a two-step pricing approach is applied. The optimize-the-price approach, on the other hand, is a one-step pricing approach because it directly determines the price of the company to be valued without requiring intermediate steps like the determination of regression coefficients.—Due to this completely different determination of prices both approaches seem to be difficult to compare.

However, duality theory might help with the comparison. The variables of the dual program of the optimize-the-price approach can be interpreted as regression coefficients (see Appendix 1.4 for the general derivation and Appendix 1.4.5 for the dual program). The regression coefficients from regressions can be compared with regression coefficients from the optimize-the-price approach[10].

– Buyer (see Appendix 1.4.5, adjusted to values of accounting characteristics and cross section)

(A1.43)

$$\min_{\mu,\beta_1,\ldots,\beta_m} \mu \cdot x - \sum_{j=1}^{m} \beta_j \cdot a_j$$

[10] This is possible because the optimize-the price problem meets the requirement of strong duality (see Appendix 1.5). Therefore, the optimal values of the objective function of the primal and the dual program coincide. Since we minimize the objective function in the (buyer's and seller's) dual program, the dual in standard form, however, is characterized by maximizing the objective function, the value of both objective functions differ by the factor−1.

s.t.

$$\left[(1-\tau)\cdot\sum_{i=1}^{n}|\varepsilon_i^+|^p+\tau\cdot\sum_{i=i}^{n}|\varepsilon_i^-|^p\right]^{\frac{1}{p}}\leq\mu$$

$$\mu\geq 0$$
$$\beta_1\geq 0,\dots,\beta_m\geq 0$$

where

$$\varepsilon_i^+=\sum_{j=1}^{m}A_{i,j}\beta_j-P_i\geq 0$$
$$\varepsilon_i^-=\sum_{j=1}^{m}A_{i,j}\beta_j-P_i<0$$

– Seller (see Appendix 1.4.7, adjusted to values of accounting characteristics and cross section)

(A1.55)

$$\min_{\mu,\beta_1,\dots,\beta_m}\mu\cdot x+\sum_{j=1}^{m}\beta_j\cdot a_j$$

s.t.

$$\left[(1-\tau)\cdot\sum_{i=1}^{n}|\varepsilon_i^+|^p+\tau\cdot\sum_{i=i}^{n}|\varepsilon_i^-|^p\right]^{\frac{1}{p}}\leq\mu$$

$$\mu\geq 0$$
$$\beta_1\geq 0,\dots,\beta_m\geq 0$$

where

$$\varepsilon_i^+=-\sum_{j=1}^{m}A_{i,j}\beta_j+P_i\geq 0$$
$$\varepsilon_i^-=-\sum_{j=1}^{m}A_{i,j}\beta_j+P_i<0$$

- Regression (see Appendix 1.2.1, adjusted to values of accounting characteristics and cross section)

(A1.4)

$$\min_{\substack{\mu_1^+,\mu_1^-,\ldots,\mu_n^+,\mu_n^-,\\ \beta_0,\beta_1,\ldots,\beta_m}} x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1}{p}}$$

s.t.

$$\begin{aligned}
\beta_0 + \sum_{j=1}^{m} A_{1,j}\beta_j - P_1 - \mu_1^+ &\leq 0\\
-\beta_0 - \sum_{j=1}^{m} A_{1,j}\beta_j + P_1 - \mu_1^- &\leq 0\\
&\vdots\\
\beta_0 + \sum_{j=1}^{m} A_{n,j}\beta_j - P_n - \mu_n^+ &\leq 0\\
-\beta_0 - \sum_{j=1}^{m} A_{n,j}\beta_j + P_n - \mu_n^- &\leq 0
\end{aligned}$$

$$\begin{aligned}
-\mu_1^+ &\leq 0\\
-\mu_1^- &\leq 0\\
&\vdots\\
-\mu_n^+ &\leq 0\\
-\mu_n^- &\leq 0
\end{aligned}$$

$$\beta_0 \in \mathbb{R}, \beta_1 \in \mathbb{R}, \ldots, \beta_m \in \mathbb{R}$$

2.6.2 Comparison

When the optimize-the-price approach and regressions are compared, four differences arise.

(i) Objective function

The objective function of the optimize-the-price approach and regressions differ by the term $\sum_{j=1}^{m} \beta_j \cdot a_j$, which is missing in regressions. This difference comes from the fact that the optimize-the-price approach focuses on arbitration and not on free lunches as regressions (see Chapter IV, Formula (4.2)). Intuitively, the difference is caused by the fact that regressions determine regression coefficients only from the group of companies and then apply them to the company to be valued. On the contrary, the optimize-the-price approach integrates the company to be valued into the determination of regression coefficients and, thus, can raise the mispricing potential not only of other companies, but also of the company to be valued. Therefore, applying more advanced regressions like generalized least squares regression cannot equalize the difference in the objective functions.

(ii) Buyer versus seller position

Regressions determine regression coefficients in a way so that the residual within the group of companies used for the regression ideally is equal to zero. Since the company to be valued cannot exert influence on regression coefficients, buyer and seller positions cannot be taken into account.

The-optimize-the-price approach by constructions optimizes the price, i.e., determines regression coefficients in a way so that buyer and seller positions are taken explicitly into account.

(iii) Regression coefficients

First, the optimize-the-price approach has no intercept β_0 as opposed to regressions. An intercept, however, is from an asset pricing perspective implausible since it means that there is a price component that is independent of company-specific accounting characteristics that serve as price drivers. Moreover, such an intercept implies (see Chapter IV, Formula (4.6)) that portfolio holdings must add to zero. Such a constraint can neither be justified from legal environment nor market usages.

Second, the optimize-the-price approach can justify a sign restriction on regression coefficients. Since the optimize-the-price approach rest upon explicitly formulated constraints on portfolio holdings (see (5.5) and (5.8)), a non-negativity constraint on portfolio holdings might be used. Then, Appendix 1.4.6.4, Formula (A1.51) shows that the non-negativity constrains on portfolio holdings result in non-negativity constraints on regression coefficients. That way, a theoretical/economic foundation of a procedure of Campbell/Thompson (2008) is delivered: They recommend using a sign restriction on regression coefficients. In particular if a regression coefficient has an unexpected sign, they set the regression coefficient equal to zero when forming forecasts.

(iv) Specification problem for the size $g(.)$ of the constraint on portfolio holdings

Regressions do not have a specification problem with respect to size $g(.)$ because they minimize pricing errors why x in Chapter IV, Formula (4.2) can be rightfully set equal to 1. On the contrary, the optimize-the-prize-approach requires the specification of the size $g(.)$ as input because otherwise it cannot find an optimal solution: Without $g(.)$ the market will not be free of an accounting arbitrage. If, however, upper limits for x must be specified to integrate institutional circumstances, the optimize-the-price approach can do this and, hence, is more flexible than regressions because in regressions just a given constraint on portfolio holdings can be used.

3 Empirical Analysis

3.1 Economic Significance of Price Differences Between Different Models

The analysis of the optimize-the-price approach in Section 2 has shown that shape $f(.)$ and size $g(.)$ of the constraints on portfolio holdings as well as buyers' and sellers' position exert influence on the optimize-the-price approach. Moreover, multi-period models are different from one-period models because the steady state assumption implied by one-period models does not hold in reality. Finally, the integrated (optimize-the-price approach) approach is conceptually different from the separated approach (regressions) and has no regression intercept; company-independent intercepts are from an asset pricing perspective implausible.—Although all these differences between models are clearly identifiable from a theoretical perspective, it is not clear whether they translate into economically significant price differences.

Therefore, this Section 3 focuses on the economic significance of model differences by answering the following questions empirically:

(i) The economic significance of price differences caused by shape $f(.)$ and size $g(.)$ of the constraints on portfolio holdings within the optimize-the-price approach.

(ii) The economic significance of price differences between one- and multi-period versions of the optimize-the-price approach.

(iii) The economic significance of price differences between integrated (optimize-the-price approach) and separated approaches (regressions) and its interplay with/without regression intercepts β_0.

(iv) The economic significance of differences between buyers' and sellers' prices.

3.2 Research Design and Data Set

The economic significance of Questions (i), (ii), (iii), and (iv) is analyzed with the help of "magnitude" (see Chapter III, Section 4.2) and "similarity" (see Chapter III, Section 4.3). In this connection, the answer to Question (iv) is then obtained as by-product to the answers of Questions (i), (ii), and (iii).

3.2.1 Research Design

- Question (i): constraints on portfolio holdings

  The shape $f(.)$ of the constraints on portfolio holdings is captured by analyzing four different constraints: short sale constraints as well as constraints in the form of L_1-, L_2-, and L_∞-norms. The size $g(.)$ of the constraint—relevant only for the L_1-, L_2-, and L_∞-norms—is addressed by specifying $g(.)$ as x and choosing three values: a small ($x = 0.5$), medium ($x = 1$), and relatively high value ($x = 2$).[11]

- Question (ii): one- versus multi-period models

  Multi-period models are constructed with 2010 as base year. The years 2011 to 2014 comprise the development of the accounting characteristics in the "future", i.e., specify the constraints at $t + 1$ to $t + 4$ in (5.12) and (5.13).—That way, empirical data can be used to represent accounting characteristics of the "future" instead of assuming arbitrary values. Both one- and multi-period models are computed with a non-negativity constraint on portfolio holdings only.

- Question (iii): comparison to regressions

  When comparing the optimize-the-price approach and regressions, it must be avoided that two effects are intermingled, namely the effect of different portfolio holding constraints and the integrated (optimize-the-price approach) versus separated (regressions) determination of regression coefficients. Therefore, the optimize-the-price approach is computed with that constraint on portfolio holdings that is implied by the specific regression model: constraints on individual portfolio holdings for quantile regressions (L_1-norm) and L_2-norm constraint for ordinary as well as weighted least squares regressions. Then, the pricing results of the (constraints-adjusted) optimize-the-price approach are compared to those of quantile (25%, 50%, and 75%), ordinary, and weighted least squares regressions with and without regression intercepts β_0.

[11] The empirical analysis in Section 3.3.2 demonstrates that these three values for x indeed create enough diversity.

3.2.2 Software

All computations are performed with RStudio Version 1.1.463 resting upon R version 3.6.0 (see R Core Team (2019)) using the following packages:

- lpSolve (version 5.6.13) for the optimize-the-price approach with non-negativity constraints on portfolio holdings
- nloptr (version 1.2.1) for the optimize-the-price approach with constraints on portfolio holdings in L_p-norm-form
- quantreg (version 5.38) for quantile regressions
- stats (version 3.6.0) for OLS and WLS regressions

3.2.3 Data Set

Principally the data set of Chapter III, Section 3.1 is used. However, the computation time for constraints on portfolio holdings in L_p-norm-form (for $x = 1$) takes 18,131.05 minutes, i.e., 12.6 days.

Thus, for L_p-norm constraints only a subset of this data set is considered: industrials from Europe in the year 2014. Since the robustness analyses regarding industry, region, and year in Chapter III, Sections 4.2.4 and 4.3.4 have shown that results regarding “magnitude” and “similarity” are not affected by industry, region, and year, such a limited data set seems to be acceptable to keep the computation time in check.

3.3 Empirical Results

3.3.1 Cleaning the Results of the Numerical Optimization

When analyzing the economic significance of the optimize-the-price approach, only solutions can be used that constitute (globally) optimal solutions and are not just local optima.

The structure of the optimization problem, in particular, the highly non-linear constraints on portfolio holdings in L_p-norm-form, however, makes numerical optimization a nontrivial task. The following two incidents give rise to local optima and, hence, call for their elimination before economic significance can be examined:

(i) The optimization algorithm exceeds the maximum number of iterations.

(ii) The alleged solution violates constraints of the optimization problem.

Appendix 5.3 contains details regarding the specific computations that had to be eliminated. Incident (i) is relevant in $\frac{47}{36{,}800} = 0.13\%$ of all cases, incident (ii) in $\frac{582}{36{,}800} = 1.58\%$. In other words, the data loss due to the inability to find a global optimum is not severe.

How can the histograms regarding “magnitude” and “similarity” be adjusted to cope with cases from incidents (i) and (ii)? To answer that question, fall back on the illustrative example of Chapter III, Section 4.1.

Assume that prices have been computed for the following companies U_j with the help of factors$_i$ and constraints on portfolio holdings in the form of L_p:

U_1 factor$_1$ L_1	U_1 factor$_2$ L_1	U_1 factor$_3$ L_1
U_1 factor$_1$ L_2	U_1 factor$_2$ L_2	U_1 factor$_3$ L_2
U_2 factor$_1$ L_1	U_2 factor$_2$ L_1	U_2 factor$_3$ L_1
U_2 factor$_1$ L_2	U_2 factor$_2$ L_2	U_2 factor$_3$ L_2

Now assume that U_1 factor$_1$ L_1 is not a global optimum because it belongs to incidents (i) or (ii). Then, it is not meaningful to compute Ratio (3.1) involving U_1

factor$_1$ L$_1$, i.e., $\frac{\mathrm{U_1\ factor_1\ L_2 - U_1\ factor_1\ L_1}}{\mathrm{market\ price\ U_1}}$ because using a non-global optimum might induce massive biases. The other ratios, however, can still be computed, i.e.,

$$\frac{\mathrm{U_2\ factor_1\ L_2 - U_2\ factor_1\ L_1}}{\mathrm{market\ price\ U_1}} \qquad \frac{\frac{\mathrm{U_1\ factor_2\ L_2 - U_1\ factor_2\ L_1}}{\mathrm{market\ price\ U_1}}}{\frac{\mathrm{U_2\ factor_2\ L_2 - U_2\ factor_2\ L_1}}{\mathrm{market\ price\ U_1}}} \qquad \frac{\frac{\mathrm{U_1\ factor_3\ L_2 - U_1\ factor_3\ L_1}}{\mathrm{market\ price\ U_1}}}{\frac{\mathrm{U_2\ factor_3\ L_2 - U_2\ factor_3\ L_1}}{\mathrm{market\ price\ U_1}}}$$

Therefore, histograms are computed based on the above five instead of originally six Ratios (3.1).

Finally, a third incident might endanger the judgement of economic significance:

(iii) The (empirical) solution of the optimize-the-price approach violates the theoretical order of Section 2.2.2, namely that
- an increase in $q \equiv \frac{p}{p-1}$ and/or
- a greater right hand side of the constraints on portfolio holdings

makes the constraints on portfolio holdings (5.5) or rather (5.8) less binding a fact that results in lower prices for buyers and higher prices for sellers.

The following two companies—Britisch American Tobacco and Bayerische Motorenwerke[12]—are examples of incident (iii). When minimizing the price of a portfolio (buyer's perspective) that offers at least the same amount of sales (factor M1) as British American Tobacco/Bayerische Motorenwerke, it is obtained:

Company	Constraint on portfolio holdings	Constraint on sales	Price (value of the objective function in the optimum)	Sales (value of the constraint in the optimum)
British American Tobacco	L$_1$, $x = 1$	16.96 (in billion EUR)	-89.15 (in billion EUR)	16.96 (in billion EUR)
	L$_2$, $x = 1$		-66.85 (in billion EUR)	31.18 (in billion EUR)
Bayerische Motorenwerke	L$_1$, $x = 1$	80.40 (in billion EUR)	5.12 (in billion EUR)	80.40 (in billion EUR)
	L$_2$, $x = 1$		5.83 (in billion EUR)	85.60 (in billion EUR)

Table 5.1: Examples where a L$_2$-norm constraint delivers higher prices for buyers than a L$_1$-norm constraint although theory suggests that prices for L$_2$-norm constraints cannot exceed those for L$_1$-norm constraints.

[12] Both companies have been chosen because they belong to the top five companies that produce incidents (i) or (ii).

Table 5.1 illustrates two seemingly puzzling results: Prices for British American Tobacco and Bayerische Motorenwerke for L_2 exceed those for L_1 even though theory states that a less binding constraint on portfolio holdings (L_2) cannot lead to higher prices that a more binding constraint (L_1). This puzzle can be resolved if the value of the accounting constraint in the optimum is taken into consideration. With a L_1-norm constraint on portfolio holdings sales of 16.96 billion EUR (British American Tobacco) or 80.40 billion EUR (Bayerische Motorenwerke) are priced. However, using the L_2-norm constraint on portfolio holdings results in sales of 31.18 billion EUR (British American Tobacco) or 85.60 billion EUR (Bayerische Motorenwerke). In other words, with a L_2-norm constraint on portfolio holdings super-replication occurs. Higher sales, however, justify a higher price meaning that incident (iii) should not be eliminated from assessing economic significance. Therefore, incident (iii) bears some similarity to the case from Section 2.2.2 where, under rare circumstances, prices of sellers were less than prices of buyers.

Note, however, two things in connection with incident (iii). First, super-replication does not necessarily lead to distorted price relations between L_1-norm and L_2-norm constraints on portfolio holdings:

Company	**Constraint on portfolio holdings**	**Constraint on sales**	**Price (value of the objective function in the optimum)**	**Sales (value of the constraint in the optimum)**
British American Tobacco	L_1, $x = 0.5$	16.96 (in billion EUR)	5.08 (in billion EUR)	17.24 (in billion EUR)
	L_2, $x = 0.5$		-26.52 (in billion EUR)	26.49 (in billion EUR)

Table 5.2: Examples where a L_2-norm constraint delivers lower prices for buyers than a L_1-norm constraint despite super-replication.

Second, sales have only a limited role in explaining prices. Theoretically, several accounting characteristics are needed to adequately reflect the multi-dimensionality of the figure cash flows that condense several value drivers into one figure. Empirically, actual market prices of British American Tobacco (79.22 billion EUR) and Bayerische Motorenwerke (58.9335 billion EUR) are far away from the prices that the optimize-the-price approach based on sales computes.

This means that market prices obviously are driven by other factors than sales alone.

3.3.2 Economic Significance of Shape and Size of Constraints on Portfolio Holdings

Analyzing the economic significance of shape and size of the constraints on portfolio holdings of the one-period model (5.5) and (5.8) requires: First, prices for several different constraints in L_p-norm-form (shape) and values of x (size) are computed for buyers and sellers. Second, these prices are compared to each other to analyze how different constraints in L_p-norm-form and x result in different "magnitudes" of price differences and, third, which L_p-norm-form and x combinations result in "similar" prices.

3.3.2.1 "Magnitude" of Price Differences of Different Shapes and Sizes of Constraints on Portfolio Holdings

"Magnitude" is computed as (normalized) difference between company values determined from all pricing approaches and the company value determined from the model under consideration. It is formalized by means of Ratio (3.1) (see Chapter III, Section 2.1.3.2) and will be regarded as high if Ratio (3.1) exceeds 10%.

3.3.2.1.1 The Role of Shape

To analyze the economic significance of shape, it is useful to break up the joint effect of position (buyer or seller) and shape of constraints in order to obtain a ceteris paribus analysis.

Consider $x = 1$ as an example (all other values for x can be found in Appendix 5.2.1.1.1). Then, it is obtained for buyers' and sellers' positions:

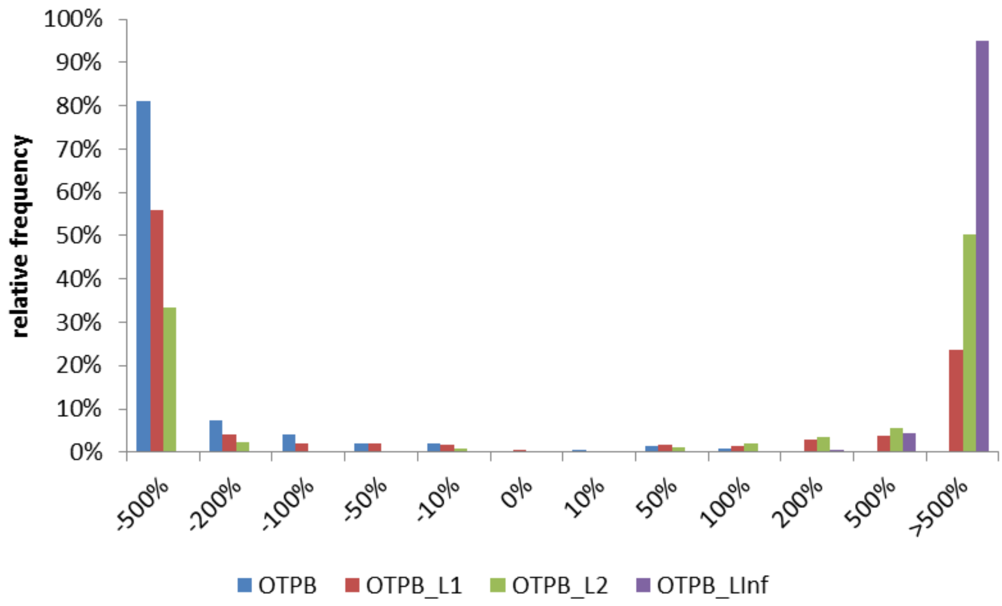


Figure 5.1: Histogram of Ratio (3.1) and x=1 for different shapes for buyers
Non-negativity constraint and valuation for buyer (OTPB))
L_1-norm constraint and valuation for buyer (OTPB_L1)
L_2-norm constraint and valuation for buyer (OTPB_L2)
L_∞-norm constraint and valuation for buyer (OTPB_LInf)

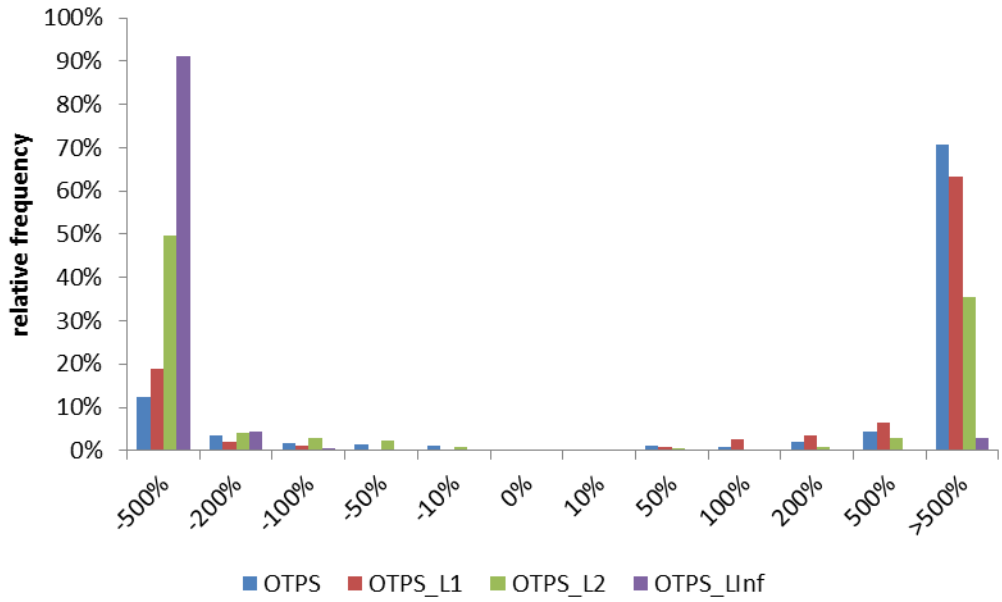


Figure 5.2: Histogram of Ratio (3.1) and x=1 for different shapes for sellers
Non-negativity constraint and valuation for the seller (OTPS)
L_1-norm constraint and valuation for the seller (OTPS_L1)
L_2-norm constraint and valuation for the seller (OTPS_L2)
L_∞-norm constraint and valuation for the seller (OTPS_LInf)

Figure 5.1 and Figure 5.2 demonstrate that the shape of constraints on portfolio holdings is economically highly significant. In addition, these differences in “magnitude” are not caused by different factors as the figures in Appendix 5.2.1.1.1 starting with Figure Appendix 5.2.1.1.1g show.

From an economic perspective, these shape effects portrayed in Figure 5.1 and Figure 5.2 can be explained as follows: Ratio (3.1) subtracts the price obtained from the reference model (the reference model is the one mentioned in the legend) from the prices of the other models. A L_1-norm constraint on portfolio holdings is more restrictive than a L_2-norm constraint which in turn is more restrictive than a L_∞-norm constraint. Therefore, prices of buyers with L_∞-norm constraints are the smallest, those with L_2-norm constraints are in the middle, and those with L_1-norm constraints are the highest. Yet, all three types of constraints allows for short sales. These are forbidden in the OTB version why this approach yields the highest (= worst) prices for buyers.

For the seller the order of L_1-, L_2-, and L_∞-norm constraints on portfolio holdings is the same. Therefore, sellers obtain best prices from L_∞-norm constraints followed by L_2- and L_1-norm constraints. Lowest prices for sellers are achieved with short sale constraints OTS: Sellers cannot sell low price companies short and invest proceeds in high price companies.

3.3.2.1.2 The Role of Size

A higher size x means less binding constraints on portfolio holdings and, thus, lower prices for buyers and higher prices for sellers. What Figure 5.3 and Figure 5.4, however, illustrate for a L_1-norm constraint on portfolio holdings—the figures for the other constraints on portfolio holdings can be found in Appendix 5.2.1.1.2—is the extreme effect of size x on Ratios (3.1), i.e., the high economic significance of size x on “magnitude”:

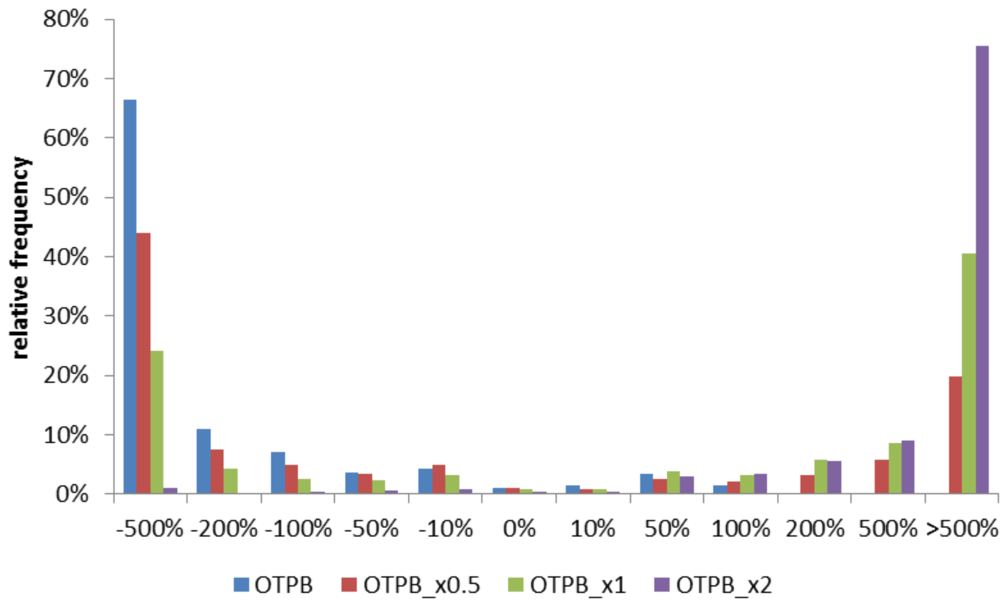


Figure 5.3: **Histogram of Ratio (3.1) and L_1 for different sizes x for buyers**

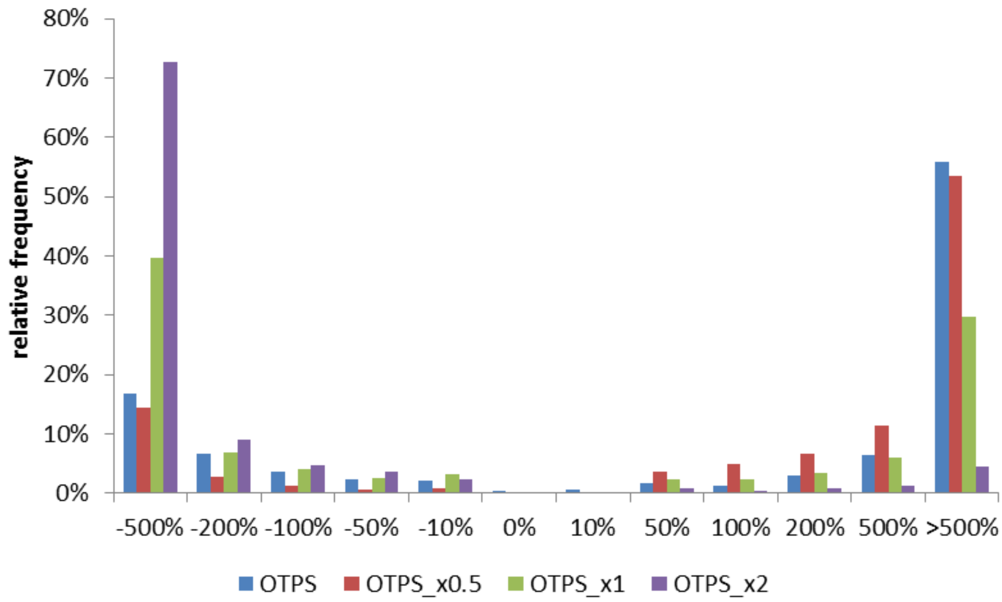


Figure 5.4: **Histogram of Ratio (3.1) and L_1 for different sizes x for sellers**

The high economic significance of size x is not limited to certain factors as the figures in Appendix 5.2.1.1.2 starting with Figure Appendix 5.2.1.1.2g show.

3.3.2.1.3 Prices of Buyers Versus Sellers

The economic significance of buyers' and sellers' positions becomes even more apparent if prices for buyers and sellers are compared directly. To that end, consider the ensuing exemplary figures, all figures can be found in Appendix 5.2.1.1.3:

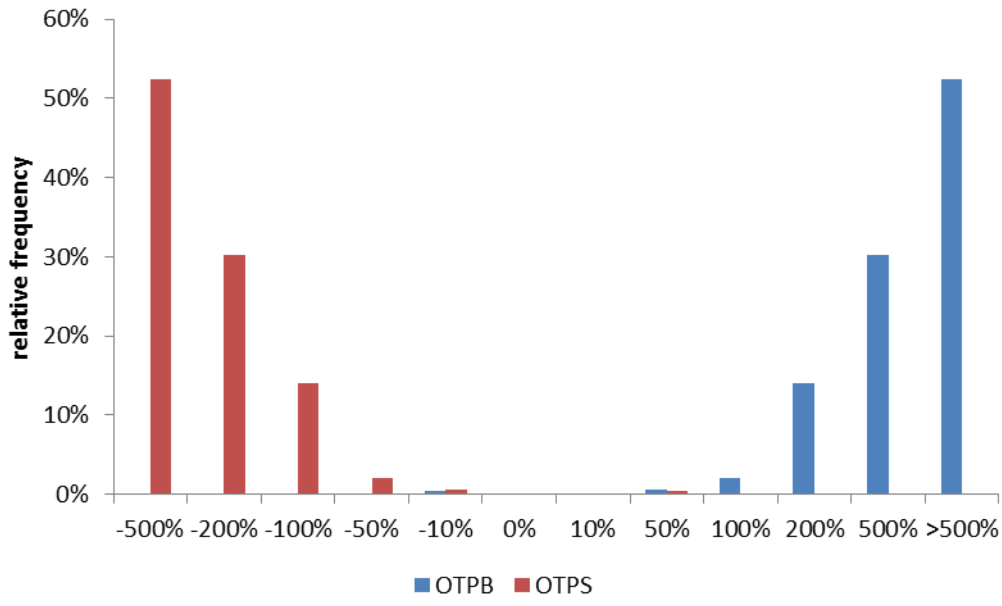


Figure 5.5: Histogram of Ratio (3.1) of buyers' and sellers' prices and portfolio holdings constraints in non-negativity form

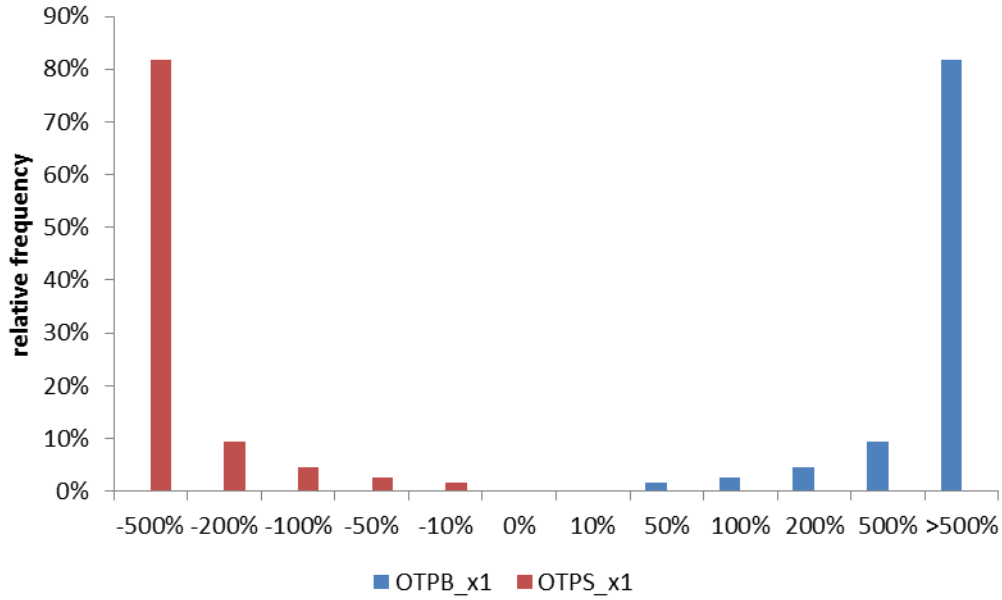


Figure 5.6: Histogram of Ratio (3.1) of buyers' and sellers' prices and portfolio holdings constraints of shape L_1 and size x=1

Figure 5.5 and Figure 5.6 clarify that it is essential with respect to economic significance to distinguish between buyers' and sellers' positions. In other words, if both positions are not analyzed separately, optimization potential—lower prices for buyers and higher prices for sellers—remains unused. Moreover, Figure 5.6 contains an example for the rare cases where buyers' prices exceed those of sellers (see, e.g., category +50% and red column; a positive number indicates that the reference model seller produces a price that is less than the one for buyers; see Section 2.2.2 for an economic explanation).

In addition, factors play a role when analyzing economic significance of buyers' and sellers' positons as the following exemplary figure demonstrates (all figures can be found in Appendix 5.2.1.1.3 starting with Figure Appendix 5.2.1.1.3k):

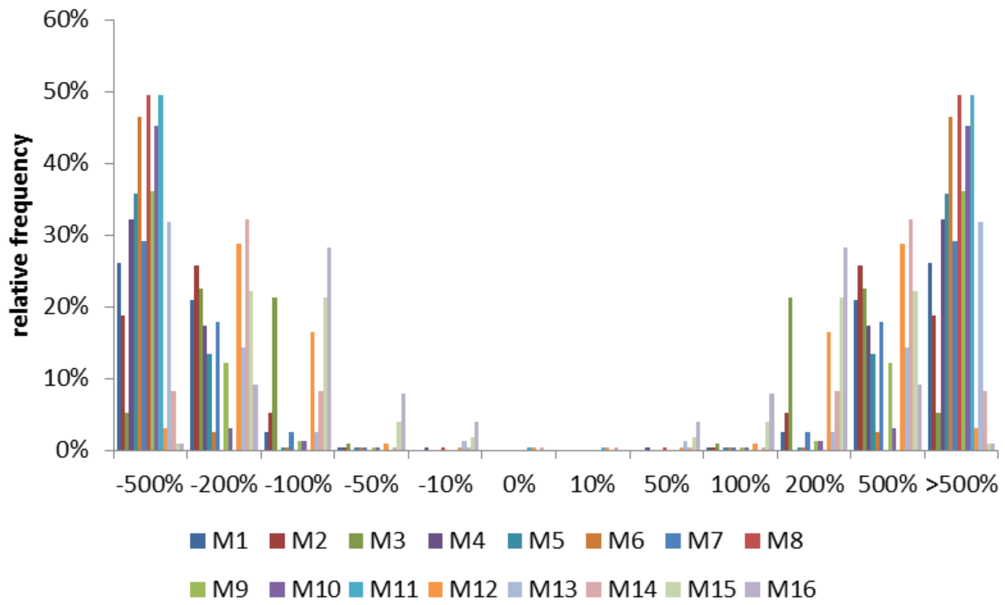


Figure 5.7: **Histogram of Ratio (3.1) of buyers' and sellers' prices and portfolio holdings constraints in non-negativity form broken down by factors**

Multi-factor models produce lesser Ratios (3.1) than single-factor models since their optimization potential is reduced due to more accounting characteristics constraints. Nevertheless, they are also unable to produce Ratios (3.1) of less than 10% that have been defined as acceptable in Chapter III, Section 2.1.3.2. In other words, even for multi-factor models it is required to separate buyers' from sellers' positions.

3.3.2.2 "Similarity" of Price Differences of Different Shapes and Sizes of Constraints on Portfolio Holdings

While "magnitude" stresses the price differences between models, i.e., focuses on dissimilarities, "similarity" concentrates on the common aspects of models. "Similarity" is computed with the help of Area (3.3) (see Chapter III, Section 2.1.3.3) and regarded as high if Area (3.3) is less than 10%.

3.3.2.2.1 The Role of Shape

Since "magnitude" results in shapes that create economically significant differences, "similarity" cannot deliver other results. In fact, Appendix 5.2.1.2.1 shows that using different shapes of constraints on portfolio holdings translates into Areas (3.3) that all are in the > 500% region (irrespective of whether it is broken down by factor or not). In other words, the criterion "similarity" also stresses the economic significance of the role of shape.

3.3.2.2.2 The Role of Size

The role of size x is of high economic significance (similarly to shape): Different sizes x on constraints on portfolio holdings translate into Areas (3.3) that all are in the > 500% region (irrespective of whether it is broken down by factor or not) as Appendix 5.2.1.2.2 shows.

3.3.2.3 Summary on the Economic Significance of Shape and Size of Constraints on Portfolio Holdings

Shape and size x of the constraints on portfolio holdings as well as decision makers' position (buyer and seller) exert huge influence on buyers' and sellers' company values, i.e., are of high economic significance.

Therefore, on the one hand, a separate price determination of buyers' and sellers' position is mandatory. On the other hand, constraints on portfolio holdings (shape and size) should be modelled wisely. In particular, they should not be implied as, e.g., it is the case with regression approaches. Given that constraints in L_p-norm-form cannot be justified from institutional circumstances (see Section 2.1.2), a non-negativity constraint on portfolio holdings as the sole constraint should be imposed. Non-negativity constraints do not have to specify the size x, constraints in L_p-norm-form sometimes result in negative prices (due to the short sale in the (super-/supra-) replication portfolio), and the practical implementation of short sales for "normal" companies might prove difficult.

3.3.3 Economic Significance of One- versus Multi-Period Versions of the Optimize-the-Price Approach

Economic significance boils down to the question whether the steady state assumption of the one-period model might not be completely unrealistic. Therefore, “similarity” (Area (3.3)) of one- and multi-period models is the decisive aspect of economic significance when one- and multi-period models are compared.

3.3.3.1 Graphical Analysis

“Similarity” figures (figures regarding “magnitude” can be found in Appendix 5.2.2.2) differentiated between buyers and sellers and with non-negativity constraints on portfolio holdings reveal:

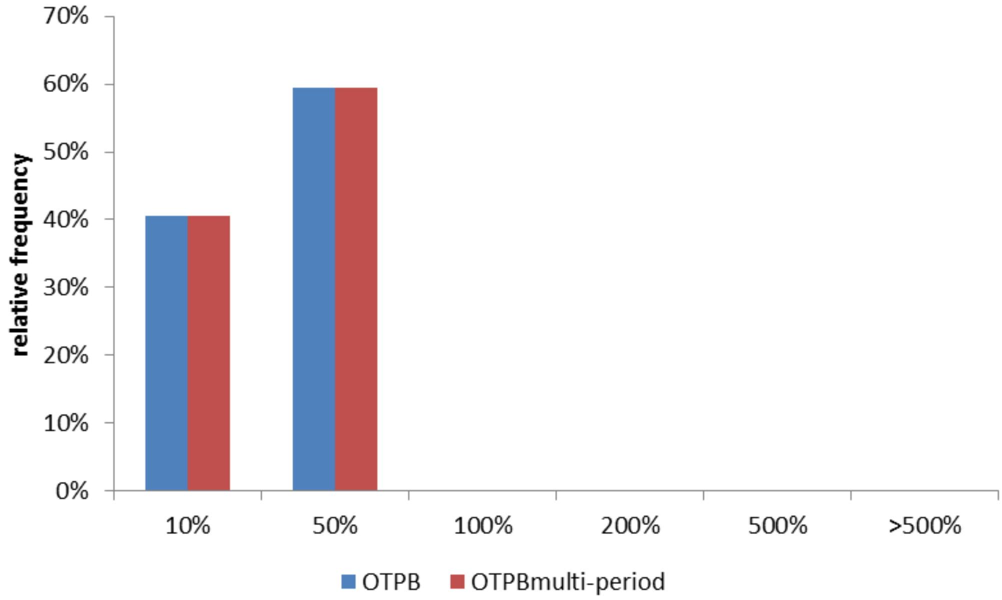


Figure 5.8: Histogram of Area (3.3) of price differences between one- and multi-period versions of the optimize-the-price approach with non-negativity constraints on portfolio holdings and buyers’ position

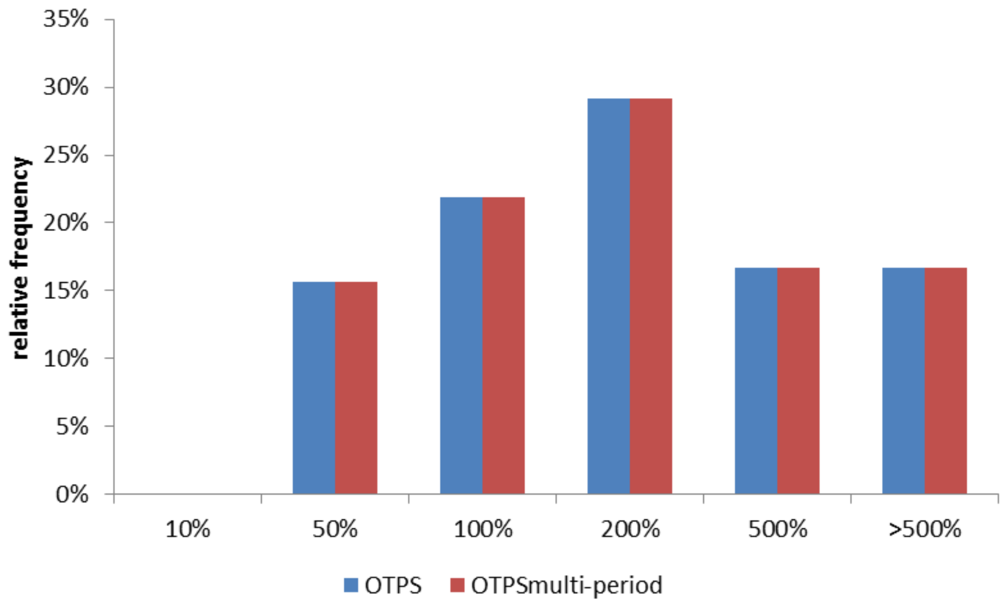


Figure 5.9: Histogram of Area (3.3) of price differences between one- and multi-period versions of the optimize-the-price approach with non-negativity constraints on portfolio holdings and sellers' position

Figure 5.8 illustrates that the steady state assumption of the one-period model partially works not bad for buyers, but is definitely bad for sellers. Obviously, the two differences between one- and multi-period models ((i) increasing number of constraints on accounting characteristics (5.12) and (5.13) and, thus, less optimization potential; (ii) time trend in the size x of constraints on portfolio holdings; see Appendix 5.2.2.1) compensate each other better in the case of buyers than in the case of sellers.

A look at different factors (see Appendix 5.2.2.3) might help to explain these results. Multi-factor models produce low "similarity" for buyers as opposed to the one-factor models M1 "Net Sales Or Revenues (SA)" and M8 "Book Value Of Common Equity". Yet these models do not possess the best explanatory power regarding company values—several accounting characteristics are needed to adequately reflect the multi-dimensionality of the figure cash flows that condense several value drivers into one figure—, a fact that might explain why there is not much of a difference between one- and multi-period variants of the optimize-the-price approach.

Finally, the different degrees of "similarity" between buyers and sellers deserve some consideration. Non-negativity constraints on portfolio holdings prevent buyers from selling expensive companies short in the (super-) replication portfolio why the price of the (super-) replication portfolio cannot be reduced that much. On the other hand, non-negativity constraints on portfolio holdings prove less restrictive for sellers since they are concerned with purchasing expensive companies in the (supra-) replication portfolio to end up with a high price for the company to be valued. In other words, non-negativity constraints exert a direct influence on the (super-), but only an indirect one (limits the number of purchases) on the (supra-) replication portfolio.

3.3.3.2 Summary on the Economic Significance of One- and Multi-Period Versions of the Optimize-the-Price Approach

Given the results regarding "similarity", it becomes apparent that price differences between one- and multi-period versions of the optimize-the-price approach are economically significant. Put differently, a one-period model is empirically not a good approximation of multi-period models and the steady state assumption implied by the one-period model is clearly violated empirically.

3.3.4 Economic Significance of Integrated (Optimize-the-Price Approach) versus Separated Approaches (Regressions)

Analyzing the integrated versus the separated approaches means examining empirically, how much optimization potential is not used if a free lunch (separated approach) is considered instead of an arbitration (integrated approach).

3.3.4.1 "Magnitude" of the Price Differences between Integrated (Optimize-the-Price Approach) versus Separated Approaches (Regressions)

Since the optimize-the-price approach is equipped with a constraint on portfolio holdings that matches the one implied by the respective regression approaches (see Section 3.2.1), it is not necessary to distinguish between shape and size of the constraints on portfolio holdings. Therefore, the results regarding "magnitude" can be directly accessed:

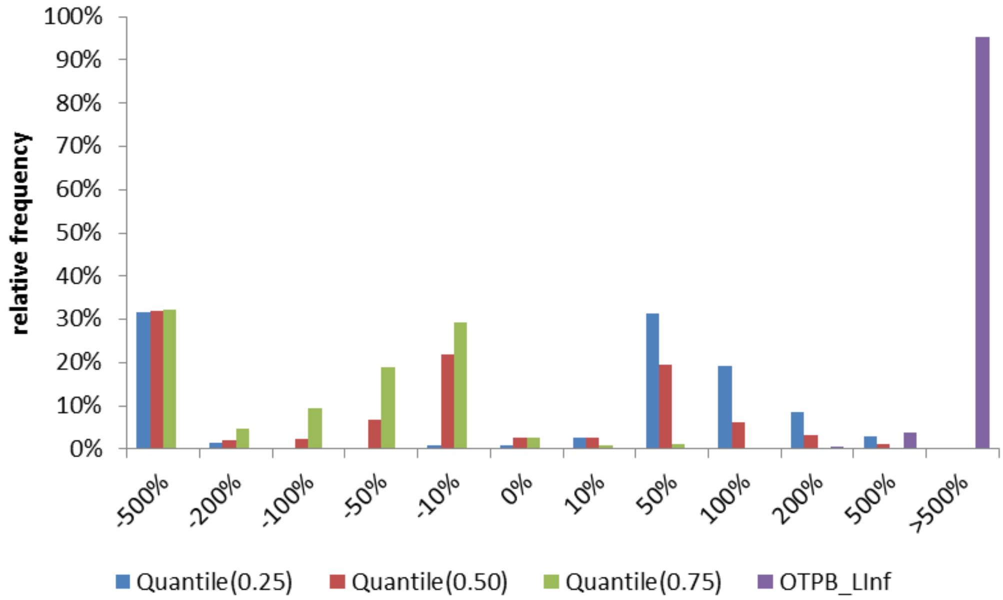


Figure 5.10: Histogram of Ratio (3.1) of buyers' price differences between quantile regressions and the optimize-the-price approach with a constraint on portfolio holdings in L_1-norm-form[13]

Figure 5.10 illustrates, first, the huge differences between prices computed with the help of regressions and the optimize-the-price approach and, second, that the prices of the optimize-the price-approach are well below the ones obtained by means of regressions.—Similar results are obtained for OLS and WLS regressions; see Appendix 5.2.3.2.1 starting with Figure Appendix 5.2.3.2.1g.

Principally the same picture is obtained for sellers. This time, however, the price of the seller following the optimize-the-price approach exceeds the one obtained from regressions (see Appendix 5.2.3.2.1):

[13] Recall that the dual program of the optimize-the-price approach must be compared to the primal program of regressions and a L_∞-constraint in the primal program translates into a L_1-constraint in the dual program.

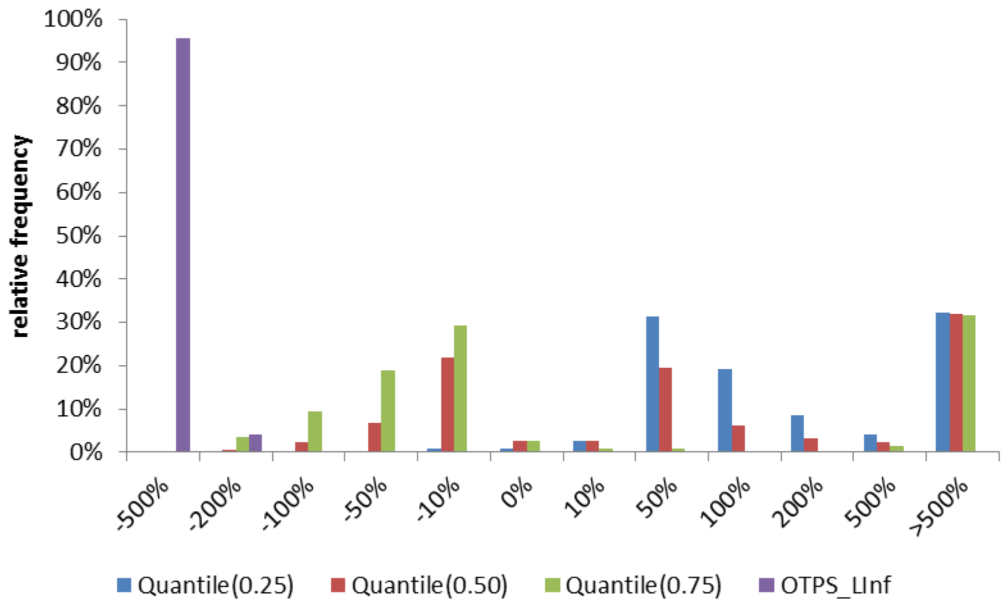


Figure 5.11: Histogram of Ratio (3.1) of sellers' price differences between quantile regressions and the optimize-the-price approach with a constraint on portfolio holdings in L_1-norm-form

In other words, focusing on a free lunch instead of following arbitration leaves an economically highly significant optimization potential on the table. Moreover, it becomes clear that regressions formally seem to represent buyers' position because they minimize prices (see Chapter IV, Formula (4.5)). However, this section illustrates that the focus on free lunches instead of arbitration cannot at all address the buyers' position.

Finally, note, first, that regressions without an intercept do not change the findings so far: Regressions without intercept are the theoretically superior model (see Chapter IV, Section 3.1.3) and produce economically significant lower prices (see Appendix 5.2.3.1.1). Nevertheless they ignore arbitration and, hence, deviate economically significantly from prices of the optimize-the-price approach. Second, all results do not change if broken down by factors, see Appendices 5.2.3.2.1 and 5.2.3.3.1.

3.3.4.2 “Similarity” of the Price Differences between Integrated (Optimize-the-Price Approach) versus Separated Approaches (Regressions)

Since “magnitude” illustrates economically significant price differences between the optimize-the-price approach and regressions, “similarity” cannot deliver different results. In fact, Appendices 5.2.3.2.2 and 5.2.3.3.2 demonstrate the economically significant low degree of “similarity” between both approaches that also holds when broken down by factors and for regressions with and without intercept.

3.3.4.3 Summary on the Economic Significance of Integrated (Optimize-the-Price Approach) versus Separated Approaches (Regressions)

The price differences between the integrated (optimize-the-price approach) and the separated approaches (regressions) are of very high economic significance. Therefore, ignoring—like regressions do—the price optimization potential of the company to be valued (arbitration) and instead focusing only on free lunches implies an economically significant mispricing. In particular, regressions can neither be associated with buyers’ nor sellers’ perspectives.

4 Conclusion

Valuing businesses by means of present values is the only correct approach from an asset pricing theory perspective because present values follow from no arbitrage, i.e., the intuition that a positive investment is needed to generate a positive return meaning that one gets nothing for free. However, no-arbitrage theory is difficult to translate into applicable valuation models for companies. Therefore, combining the practicability of accounting-based valuation models with the theoretical rigor of asset pricing theory might bring business valuation a visible step forward.—The optimize-the-price approach is our solution to the business valuation task:

From a theoretical perspective, the optimize-the-price approach is based on the economic principle and is able to integrate constraints on portfolio holdings that are in line with the institutional environment and market usages. Moreover, the optimize-the-price approach can distinguish between buyers' and sellers' position, use the mispricing potential of the company to be valued (arbitration) instead of focusing only on mispricing of other companies (free lunch), and can integrate synergies, multi-period valuations as well as risk.

From an empirical perspective, the price differences between the integrated (optimize-the-price approach) and the separated approaches (regressions) as well as price differences between buyers and sellers are of very high economic significance measured with the help of "magnitude" and "similarity".

From the perspective of a practical business valuation, a multi-period version of the optimize-the-price approach together with a non-negativity constraint on portfolio holdings is suited best: First, the steady state assumption of a one-period model is not given in reality. Second, constraints in L_p-norm-form cannot be justified from institutional circumstances and sometimes result in negative prices (due to the short sale in the (super-/supra-) replication portfolio). Moreover, the practical implementation of short sales for "normal" companies might prove difficult.

Appendix

Appendix 1 Lagrange Duality

Appendix 1.1 A Primer on the Implementation Steps of Lagrange Duality

Before Lagrange duality is applied to the specific problems of Chapter II, it is useful to present a general overview of how primal programs are transformed into dual programs. Once such a central theme is developed, it will be easier to follow the specific derivations.

By giving this primer, we follow the textbook of Boyd/Vandenberghe (2009).

Appendix 1.1.1 Fundamental Relation Between Primal and Dual Program

Starting point is the primal problem in standard form, i.e., a minimization problem where inequality constraints are in less-than-or-equal-to-zero form (see Boyd/Vandenberghe (2009), Formula 5.1, p. 215)

(A1.1)

$$\begin{aligned} &\text{minimize } f_0(z) \\ &\text{subject to } f_i(z) \leq 0, \quad i = 1, \dots, m \\ &\qquad\qquad\quad h_i(z) = 0, \quad i = 1, \dots, p \end{aligned}$$

with decision variables $z \in \mathbb{R}^n$.

The constraints are connected with the objective function by means of a Lagrange function (see Boyd/Vandenberghe (2009), p. 216).

(A1.2)

$$g(\lambda, \nu) = \inf_z L(x, \lambda, \nu) = \inf_z \left(f_0(z) + \sum_{i=1}^{m} \lambda_i f_i(z) + \sum_{i=1}^{p} \nu_i h_i(z) \right)$$

After optimizing the Lagrange function, the dual problem emerges (see Boyd/Vandenberghe (2009), Formula 5.16, p. 223):

(A1.3)

$$\begin{aligned} &\text{maximize } g(\lambda, \nu) \\ &\text{subject to } \lambda \geq 0 \end{aligned}$$

where ν is not sign-constrained since it is the Langrange multiplier of the constraints in equation form of the primal problem (A1.1).

Appendix 1.1.2 Steps of Lagrange Duality

To transform a primal problem to the dual problem (A1.3), the following steps must be mastered:

1st step: reformulate the primal problem in standard form where necessary

2nd step: write the Lagrange function

3rd step: differentiate the Lagrange function with respect to decision variables z (but not with respect to the Lagrange multipliers λ and ν)

4th step: group the Lagrange function by the variables z

5th step: insert the necessary conditions into the Lagrange function

6th step: formulate the dual problem in accordance with (A1.3)
Note in this connection that the necessary conditions from the 5th step guarantee that the objective function of the dual problem assumes a finite value.

7th step: transform the maximization problem into a minimization problem by multiplying the objective by -1
The original dual program of empirical asset pricing models is a minimization problem (valuation errors are minimized). Consequently, the dual problem is a maximization problem. For economic interpreta-

tions—relation to the economic principle—a minimization problem is superior. Hence, the maximization problem is multiplied by -1.

8th step: reformulate some of the constraints of the dual program to obtain a better economic interpretation for them

Appendix 1.2 Lagrange Dual of the Superordinate Category Regression Approaches

Appendix 1.2.1 Primal Program

According to Chapter II (Section 3.1, Formulas (2.52) to (2.21)) the general optimization problem of regression approaches reads

(2.52)

$$\min_{\substack{\mu_1^+,\mu_1^-,\ldots,\mu_n^+,\mu_n^-,\\ \beta_0,\beta_1,\ldots,\beta_m}} x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1}{p}}$$

s.t.

(2.53)

overestimation: $\varepsilon_i^{*+} = \beta_0 + \sum_{j=1}^{m} A_{1,j}^* \beta_j - y_1^* > 0$

$$\varepsilon_1^{*+} \leq \mu_1^+$$
$$\vdots$$
$$\varepsilon_n^{*+} \leq \mu_n^+$$

underestimation: $\varepsilon_i^{*-} = \beta_0 + \sum_{j=1}^{m} A_{1,j}^* \beta_j - y_1^* < 0$

$$\varepsilon_1^{*-} \geq -\mu_1^- \text{ or } -\varepsilon_1^{*-} \leq \mu_1^-$$
$$\vdots$$
$$\varepsilon_n^{*-} \geq -\mu_n^- \text{ or } -\varepsilon_n^{*-} \leq \mu_n^-$$

(2.54)

$$\mu_1^+ \geq 0, \mu_1^- \geq 0, \ldots, \mu_n^+ \geq 0, \mu_n^- \geq 0, \beta_0 \in \mathbb{R}, \beta_1 \in \mathbb{R}, \ldots, \beta_m \in \mathbb{R}$$

with

(2.21)

$$\begin{pmatrix} y_1^* \\ \vdots \\ y_n^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

$$\begin{pmatrix} A_{1,1}^* & \cdots & A_{1,m}^* \\ \vdots & \ddots & \vdots \\ A_{n,1}^* & \cdots & A_{n,m}^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,m} \end{pmatrix}$$

$$\begin{pmatrix} \varepsilon_1^* \\ \vdots \\ \varepsilon_n^* \end{pmatrix} = \underbrace{\begin{pmatrix} \omega_{1,1} & \cdots & \omega_{1,n} \\ \vdots & \ddots & \vdots \\ \omega_{n,1} & \cdots & \omega_{n,n} \end{pmatrix}}_{\equiv \omega} \cdot \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

where x is greater than zero and denotes a scaling factor.

The primal program (2.52) to (2.21) is then formulated in standard form (1st step) by plugging in for ε and expressing all constraints in less-than-or-equal-to-zero form:

(A1.4)

$$\min_{\substack{\mu_1^+,\mu_1^-,\ldots,\mu_n^+,\mu_n^-, \\ \beta_0,\beta_1,\ldots,\beta_m}} x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p \right]^{\frac{1}{p}}$$

s.t.

$$\beta_0 + \sum_{j=1}^{m} A_{1,j}^* \beta_j - y_1^* - \mu_1^+ \leq 0$$

$$-\beta_0 - \sum_{j=1}^{m} A_{1,j}^* \beta_j + y_1^* - \mu_1^- \leq 0$$

$$\vdots$$

$$\beta_0 + \sum_{j=1}^{m} A_{n,j}^* \beta_j - y_n^* - \mu_n^+ \leq 0$$

$$-\beta_0 - \sum_{j=1}^{m} A_{n,j}^* \beta_j + y_n^* - \mu_n^- \leq 0$$

$$-\mu_1^+ \leq 0$$
$$-\mu_1^- \leq 0$$
$$\vdots$$
$$-\mu_n^+ \leq 0$$
$$-\mu_n^- \leq 0$$

$$\beta_0 \in \mathbb{R}, \beta_1 \in \mathbb{R}, \dots, \beta_m \in \mathbb{R}$$

Appendix 1.2.2 Preparing for Dualization

The Lagrange function of the primal problem in standard form (A1.4) reads (2nd step)

(A1.5)

$$
\begin{aligned}
L = {} & x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}} \\
& + \lambda_1^+ \left(\beta_0 + \sum_{j=1}^{m} A_{1,j}^* \beta_j - y_1^* - \mu_1^+\right) + \lambda_1^- \left(-\beta_0 - \sum_{j=1}^{m} A_{1,j}^* \beta_j + y_1^* - \mu_1^-\right) \\
& + \cdots \\
& + \lambda_n^+ \left(\beta_0 + \sum_{j=1}^{m} A_{n,j}^* \beta_j - y_n^* - \mu_n^+\right) + \lambda_n^- \left(-\beta_0 - \sum_{j=1}^{m} A_{n,j}^* \beta_j + y_n^* - \mu_n^-\right) \\
& + \gamma_1^+(-\mu_1^+) + \gamma_1^-(-\mu_1^-) + \cdots + \gamma_n^+(-\mu_n^+) + \gamma_n^-(-\mu_n^-)
\end{aligned}
$$

Forming necessary conditions (3rd step) delivers

$$
\begin{aligned}
\frac{\partial L}{\partial \mu_1^+} = {} & x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot p \cdot (\mu_1^+)^{p-1} \\
& - \lambda_1^+ - \gamma_1^+ = 0 \\
\frac{\partial L}{\partial \mu_1^-} = {} & x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot p \cdot (\mu_1^-)^{p-1} \\
& - \lambda_1^- - \gamma_1^- = 0 \\
& \vdots
\end{aligned}
$$

$$\frac{\partial L}{\partial \mu_n^+} = x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot p \cdot (\mu_n^+)^{p-1}$$
$$-\lambda_n^+ - \gamma_n^+ = 0$$

$$\frac{\partial L}{\partial \mu_n^-} = x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot p \cdot (\mu_n^-)^{p-1}$$
$$-\lambda_n^- - \gamma_n^- = 0$$

$$\frac{\partial L}{\partial \beta_0} = \lambda_1^+ - \lambda_1^- + \cdots + \lambda_n^+ - \lambda_n^- = 0$$
$$\frac{\partial L}{\partial \beta_1} = \lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* = 0$$
$$\vdots$$
$$\frac{\partial L}{\partial \beta_m} = \lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* = 0$$

because $\frac{1}{p} - 1 = \frac{1-p}{p}$.

Collecting decision variables $\mu_1^+, \mu_1^-, \dots, \mu_n^+, \mu_n^-$ and $\beta_0, \beta_1, \dots, \beta_m$ in the Lagrange function (A1.5) delivers (4th step)

(A1.6)

$$L = x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}}$$
$$-\lambda_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \cdots - \lambda_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^-$$
$$-\gamma_1^+ \cdot \mu_1^+ - \gamma_1^- \cdot \mu_1^- - \cdots - \gamma_n^+ \cdot \mu_n^+ - \gamma_n^- \cdot \mu_n^-$$
$$+\beta_0 \sum_{i=1}^{n} \lambda_i^+ - \beta_0 \sum_{i=1}^{n} \lambda_i^-$$
$$+\beta_1 \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,1}^* - \beta_1 \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,1}^*$$
$$+\cdots$$
$$+\beta_m \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,m}^* - \beta_m \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,m}^*$$

$$-\sum_{i=1}^{n} \lambda_i^+ \cdot y_i^* + \sum_{i=1}^{n} \lambda_i^- \cdot y_i^*$$

To prepare for inserting the necessary conditions into the Lagrange function (5th step), an intermediate transformation is recommended. Each necessary condition with respect to μ_i^+ and μ_i^- is multiplied by its μ_i^+ or μ_i^- respectively. Then, all these multiplied necessary conditions are added, a procedure that results in

$$\begin{aligned}
&x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_1^+)^p \\
&-\lambda_1^+ \cdot \mu_1^+ - \gamma_1^+ \cdot \mu_1^+ \\
&+x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_1^-)^p \\
&-\lambda_1^- \cdot \mu_1^- - \gamma_1^- \cdot \mu_1^- \\
&+\cdots \\
&+x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_n^+)^p \\
&-\lambda_n^+ \cdot \mu_n^+ - \gamma_n^+ \cdot \mu_n^+ \\
&+x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_n^-)^p \\
&-\lambda_n^- \cdot \mu_n^- - \gamma_n^- \cdot \mu_n^- = 0
\end{aligned}$$

i.e., after collecting associated μ_i^+ and μ_i^-

$$\begin{aligned}
&x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \\
&\cdot \left[\underbrace{(1-\tau) \cdot (\mu_1^+)^p + \tau \cdot (\mu_1^-)^p + \cdots + (1-\tau) \cdot (\mu_n^+)^p + \tau \cdot (\mu_n^-)^p}_{=(1-\tau)\cdot\sum_{i=1}^{n}\left(\mu_i^+\right)^p + \tau\cdot\sum_{i=1}^{n}\left(\mu_i^-\right)^p}\right] \\
&-\lambda_1^+ \cdot \mu_1^+ - \gamma_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \gamma_1^- \cdot \mu_1^- \\
&-\cdots \\
&-\lambda_n^+ \cdot \mu_n^+ - \gamma_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- - \gamma_n^- \cdot \mu_n^- = 0
\end{aligned}$$

and thus

(A1.7)

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}}$$
$$-\lambda_1^+ \cdot \mu_1^+ - \gamma_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \gamma_1^- \cdot \mu_1^-$$
$$- \cdots$$
$$-\lambda_n^+ \cdot \mu_n^+ - \gamma_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- - \gamma_n^- \cdot \mu_n^- = 0$$

because $\frac{1-p}{p} + 1 = \frac{1}{p}$.

Plugging the aggregated necessary conditions for μ (A1.7) and the necessary conditions for β into Lagrange function (A1.6) yields (5th step)

$$L = x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}}$$
$$\underbrace{\begin{aligned}&-\lambda_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \cdots - \lambda_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- \\ &-\gamma_1^+ \cdot \mu_1^+ - \gamma_1^- \cdot \mu_1^- - \cdots - \gamma_n^+ \cdot \mu_n^+ - \gamma_n^- \cdot \mu_n^-\end{aligned}}_{=0\ because\ of\ \text{(A1.7)}}$$
$$\underbrace{+\beta_0 \sum_{i=1}^{n} \lambda_i^+ - \beta_0 \sum_{i=1}^{n} \lambda_i^-}_{=0\ because\ of\ \frac{\partial L}{\partial \beta_0}=0}$$
$$\underbrace{+\beta_1 \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,1}^* - \beta_1 \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,1}^*}_{=0\ because\ of\ \frac{\partial L}{\partial \beta_1}=0}$$
$$+ \cdots$$
$$\underbrace{+\beta_m \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,m}^* - \beta_m \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,m}^*}_{=0\ because\ of\ \frac{\partial L}{\partial \beta_m}=0}$$
$$-\sum_{i=1}^{n} \lambda_i^+ \cdot y_i^* + \sum_{i=1}^{n} \lambda_i^- \cdot y_i^*$$

Consequently, it is obtained

(A1.8)

$$L = -\sum_{i=1}^{n} \lambda_i^+ \cdot y_i^* + \sum_{i=1}^{n} \lambda_i^- \cdot y_i^*$$

Appendix 1.2.3 Dual Program: First Version

The adaption of the general dual program (A1.3) to the regression environment (6^{th} step) reads

(A1.9)

$$\max_{\substack{\lambda_1^+,\lambda_1^-,\dots,\lambda_n^+,\lambda_n^-,\\ \gamma_1^+,\gamma_1^-,\dots,\gamma_n^+,\gamma_n^-}} -\sum_{i=1}^{n} \lambda_i^+ \cdot y_i^* + \sum_{i=1}^{n} \lambda_i^- \cdot y_i^*$$

s.t.

$$\frac{\partial L}{\partial \mu_1^+} = x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot p \cdot (\mu_1^+)^{p-1}$$
$$-\lambda_1^+ - \gamma_1^+ = 0$$

$$\frac{\partial L}{\partial \mu_1^-} = x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot p \cdot (\mu_1^-)^{p-1}$$
$$-\lambda_1^- - \gamma_1^- = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial \mu_n^+} = x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot p \cdot (\mu_n^+)^{p-1}$$
$$-\lambda_n^+ - \gamma_n^+ = 0$$

$$\frac{\partial L}{\partial \mu_n^-} = x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot p \cdot (\mu_n^-)^{p-1}$$
$$-\lambda_n^- - \gamma_n^- = 0$$

$$\frac{\partial L}{\partial \beta_0} = \lambda_1^+ - \lambda_1^- + \cdots + \lambda_n^+ - \lambda_n^- = 0$$

$$\frac{\partial L}{\partial \beta_1} = \lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial \beta_m} = \lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* = 0$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \ldots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$
$$\gamma_1^+ \geq 0, \gamma_1^- \geq 0, \ldots, \gamma_n^+ \geq 0, \gamma_n^- \geq 0$$

The objective function of the maximization problem (A1.9) reads

$$\max_{\substack{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-,\\ \gamma_1^+,\gamma_1^-,\ldots,\gamma_n^+,\gamma_n^-}} -\sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^*$$

Hence, the maximization problem (A1.9) easily translates into a minimization problem (7th step) and the first version of the dual program is obtained.

(A1.10)

$$\min_{\substack{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-,\\ \gamma_1^+,\gamma_1^-,\ldots,\gamma_n^+,\gamma_n^-}} \sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^*$$

s.t.

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_1^+)^{p-1} - \lambda_1^+ - \gamma_1^+ = 0$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_1^-)^{p-1} - \lambda_1^- - \gamma_1^- = 0$$

$$\vdots$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_n^+)^{p-1} - \lambda_n^+ - \gamma_n^+ = 0$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_n^-)^{p-1} - \lambda_n^- - \gamma_n^- = 0$$

$$\lambda_1^+ - \lambda_1^- + \cdots + \lambda_n^+ - \lambda_n^- = 0$$

$$\lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* = 0$$
$$\vdots$$
$$\lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* = 0$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \ldots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$
$$\gamma_1^+ \geq 0, \gamma_1^- \geq 0, \ldots, \gamma_n^+ \geq 0, \gamma_n^- \geq 0$$

Appendix 1.2.4 Dual Program: Final Version (8th step)

The necessary conditions of the dual program: first version (A1.10) are a function of the decision variables μ_i^+ and μ_i^- of the primal program (A1.4). For that reason, dual program (A1.10) has no ready economic interpretation and, thus, cannot serve as economic model selection criterion. Instead, μ_i^+ and μ_i^- must be removed from the constraints of dual program (A1.10).

Appendix 1.2.4.1 Determination and thus Elimination of $\boldsymbol{\mu_i^+}$ and $\boldsymbol{\mu_i^-}$

Appendix 1.2.4.1.1 Equations that contain $\boldsymbol{\mu_i^+}$ and $\boldsymbol{\mu_i^-}$

The following equations contain μ_i^+ and μ_i^- and, hence, are candidates for the elimination of μ_i^+ and μ_i^-.

– Aggregated necessary condition for μ_i^+ and μ_i^-

  (A1.7)

$$\begin{aligned} & x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p \right]^{\frac{1}{p}} \\ & -\lambda_1^+ \cdot \mu_1^+ - \gamma_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \gamma_1^- \cdot \mu_1^- \\ & - \cdots \\ & -\lambda_n^+ \cdot \mu_n^+ - \gamma_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- - \gamma_n^- \cdot \mu_n^- = 0 \end{aligned}$$

– Necessary conditions with respect to μ_i^+ and μ_i^-

(A1.11)

$$x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau)\cdot(\mu_1^+)^{p-1} - \lambda_1^+ - \gamma_1^+ = 0$$

$$x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau\cdot(\mu_1^-)^{p-1} - \lambda_1^- - \gamma_1^- = 0$$

$$\vdots$$

$$x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau)\cdot(\mu_n^+)^{p-1} - \lambda_n^+ - \gamma_n^+ = 0$$

$$x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau\cdot(\mu_n^-)^{p-1} - \lambda_n^- - \gamma_n^- = 0$$

Appendix 1.2.4.1.2 Solving the Equation System with respect to $\boldsymbol{\mu_i^+}$ and $\boldsymbol{\mu_i^-}$

The equation system (A1.7) and (A1.11) is rich enough to allow for a complete determination μ_i^+ and μ_i^- and, hence, elimination of μ_i^+ and μ_i^- from the constraints of dual program (A1.10). Intuitively, it is proceeded as follows: From the necessary conditions (A1.11) $\mu_1^-, \dots, \mu_n^-$ are expressed as functions of μ_1^+. These expressions are then inserted into (A1.7).

Restructuring (A1.7) leads to

$$\left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1}{p}} = \frac{(\lambda_1^+ + \gamma_1^+)\cdot\mu_1^+ + (\lambda_1^- + \gamma_1^-)\cdot\mu_1^- + \cdots + (\lambda_n^+ + \gamma_n^+)\cdot\mu_n^+ + (\lambda_n^- + \gamma_n^-)\cdot\mu_n^-}{x}$$

since x is by definition greater than zero.

Reshuffling (A1.11) gains

$$x \cdot \left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p + \tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau)\cdot(\mu_1^+)^{p-1} = \lambda_1^+ + \gamma_1^+$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_1^-)^{p-1} = \lambda_1^- + \gamma_1^-$$

$$\vdots$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_n^+)^{p-1} = \lambda_n^+ + \gamma_n^+$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_n^-)^{p-1} = \lambda_n^- + \gamma_n^-$$

or rather (after taking the $\frac{1}{1-p}$th power)

(A1.12)

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}} = x^{\frac{1}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_1^+ \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$$

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}} = x^{\frac{1}{p-1}} \cdot \tau^{\frac{1}{p-1}} \cdot \mu_1^- \cdot (\lambda_1^- + \gamma_1^-)^{\frac{1}{1-p}}$$

$$\vdots$$

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}} = x^{\frac{1}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_n^+ \cdot (\lambda_n^+ + \gamma_n^+)^{\frac{1}{1-p}}$$

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}} = x^{\frac{1}{p-1}} \cdot \tau^{\frac{1}{p-1}} \cdot \mu_n^- \cdot (\lambda_n^- + \gamma_n^-)^{\frac{1}{1-p}}$$

τ is located between 0 and 1 (see, e.g., Koenker (2005), p. 5); λ_i^+ and λ_i^- as well as γ_i^+ and γ_i^- are nonnegative in the dual program (A1.10). Moreover, λ_i^+ and γ_i^+ (λ_i^- and γ_i^-) will not be both identical to zero. If a purchase occurs in the optimum, λ_i^+ will be greater than zero and the nonnegativity constraint on purchases will not bind, i.e., $\gamma_i^+ = 0$. If a sale occurs in the optimum, λ_i^+ will be equal to zero, but γ_i^+ will be greater than zero due to the now binding nonnegativity constraint.

Since the left-hand sides of (A1.12) are identical, a relation between $\mu_1^-, \dots, \mu_n^-$ and μ_1^+ can be established:

$$x^{\frac{1}{p-1}} \cdot \tau^{\frac{1}{p-1}} \cdot \mu_1^- \cdot (\lambda_1^- + \gamma_1^-)^{\frac{1}{1-p}} = x^{\frac{1}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_1^+ \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$$

$$\vdots$$

$$x^{\frac{1}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_n^+ \cdot (\lambda_n^+ + \gamma_n^+)^{\frac{1}{1-p}} = x^{\frac{1}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_1^+ \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$$

$$x^{\frac{1}{p-1}} \cdot \tau^{\frac{1}{p-1}} \cdot \mu_n^- \cdot (\lambda_n^- + \gamma_n^-)^{\frac{1}{1-p}} = x^{\frac{1}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_1^+ \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$$

i.e.,

(A1.13)

$$\mu_1^- = \frac{(1-\tau)^{\frac{1}{p-1}}}{\tau^{\frac{1}{p-1}}} \cdot \mu_1^+ \cdot \frac{(\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}}{(\lambda_1^- + \gamma_1^-)^{\frac{1}{1-p}}}$$

$$\vdots$$

$$\mu_n^+ = \frac{(1-\tau)^{\frac{1}{p-1}}}{(1-\tau)^{\frac{1}{p-1}}} \cdot \mu_1^+ \cdot \frac{(\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}}{(\lambda_n^+ + \gamma_n^+)^{\frac{1}{1-p}}}$$

$$\mu_n^- = \frac{(1-\tau)^{\frac{1}{p-1}}}{\tau^{\frac{1}{p-1}}} \cdot \mu_1^+ \cdot \frac{(\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}}{(\lambda_n^- + \gamma_n^-)^{\frac{1}{1-p}}}$$

and from (A1.7)

(A1.14)

$$(\lambda_1^+ + \gamma_1^+) \cdot \mu_1^+ + (\lambda_1^- + \gamma_1^-) \cdot \mu_1^- + \cdots + (\lambda_n^+ + \gamma_n^+) \cdot \mu_n^+ + (\lambda_n^- + \gamma_n^-) \cdot \mu_n^-$$
$$= x^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_1^+ \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$$

since it holds $x^{\frac{1}{p-1}+1} = x^{\frac{1+p-1}{p-1}} = x^{\frac{p}{p-1}}$.

Inserting (A1.13) into (A1.14) gains

$$(\lambda_1^+ + \gamma_1^+) \cdot \mu_1^+ + (\lambda_1^- + \gamma_1^-) \cdot \frac{(1-\tau)^{\frac{1}{p-1}}}{\tau^{\frac{1}{p-1}}} \cdot \mu_1^+ \cdot \frac{(\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}}{(\lambda_1^- + \gamma_1^-)^{\frac{1}{1-p}}}$$
$$+ \cdots$$
$$+(\lambda_n^+ + \gamma_n^+) \cdot \frac{(1-\tau)^{\frac{1}{p-1}}}{(1-\tau)^{\frac{1}{p-1}}} \cdot \mu_1^+ \cdot \frac{(\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}}{(\lambda_n^+ + \gamma_n^+)^{\frac{1}{1-p}}}$$

$$+(\lambda_n^- + \gamma_n^-) \cdot \frac{(1-\tau)^{\frac{1}{p-1}}}{\tau^{\frac{1}{p-1}}} \cdot \mu_1^+ \cdot \frac{(\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}}{(\lambda_n^- + \gamma_n^-)^{\frac{1}{1-p}}}$$

$$= x^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{p-1}} \cdot \mu_1^+ \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$$

Division[14] by μ_1^+ and $(1-\tau)^{\frac{1}{p-1}} \cdot (\lambda_1^+ + \gamma_1^+)^{\frac{1}{1-p}}$ yields

$$\begin{aligned}
&(\lambda_1^+ + \gamma_1^+)^{1-\frac{1}{1-p}} \cdot (1-\tau)^{\frac{1}{1-p}} \\
&+(\lambda_1^- + \gamma_1^-) \cdot \frac{1}{\tau^{\frac{1}{p-1}}} \cdot \frac{1}{(\lambda_1^- + \gamma_1^-)^{\frac{1}{1-p}}} \\
&+(\lambda_n^+ + \gamma_n^+) \cdot \frac{1}{(1-\tau)^{\frac{1}{p-1}}} \cdot \frac{1}{(\lambda_n^+ + \gamma_n^+)^{\frac{1}{1-p}}} \\
&+(\lambda_n^- + \gamma_n^-) \cdot \frac{1}{\tau^{\frac{1}{p-1}}} \cdot \frac{1}{(\lambda_n^- + \gamma_n^-)^{\frac{1}{1-p}}} = x^{\frac{p}{p-1}}
\end{aligned}$$

i.e.,

(A1.15)

$$\begin{aligned}
&(\lambda_1^+ + \gamma_1^+)^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{1-p}} + (\lambda_1^- + \gamma_1^-)^{\frac{p}{p-1}} \cdot \tau^{\frac{1}{1-p}} \\
&+\cdots \\
&+(\lambda_n^+ + \gamma_n^+)^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{1-p}} + (\lambda_n^- + \gamma_n^-)^{\frac{p}{p-1}} \cdot \tau^{\frac{1}{1-p}} = x^{\frac{p}{p-1}}
\end{aligned}$$

since it holds

$$(\lambda_1^- + \gamma_1^-)^1 \cdot (\lambda_1^- + \gamma_1^-)^{\frac{1}{p-1}} = (\lambda_1^- + \gamma_1^-)^{1+\frac{1}{p-1}} = (\lambda_1^- + \gamma_1^-)^{\frac{p}{p-1}}$$

Taking the $\frac{p}{p-1}$th root of (A1.15) finally delivers

(A1.16)

$$\begin{bmatrix}
(\lambda_1^+ + \gamma_1^+)^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{1-p}} + (\lambda_1^- + \gamma_1^-)^{\frac{p}{p-1}} \cdot \tau^{\frac{1}{1-p}} \\
+\cdots \\
+(\lambda_n^+ + \gamma_n^+)^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{1-p}} + (\lambda_n^- + \gamma_n^-)^{\frac{p}{p-1}} \cdot \tau^{\frac{1}{1-p}}
\end{bmatrix}^{\frac{p-1}{p}} = x$$

[14] In the case that asset 1 is overvalued, μ_1^+ will be zero. In that case, however, all equations will be expressed as a function of μ_1^- and the computations will follow the exactly same path.

(A1.16) contains the unobservable variables γ_i^+ and γ_i^-. However, since λ_i^+ and λ_i^- as well as γ_i^+ and γ_i^- are nonnegative in the dual program (A1.10), it holds

$$\lambda_i^+ + \gamma_i^+ \geq \lambda_i^+ \text{ and } \lambda_i^- + \gamma_i^- \geq \lambda_i^-$$

and, hence,

(A1.17)

$$\begin{bmatrix} (\lambda_1^+)^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{1-p}} + (\lambda_1^-)^{\frac{p}{p-1}} \cdot \tau^{\frac{1}{1-p}} \\ + \cdots \\ + (\lambda_n^+)^{\frac{p}{p-1}} \cdot (1-\tau)^{\frac{1}{1-p}} + (\lambda_n^-)^{\frac{p}{p-1}} \cdot \tau^{\frac{1}{1-p}} \end{bmatrix}^{\frac{p-1}{p}} \leq x$$

(A1.17) is the form of a constraint useful for integrating into the dual program (A1.10) because it no longer depends on μ_i^+ and μ_i^- as well as γ_i^+ and γ_i^-.

Appendix 1.2.4.2 Dual Program: Final Version

Substituting (A1.17) for the necessary conditions with respect to μ_i^+ and μ_i^- in the first form of dual program (A1.10), it is finally gained

(4.1)

$$\min_{\lambda_1^+, \lambda_1^-, \ldots, \lambda_n^+, \lambda_n^-} \sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^*$$

s.t.

(4.2)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$\begin{aligned} \lambda_1^+ - \lambda_1^- + \cdots + \lambda_n^+ - \lambda_n^- &= 0 \\ \lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* &= 0 \\ &\vdots \\ \lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* &= 0 \end{aligned}$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \ldots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$

(4.1) is the desired final form of the dual program.

Appendix 1.2.5 Dual Program for the Special Case: p=1

Since p is in the denominator of the final form of the dual program (4.1), the special case $p = 1$ cannot be directly seen from (4.1). However, this special case can be directly derived from the necessary conditions with respect to μ_i^+ and μ_i^- of the first form of dual program (A1.10):

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_1^+)^{p-1} - \lambda_1^+ - \gamma_1^+ = 0$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_1^-)^{p-1} - \lambda_1^- - \gamma_1^- = 0$$

$$\vdots$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot (\mu_n^+)^{p-1} - \lambda_n^+ - \gamma_n^+ = 0$$

$$x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot (\mu_n^-)^{p-1} - \lambda_n^- - \gamma_n^- = 0$$

With $p = 1$ the above necessary conditions simplify to

(A1.18)

$$\begin{aligned} x \cdot (1-\tau) - \lambda_1^+ - \gamma_1^+ &= 0 \\ x \cdot \tau - \lambda_1^- - \gamma_1^- &= 0 \\ &\vdots \\ x \cdot (1-\tau) - \lambda_n^+ - \gamma_n^+ &= 0 \\ x \cdot \tau - \lambda_n^- - \gamma_n^- &= 0 \end{aligned}$$

Since γ_i^+ and γ_i^- are nonnegative in the dual program (A1.10), (A1.18) can be written without explicit reference to γ_i^+ and γ_i^- as

$$\begin{aligned} x \cdot (1-\tau) - \lambda_1^+ &\leq 0 \\ x \cdot \tau - \lambda_1^- &\leq 0 \end{aligned}$$

$$\vdots$$

$$x \cdot (1-\tau) - \lambda_n^+ \leq 0$$

$$x \cdot \tau - \lambda_n^- \leq 0$$

and, finally,

(A1.19)

$$\begin{aligned}
&\lambda_1^+ \leq x \cdot (1-\tau) \\
&\lambda_1^- \leq x \cdot \tau \\
&\quad \vdots \\
&\lambda_n^+ \leq x \cdot (1-\tau) \\
&\lambda_n^- \leq x \cdot \tau
\end{aligned}$$

(A1.19) are the desired constraints for integrating into the dual program (A1.10).

Appendix 1.3 Lagrange Duality of the Superordinate Category Method of Multiples

Appendix 1.3.1 Primal Program

The method of multiples does not involve optimization (see Chapter II, Section 2.2.1) since the multiple—and, thus, company values—is determined completely as a function of company characteristics (see (2.23) to (2.27)).

Put differently, an optimization problem will be adequate to capture the method of multiples if it yields as outcome of the optimization the multiple

(A1.20)

$$\beta_j = f(y, A_j)$$

where $f(.)$ is defined in a way so that β_j always remains positive.

Note in addition that $f(.)$ has no subscript because the function is independent of specific accounting figures. It is, e.g., an arithmetic average meaning that arithmetic averages are computed for all accounting figures A_j.

Note in this connection that several multiples can be used in combination to determine prices (see (2.33)). However, each multiple is determined independently of other multiples, a fact that can be formalized as follows: The necessary condition of multiple β_j reads

(A1.21)

$$\frac{\partial G_j}{\partial \beta_j} = 0 = \beta_j - f(y, A_j)$$

where G is the yet unknown objective function of the optimization problem.

G_j can be determined from (A1.21) by means of integration:

$$G_j = \int \beta_j - f(y, A_j)\, d\beta_j = \frac{1}{2} \cdot \beta_j^2 - f(y, A_j) \cdot \beta_j + const$$

where $const$ denotes an arbitrary constant.

Even though technically arbitrary, $const$ should be specified so that a comparison to regression approaches can be established.—As long as $const$ does not depend on β_j this procedure will be innocuous. The same is true for adding constraints that do not influence the optimal value of β_j.

Such innocuous $const$ and constraints are any of those that optimize with respect to valuation errors μ_i^+ and μ_i^- since they play no role in the method of multiples. To remain comparable to the superordinate category of regression approaches, its objective function and constraints regarding over- and underestimation might be added.

This signifies that the following primal program in standard form can be used to describe the optimization problem of the method of multiples (1st step):

(A1.22)

$$\min_{\substack{\mu_1^+, \mu_1^-, \ldots, \mu_n^+, \mu_n^-, \\ \beta_1, \ldots, \beta_m, \\ \theta_1, \ldots, \theta_m}} \sum_{j=1}^{m} \frac{1}{2} \cdot \beta_j^2 - f(y, A_j) \cdot \beta_j + x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p \right]^{\frac{1}{p}}$$

s.t.

$$\sum_{j=1}^{m} A_{1,j}^{*}\theta_j - y_1^{*} - \mu_1^{+} \leq 0$$

$$-\sum_{j=1}^{m} A_{1,j}^{*}\theta_j + y_1^{*} - \mu_1^{-} \leq 0$$

$$\vdots$$

$$\sum_{j=1}^{m} A_{n,j}^{*}\theta_j - y_n^{*} - \mu_n^{+} \leq 0$$

$$-\sum_{j=1}^{m} A_{n,j}^{*}\theta_j + y_n^{*} - \mu_n^{-} \leq 0$$

$$-\mu_1^{+} \leq 0$$

$$-\mu_1^{-} \leq 0$$

$$\vdots$$

$$-\mu_n^{+} \leq 0$$

$$-\mu_n^{-} \leq 0$$

Note that adding a nonnegativity constraint for β_j is not adequate. On the one hand, because the method of multiples specifies the function $f(y, A_j)$ to be nonnegative. Adding a nonnegativity constraint for β_j overlooks this institutional feature of the method of multiples. On the other hand for formal reasons: Adding a nonnegativity constraint for β_j (or $-\beta_j \leq 0$ in standard form) would result in the following necessary condition for β_j

$$\frac{\partial L}{\partial \beta_\mathrm{j}} = \beta_j - f(y, A_j) - \lambda_{\beta_j} = 0$$

where λ_{β_j} (with $\lambda_{\beta_j} \geq 0$) denotes the Lagrange multiplier of the nonnegativity constraint for β_j.

Therefore a nonnegativity constraint on β_j cannot reproduce the core result of the method of multiples $\beta_j = f(y, A_j)$, but only $\beta_j \geq f(y, A_j)$.

Appendix 1.3.2 Preparing for Dualization

The Lagrange function of the primal problem in standard form (A1.22) reads (2nd step)

(A1.23)

$$
\begin{aligned}
L = & \sum_{j=1}^{m} \frac{1}{2} \cdot \beta_j^2 - f\left(y, A_j\right) \cdot \beta_j + x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1}{p}} \\
& + \lambda_1^+ \left(\sum_{j=1}^{m} A_{1,j}^* \theta_j - y_1^* - \mu_1^+ \right) + \lambda_1^- \left(- \sum_{j=1}^{m} A_{1,j}^* \theta_j + y_1^* - \mu_1^- \right) \\
& + \cdots \\
& + \lambda_n^+ \left(\sum_{j=1}^{m} A_{n,j}^* \theta_j - y_n^* - \mu_n^+ \right) + \lambda_n^- \left(- \sum_{j=1}^{m} A_{n,j}^* \theta_j + y_n^* - \mu_n^- \right) \\
& + \gamma_1^+(-\mu_1^+) + \gamma_1^-(-\mu_1^-) + \cdots + \gamma_n^+(-\mu_n^+) + \gamma_n^-(-\mu_n^-)
\end{aligned}
$$

Forming necessary conditions (3rd step) delivers

(A1.24)

$$
\begin{aligned}
\frac{\partial L}{\partial \mu_1^+} = & \, x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot p \cdot (\mu_1^+)^{p-1} \\
& -\lambda_1^+ - \gamma_1^+ = 0 \\
\frac{\partial L}{\partial \mu_1^-} = & \, x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot p \cdot (\mu_1^-)^{p-1} \\
& -\lambda_1^- - \gamma_1^- = 0 \\
& \vdots \\
\frac{\partial L}{\partial \mu_n^+} = & \, x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot (1-\tau) \cdot p \cdot (\mu_n^+)^{p-1} \\
& -\lambda_n^+ - \gamma_n^+ = 0 \\
\frac{\partial L}{\partial \mu_n^-} = & \, x \cdot \frac{1}{p} \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p\right]^{\frac{1-p}{p}} \cdot \tau \cdot p \cdot (\mu_n^-)^{p-1} \\
& -\lambda_n^- - \gamma_n^- = 0
\end{aligned}
$$

$$\frac{\partial L}{\partial \theta_1} = \lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial \theta_m} = \lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* = 0$$

$$\frac{\partial L}{\partial \beta_1} = \beta_1 - f(y, A_1) = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial \beta_m} = \beta_m - f(y, A_m) = 0$$

Collecting decision variables $\mu_1^+, \mu_1^-, \dots, \mu_n^+, \mu_n^-$ and $\theta_1, \dots, \theta_m$ in the Lagrange function (A1.23) produces (4th step)

(A1.25)

$$\begin{aligned}
L = x \cdot &\left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p \right]^{\frac{1}{p}} \\
&- \lambda_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \cdots - \lambda_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- \\
&- \gamma_1^+ \cdot \mu_1^+ - \gamma_1^- \cdot \mu_1^- - \cdots - \gamma_n^+ \cdot \mu_n^+ + \gamma_n^- \cdot \mu_n^- \\
&+ \theta_1 \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,1}^* - \theta_1 \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,1}^* + \frac{1}{2} \cdot \beta_1^2 - f(y, A_1) \cdot \beta_1 \\
&+ \cdots \\
&+ \theta_m \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,m}^* - \theta_m \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,m}^* + \frac{1}{2} \cdot \beta_m^2 - f(y, A_m) \cdot \beta_m \\
&- \sum_{i=1}^{n} \lambda_i^+ \cdot y_i^* + \sum_{i=1}^{n} \lambda_i^- \cdot y_i^*
\end{aligned}$$

To prepare for inserting the necessary conditions into the Lagrange function (5th step), an intermediate transformation is recommended. Each necessary condition with respect to μ_i^+ and μ_i^- is multiplied by its μ_i^+ and μ_i^- respectively. Then, all these multiplied necessary conditions are added, a procedure that results in

(A1.7)

$$
\begin{aligned}
& x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p \right]^{\frac{1}{p}} \\
& -\lambda_1^+ \cdot \mu_1^+ - \gamma_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \gamma_1^- \cdot \mu_1^- \\
& - \cdots \\
& -\lambda_n^+ \cdot \mu_n^+ - \gamma_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- - \gamma_n^- \cdot \mu_n^- = 0
\end{aligned}
$$

In a similar vein, each necessary condition with respect to θ_i is multiplied by its θ_i.

(A1.26)

$$
\begin{aligned}
\frac{\partial L}{\partial \theta_1} &= \lambda_1^+ A_{1,1}^* \cdot \theta_1 - \lambda_1^- A_{1,1}^* \cdot \theta_1 + \cdots + \lambda_n^+ A_{n,1}^* \cdot \theta_1 - \lambda_n^- A_{n,1}^* \cdot \theta_1 = 0 \\
&\vdots \\
\frac{\partial L}{\partial \theta_m} &= \lambda_1^+ A_{1,m}^* \cdot \theta_m - \lambda_1^- A_{1,m}^* \cdot \theta_m + \cdots + \lambda_n^+ A_{n,m}^* \cdot \theta_m - \lambda_n^- A_{n,m}^* \cdot \theta_m = 0
\end{aligned}
$$

Plugging the aggregated necessary conditions for μ (A1.7) and the multiplied necessary conditions for θ (A1.26) into Lagrange function (A1.25) yields (5th step)

$$
\begin{aligned}
L = {} & x \cdot \left[(1-\tau) \cdot \sum_{i=1}^{n} (\mu_i^+)^p + \tau \cdot \sum_{i=1}^{n} (\mu_i^-)^p \right]^{\frac{1}{p}} \\
& -\lambda_1^+ \cdot \mu_1^+ - \lambda_1^- \cdot \mu_1^- - \cdots - \lambda_n^+ \cdot \mu_n^+ - \lambda_n^- \cdot \mu_n^- \\
& \underbrace{-\gamma_1^+ \cdot \mu_1^+ - \gamma_1^- \cdot \mu_1^- - \cdots - \gamma_n^+ \cdot \mu_n^+ + \gamma_n^- \cdot \mu_n^-}_{=0 \; because \; of \; \text{(A1.7)}} \\
& \underbrace{+ \theta_1 \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,1}^* - \theta_1 \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,1}^*}_{=0 \; because \; of \; \text{(A1.26)}} \\
& + \cdots \\
& \underbrace{+ \theta_m \sum_{i=1}^{n} \lambda_i^+ \cdot A_{i,m}^* - \theta_m \sum_{i=1}^{n} \lambda_i^- \cdot A_{i,m}^*}_{=0 \; because \; of \; \text{(A1.26)}} \\
& - \sum_{i=1}^{n} \lambda_i^+ \cdot y_i^* + \sum_{i=1}^{n} \lambda_i^- \cdot y_i^*
\end{aligned}
$$

$$+\underbrace{\sum_{j=1}^{m}\frac{1}{2}\cdot\beta_j^2-f(y,A_j)\cdot\beta_j}_{=\sum_{j=1}^{m}\beta_j\cdot\underbrace{\left(\beta_j-f(y,A_j)\right)}_{=0\ because\ of\ (A1.24)}-\frac{1}{2}\cdot\beta_j^2}$$

Consequently,

$$L=-\sum_{i=1}^{n}\lambda_i^+\cdot y_i^*+\sum_{i=1}^{n}\lambda_i^-\cdot y_i^*-\frac{1}{2}\cdot\sum_{j=1}^{m}\beta_j^2$$

Appendix 1.3.3 Dual Program: First Version

Adapting the general dual program (A1.3) to the method of multiple environment (6th step) and translating the maximization problem into a minimization problem (7th step) leads to the first version of the dual program:

(A1.27)

$$\min_{\substack{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-,\\ \beta_1,\ldots,\beta_m,\\ \gamma_1^+,\gamma_1^-,\ldots,\gamma_n^+,\gamma_n^-}}\sum_{i=1}^{n}(\lambda_i^+-\lambda_i^-)\cdot y_i^*+\frac{1}{2}\cdot\sum_{j=1}^{m}\beta_j^2$$

s.t.

$$x\cdot\left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p+\tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}}\cdot(1-\tau)\cdot(\mu_1^+)^{p-1}-\lambda_1^+-\gamma_1^+=0$$

$$x\cdot\left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p+\tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}}\cdot\tau\cdot(\mu_1^-)^{p-1}-\lambda_1^--\gamma_1^-=0$$

$$\vdots$$

$$x\cdot\left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p+\tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}}\cdot(1-\tau)\cdot(\mu_n^+)^{p-1}-\lambda_n^+-\gamma_n^+=0$$

$$x\cdot\left[(1-\tau)\cdot\sum_{i=1}^{n}(\mu_i^+)^p+\tau\cdot\sum_{i=1}^{n}(\mu_i^-)^p\right]^{\frac{1-p}{p}}\cdot\tau\cdot(\mu_n^-)^{p-1}-\lambda_n^--\gamma_n^-=0$$

$$\lambda_1^+ A_{1,1}^* - \lambda_1^- A_{1,1}^* + \cdots + \lambda_n^+ A_{n,1}^* - \lambda_n^- A_{n,1}^* = 0$$
$$\vdots$$
$$\lambda_1^+ A_{1,m}^* - \lambda_1^- A_{1,m}^* + \cdots + \lambda_n^+ A_{n,m}^* - \lambda_n^- A_{n,m}^* = 0$$

$$\beta_1 - f(y, A_1) = 0$$
$$\vdots$$
$$\beta_m - f(y, A_m) = 0$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \dots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$
$$\gamma_1^+ \geq 0, \gamma_1^- \geq 0, \dots, \gamma_n^+ \geq 0, \gamma_n^- \geq 0$$

Appendix 1.3.4 Dual Program: Final Version (8th Step)

The necessary conditions of the dual program: first version (A1.27) are a function of the decision variables μ_i^+ and μ_i^- of the primal program (A1.22). For that reason, dual program (A1.27) has no ready economic interpretation and, thus, cannot serve as economic model selection criterion. Instead, μ_i^+ and μ_i^- must be removed from the constraints of dual program (A1.27). In this, connection, the same procedure as in Appendix 1.2.4 can be followed because the necessary conditions are identical. For that reason it is obtained

(4.3)

$$\min_{\substack{\lambda_1^+, \lambda_1^-, \dots, \lambda_n^+, \lambda_n^-, \\ \beta_1, \dots, \beta_m}} \sum_{i=1}^{n} (\lambda_i^+ - \lambda_i^-) \cdot y_i^* + \frac{1}{2} \cdot \sum_{j=1}^{m} \beta_j^2$$

s.t.

(4.4)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}} \leq x$$

$$\beta_1 - f(y, A_1) = 0$$
$$\vdots$$
$$\beta_m - f(y, A_m) = 0$$

$$\lambda_1^+ \geq 0, \lambda_1^- \geq 0, \dots, \lambda_n^+ \geq 0, \lambda_n^- \geq 0$$

β_j remains as artefact in the objective function.

Appendix 1.4 Lagrange Dual of Optimize-the-Price Approaches

Appendix 1.4.1 Primal Program

The primal program of the buyer's optimize-the-price-approach reads (see Chapter V, Formulas (5.3) to (5.5))

(5.3)

$$\min_{\lambda_1^+, \lambda_1^-, \dots, \lambda_n^+, \lambda_n^-} P_1 \cdot (\lambda_1^+ - \lambda_1^-) + \cdots + P_n \cdot (\lambda_n^+ - \lambda_n^-)$$

s.t.

(5.4)

$$A_{1,1}(\lambda_1^+ - \lambda_1^-) + \cdots + A_{n,1}(\lambda_n^+ - \lambda_n^-) \geq a_1$$
$$\vdots$$
$$A_{1,m}(\lambda_1^+ - \lambda_1^-) + \dots + A_{n,m}(\lambda_n^+ - \lambda_n^-) \geq a_m$$

(5.5)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$\lambda_1^+ \geq 0$$
$$\lambda_1^- \geq 0$$
$$\vdots$$
$$\lambda_n^+ \geq 0$$
$$\lambda_n^- \geq 0$$

The primal program (5.3) to (5.5) is then formulated in standard form (1st step)

(A1.28)

$$\min_{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-} P_1 \cdot (\lambda_1^+ - \lambda_1^-) + \cdots + P_n \cdot (\lambda_n^+ - \lambda_n^-)$$

s.t.

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$-A_{1,1} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,1} \cdot (\lambda_n^+ - \lambda_n^-) \leq -a_1$$

$$\vdots$$

$$-A_{1,m} \cdot (\lambda_1^+ - \lambda_1^-) - \ldots - A_{n,m} \cdot (\lambda_n^+ - \lambda_n^-) \leq -a_m$$

$$-\lambda_1^+ \leq 0$$
$$-\lambda_1^- \leq 0$$
$$\vdots$$
$$-\lambda_n^+ \leq 0$$
$$-\lambda_n^- \leq 0$$

Appendix 1.4.2 Preparing for Dualization

The Lagrange function of the primal problem in standard form (A1.28) reads (2nd step)

(A1.29)

$$\begin{aligned} L = {} & P_1 \cdot (\lambda_1^+ - \lambda_1^-) + \cdots + P_n \cdot (\lambda_n^+ - \lambda_n^-) \\ & + \mu \cdot \left(\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} - x \right) \\ & + \beta_1 \cdot \left(-A_{1,1} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,1} \cdot (\lambda_n^+ - \lambda_n^-) + a_1\right) \\ & + \cdots \\ & + \beta_m \cdot \left(-A_{1,m} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,m} \cdot (\lambda_n^+ - \lambda_n^-) + a_m\right) \\ & + \nu_1^+ \cdot (-\lambda_1^+) + \nu_1^- \cdot (-\lambda_1^-) + \cdots + \nu_n^+ \cdot (-\lambda_n^+) + \nu_n^- \cdot (-\lambda_n^-) \end{aligned}$$

Forming necessary conditions (3rd step) delivers

$$\frac{\partial L}{\partial \lambda_1^+} = P_1 + \mu \cdot \frac{p-1}{p} \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}-1}$$

$$\cdot (1-\tau)^{\frac{1}{1-p}} \cdot \frac{p}{p-1} \cdot (\lambda_1^+)^{\frac{p}{p-1}-1} - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m} - \nu_1^+ = 0$$

$$\frac{\partial L}{\partial \lambda_1^-} = -P_1 + \mu \cdot \frac{p-1}{p} \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}-1}$$

$$\cdot (\tau)^{\frac{1}{1-p}} \cdot \frac{p}{p-1} \cdot (\lambda_1^-)^{\frac{p}{p-1}-1} + \beta_1 \cdot A_{1,1} + \cdots + \beta_m \cdot A_{1,m} - \nu_1^- = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial \lambda_n^+} = P_n + \mu \cdot \frac{p-1}{p} \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}-1}$$

$$\cdot (1-\tau)^{\frac{1}{1-p}} \cdot \frac{p}{p-1} \cdot (\lambda_n^+)^{\frac{p}{p-1}-1} - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m} - \nu_n^+ = 0$$

$$\frac{\partial L}{\partial \lambda_n^-} = -P_n + \mu \cdot \frac{p-1}{p} \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{p-1}{p}-1}$$

$$\cdot (\tau)^{\frac{1}{1-p}} \cdot \frac{p}{p-1} \cdot (\lambda_n^-)^{\frac{p}{p-1}-1} + \beta_1 \cdot A_{n,1} + \cdots + \beta_m \cdot A_{n,m} - \nu_n^- = 0$$

i.e.,

(A1.30)

$$\frac{\partial L}{\partial \lambda_1^+} = P_1 + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{-1}{p}}$$

$$\cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^+)^{\frac{1}{p-1}} - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m} - \nu_1^+ = 0$$

$$\frac{\partial L}{\partial \lambda_1^-} = -P_1 + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{-1}{p}}$$

$$\cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^-)^{\frac{1}{p-1}} + \beta_1 \cdot A_{1,1} + \cdots + \beta_m \cdot A_{1,m} - \nu_1^- = 0$$

$$\vdots$$

$$\frac{\partial L}{\partial \lambda_n^+} = P_n + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{-1}{p}}$$

$$\cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{1}{p-1}} - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m} - \nu_n^+ = 0$$

$$\frac{\partial L}{\partial \lambda_n^-} = -P_n + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$

$$\cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{1}{p-1}} + \beta_1 \cdot A_{n,1} + \cdots + \beta_m \cdot A_{n,m} - \nu_n^- = 0$$

because $\frac{p-1}{p} - 1 = \frac{-1}{p}$ and $\frac{p}{p-1} - 1 = \frac{1}{p-1}$.

Collecting decision variables λ_i^+ and λ_i^- in the Lagrange function (A1.29) delivers (4th step)

(A1.31)

$$\begin{aligned} L = &\left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right)(\lambda_1^+ - \lambda_1^-) \\ &+ \cdots \\ &+ \left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right)(\lambda_n^+ - \lambda_n^-) \\ &- \nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\ &+ \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \\ &+ \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}} \\ &- \mu \cdot x \end{aligned}$$

To prepare for inserting the necessary conditions into the Lagrange function (5th step), an intermediate transformation is recommended. Each necessary condition with respect to λ_i^+ and λ_i^- is multiplied by its λ_i^+ and λ_i^- respectively. Then, all these multiplied necessary conditions are added, a procedure that results in (because $\frac{1}{p-1} + 1 = \frac{p}{p-1}$)

$$0 = P_1 \cdot \lambda_1^+ + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$

$$\cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^+)^{\frac{p}{p-1}} - \beta_1 \cdot A_{1,1} \cdot \lambda_1^+ - \cdots - \beta_m \cdot A_{1,m} \cdot \lambda_1^+ - \nu_1^+ \cdot \lambda_1^+$$

$$-P_1 \cdot \lambda_1^- + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$

$$\cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^-)^{\frac{p}{p-1}} + \beta_1 \cdot A_{1,1} \cdot \lambda_1^- + \cdots + \beta_m \cdot A_{1,m} \cdot \lambda_1^- - \nu_1^- \cdot \lambda_1^-$$

$$+\cdots$$

$$+P_n \cdot \lambda_n^+ + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$

$$\cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{p}{p-1}} - \beta_1 \cdot A_{n,1} \cdot \lambda_n^+ - \cdots - \beta_m \cdot A_{n,m} \cdot \lambda_n^+ - \nu_n^+ \cdot \lambda_n^+$$

$$-P_n \cdot \lambda_n^- + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$

$$\cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{p}{p-1}} + \beta_1 \cdot A_{n,1} \cdot \lambda_n^- + \cdots + \beta_m \cdot A_{n,m} \cdot \lambda_n^- - \nu_n^- \cdot \lambda_n^-$$

i.e.,

$$\begin{aligned}
0 = {} & \left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right)(\lambda_1^+ - \lambda_1^-) \\
& + \cdots \\
& + \left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right)(\lambda_n^+ - \lambda_n^-) \\
& - \nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\
& + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \\
& \cdot \begin{bmatrix} (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^-)^{\frac{p}{p-1}} \\ + \cdots \\ +(1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{p}{p-1}} \end{bmatrix}
\end{aligned}$$

Note that the last terms in brackets can be written as

$$\begin{aligned}
& (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^-)^{\frac{p}{p-1}} \\
& + \cdots \\
& +(1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{p}{p-1}} \\
& = (1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}
\end{aligned}$$

Therefore, the multiplied and aggregated necessary condition reads

(A1.32)

$$
\begin{aligned}
0 = & \left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right)(\lambda_1^+ - \lambda_1^-) \\
& + \cdots \\
& + \left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right)(\lambda_n^+ - \lambda_n^-) \\
& - \nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\
& + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}
\end{aligned}
$$

Plugging the aggregated necessary conditions for λ (A1.32) into Lagrange function (A1.31) yields (5th step)

$$
\begin{aligned}
L = & \left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right)(\lambda_1^+ - \lambda_1^-) \\
& + \cdots \\
& + \left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right)(\lambda_n^+ - \lambda_n^-) \\
& - \nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\
& \underbrace{+ \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}}_{=0\ because\ of\ (\text{A1.32})} \\
& + \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}} \\
& - \mu \cdot x
\end{aligned}
$$

Consequently, it is obtained

(A1.33)

$$
L = -\mu \cdot x + \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}}
$$

Appendix 1.4.3 Dual Program: First Version

Adapting the general dual program (A1.3) to the optimize-the-price-approach environment (6th step) and translating the maximization problem into a minimization problem (7th step) leads to

$$\max_{\mu,\beta_1,\ldots,\beta_m} -\mu \cdot x + \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}}$$

which is equivalent to

$$\min_{\mu,\beta_1,\ldots,\beta_m} \mu \cdot x - \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}}$$

and, hence,

(A1.34)

$$\min_{\mu,\beta_1,\ldots,\beta_m} \mu \cdot x - \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}}$$

s.t.

$$\begin{aligned}
&P_1 + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \\
&\quad \cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^+)^{\frac{1}{p-1}} - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m} - \nu_1^+ = 0 \\
&-P_1 + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \\
&\quad \cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^-)^{\frac{1}{p-1}} + \beta_1 \cdot A_{1,1} + \cdots + \beta_m \cdot A_{1,m} - \nu_1^- = 0 \\
&\quad \vdots \\
&P_n + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \\
&\quad \cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{1}{p-1}} - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m} - \nu_n^+ = 0
\end{aligned}$$

$$-P_n + \mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}} \right]^{\frac{-1}{p}}$$
$$\cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{1}{p-1}} + \beta_1 \cdot A_{n,1} + \cdots + \beta_m \cdot A_{n,m} - \nu_n^- = 0$$

$$\mu \geq 0$$
$$\beta_1 \geq 0, \dots, \beta_m \geq 0$$
$$\nu_1^+ \geq 0, \dots, \nu_n^+ \geq 0$$
$$\nu_1^- \geq 0, \dots, \nu_n^- \geq 0$$

Appendix 1.4.4 Dual Program: Final Version (8th step)

The necessary conditions of the dual program: first version (A1.34) are a function of the decision variables λ_i^+ and λ_i^- of the primal program (A1.28). For that reason, dual program (A1.34) has no ready interpretation as error minimization problem/cannot be related to the superordinate categories regression approaches and method of multiples. Instead, λ_i^+ and λ_i^- must be removed from the constraints of dual program (A1.34).

Appendix 1.4.4.1 Determination and thus Elimination of $\boldsymbol{\lambda_i^+}$ and $\boldsymbol{\lambda_i^-}$

Appendix 1.4.4.1.1 Equations that contain $\boldsymbol{\lambda_i^+}$ and $\boldsymbol{\lambda_i^-}$

The following equations contain λ_i^+ and λ_i^- and, hence, are candidates for the elimination of λ_i^+ and λ_i^-.

– Aggregated necessary condition for λ_i^+ and λ_i^-
  (A1.32)

$$\begin{aligned} 0 = &\left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right)(\lambda_1^+ - \lambda_1^-) \\ &+ \cdots \\ &+\left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right)(\lambda_n^+ - \lambda_n^-) \\ &-\nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \end{aligned}$$

$$+\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

– Necessary conditions with respect to λ_i^+ and λ_i^-

(A1.30)

$$P_1+\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$
$$\cdot(1-\tau)^{\frac{1}{1-p}}\cdot(\lambda_1^+)^{\frac{1}{p-1}}-\beta_1\cdot A_{1,1}-\cdots-\beta_m\cdot A_{1,m}-\nu_1^+=0$$
$$-P_1+\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$
$$\cdot(\tau)^{\frac{1}{1-p}}\cdot(\lambda_1^-)^{\frac{1}{p-1}}+\beta_1\cdot A_{1,1}+\cdots+\beta_m\cdot A_{1,m}-\nu_1^-=0$$
$$\vdots$$
$$P_n+\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$
$$\cdot(1-\tau)^{\frac{1}{1-p}}\cdot(\lambda_n^+)^{\frac{1}{p-1}}-\beta_1\cdot A_{n,1}-\cdots-\beta_m\cdot A_{n,m}-\nu_n^+=0$$
$$-P_n+\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}$$
$$\cdot(\tau)^{\frac{1}{1-p}}\cdot(\lambda_n^-)^{\frac{1}{p-1}}+\beta_1\cdot A_{n,1}+\cdots+\beta_m\cdot A_{n,m}-\nu_n^-=0$$

or

(A1.35)

$$\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}\cdot(1-\tau)^{\frac{1}{1-p}}\cdot(\lambda_1^+)^{\frac{1}{p-1}}$$
$$=\underbrace{\beta_1\cdot A_{1,1}+\cdots+\beta_m\cdot A_{1,m}-P_1}_{=\varepsilon_1^+\ according\ to\ (2.53)}+\nu_1^+$$

$$\mu\cdot\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}}\cdot(\tau)^{\frac{1}{1-p}}\cdot(\lambda_1^-)^{\frac{1}{p-1}}$$
$$=\underbrace{P_1-\beta_1\cdot A_{1,1}-\cdots-\beta_m\cdot A_{1,m}}_{=-\varepsilon_1^-\ according\ to\ (2.53)}+\nu_1^-$$

$\vdots$

$$\mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{1}{p-1}}$$
$$= \underbrace{\beta_1 \cdot A_{n,1} + \cdots + \beta_m \cdot A_{n,m} - P_n}_{=\varepsilon_n^+ \ according\ to\ (2.53)} + \nu_n^+$$

$$\mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{1}{p-1}}$$
$$= \underbrace{P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}}_{=-\varepsilon_n^- \ according\ to\ (2.53)} + \nu_n^-$$

Appendix 1.4.4.1.2 Solving the Equation System with respect to λ_i^+ and λ_i^-

The equation system (A1.35) and (A1.32) is rich enough to allow for a complete determination of λ_i^+ and λ_i^- and, hence, elimination of λ_i^+ and λ_i^- from the constraints of dual program (A1.34). Intuitively, it is proceeded as follows: From the necessary conditions (A1.35) $\lambda_1^-, \ldots, \lambda_n^-$ are expressed as functions of λ_1^+. These expressions are then inserted into (A1.32).

Restructuring (A1.32) leads to

$$\begin{aligned}
&\mu \cdot \left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \\
&+(-\varepsilon_1^+ - \nu_1^+) \cdot \lambda_1^+ + (-\varepsilon_1^- + \nu_1^-) \cdot (-\lambda_1^-) \\
&+\cdots \\
&+(-\varepsilon_n^+ - \nu_n^+) \cdot \lambda_n^+ + (-\varepsilon_n^- + \nu_n^-) \cdot (-\lambda_n^-) = 0
\end{aligned}$$

or

(A1.36)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$
$$= \frac{(\varepsilon_1^+ + \nu_1^+) \cdot \lambda_1^+ + (-\varepsilon_1^- + \nu_1^-) \cdot \lambda_1^- + \cdots + (\varepsilon_n^+ + \nu_n^+) \cdot \lambda_n^+ + (-\varepsilon_n^- + \nu_n^-) \cdot \lambda_n^-}{\mu}$$

since μ is by definition greater than zero. μ is the Langrange multiplier of the constraint on portfolio holdings. $\mu = 0$ implies that this constraint is not binding. From an economic perspective this means that no company is purchased and sold, i.e., all companies are correctly valued. In such a case, the profit from setting up an "accounting arbitrage" would be zero.—This is, however, a rather unrealistic case.

Reshuffling (A1.35) gains

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^+)^{\frac{1}{p-1}}$$
$$= \frac{\varepsilon_1^+ + \nu_1^+}{\mu}$$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_1^-)^{\frac{1}{p-1}}$$
$$= \frac{-\varepsilon_1^- + \nu_1^-}{\mu}$$

$\vdots$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \cdot (1-\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^+)^{\frac{1}{p-1}}$$
$$= \frac{\varepsilon_n^+ + \nu_n^+}{\mu}$$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{-1}{p}} \cdot (\tau)^{\frac{1}{1-p}} \cdot (\lambda_n^-)^{\frac{1}{p-1}}$$
$$= \frac{-\varepsilon_n^- + \nu_n^-}{\mu}$$

or rather (after taking the $1-p^{\text{th}}$ power)

(A1.37)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

$$= \left(\frac{\varepsilon_1^+ + \nu_1^+}{\mu}\right)^{1-p} \cdot (1-\tau)^{-1} \cdot \lambda_1^+$$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

$$= \left(\frac{-\varepsilon_1^- + \nu_1^-}{\mu}\right)^{1-p} \cdot (\tau)^{-1} \cdot \lambda_1^-$$

$$\vdots$$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

$$= \left(\frac{\varepsilon_n^+ + \nu_n^+}{\mu}\right)^{1-p} \cdot (1-\tau)^{-1} \cdot \lambda_n^+$$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}$$

$$= \left(\frac{-\varepsilon_n^- + \nu_n^-}{\mu}\right)^{1-p} \cdot (\tau)^{-1} \cdot \lambda_n^-$$

τ is located between 0 and 1 (see, e.g., Koenker (2005), p. 5); λ_i^+ and λ_i^- are nonnegative in the primal program (5.3) to (5.5). Moreover, $\varepsilon_i^+ + \nu_i^+$ $(-\varepsilon_i^- + \nu_i^-)$ will not be both identical to zero. If the estimated value exceeds the observed value ($\varepsilon_i^+ > 0$, undervaluation), the Langrage multiplier of the nonnegativity constraint on ε_i^+ (ν_i^+) is equal to zero because this constraint in not binding. In that case, no overvaluation will occur ($\varepsilon_i^- = 0$) and the Langagnre multiplier on nonnegative overvaluations ν_i^- will be greater than zero because the nonnegativity constraint on ε_i^- will be binding.

Since the left-hand sides of (A1.37) are identical, a relation between $\lambda_1^-, \dots, \lambda_n^-$ and λ_1^+ can be established:

(A1.38)

$$\lambda_1^- = \frac{(\varepsilon_1^+ + \nu_1^+)^{1-p}}{(-\varepsilon_1^- + \nu_1^-)^{1-p}} \cdot \frac{\tau}{1-\tau} \cdot \lambda_1^+$$

$$\vdots$$

$$\lambda_n^+ = \frac{(\varepsilon_1^+ + \nu_1^+)^{1-p}}{(\varepsilon_n^+ + \nu_n^+)^{1-p}} \cdot \lambda_1^+$$

$$\lambda_n^- = \frac{(\varepsilon_1^+ + \nu_1^+)^{1-p}}{(-\varepsilon_n^- + \nu_n^-)^{1-p}} \cdot \frac{\tau}{1-\tau} \cdot \lambda_1^+$$

and from (A1.36)

(A1.39)

$$\frac{(\varepsilon_1^+ + \nu_1^+) \cdot \lambda_1^+ + (-\varepsilon_1^- + \nu_1^-) \cdot \lambda_1^- + \cdots + (\varepsilon_n^+ + \nu_n^+) \cdot \lambda_n^+ + (-\varepsilon_n^- + \nu_n^-) \cdot \lambda_n^-}{\mu}$$
$$= (\varepsilon_1^+ + \nu_1^+)^{1-p} \cdot (1-\tau)^{-1} \cdot \lambda_1^+$$

Inserting (A1.38) into (A1.39) gains

$$\frac{\begin{gathered}(\varepsilon_1^+ + \nu_1^+) \cdot \lambda_1^+ + (-\varepsilon_1^- + \nu_1^-) \cdot \frac{(\varepsilon_1^+ + \nu_1^+)^{1-p}}{(-\varepsilon_1^- + \nu_1^-)^{1-p}} \cdot \frac{\tau}{1-\tau} \cdot \lambda_1^+ \\ + \cdots + (\varepsilon_n^+ + \nu_n^+) \cdot \frac{(\varepsilon_1^+ + \nu_1^+)^{1-p}}{(\varepsilon_n^+ + \nu_n^+)^{1-p}} \cdot \lambda_1^+ \\ +(-\varepsilon_n^- + \nu_n^-) \cdot \frac{(\varepsilon_1^+ + \nu_1^+)^{1-p}}{(-\varepsilon_n^- + \nu_n^-)^{1-p}} \cdot \frac{\tau}{1-\tau} \cdot \lambda_1^+\end{gathered}}{\mu}$$
$$= (\varepsilon_1^+ + \nu_1^+)^{1-p} \cdot (1-\tau)^{-1} \cdot \lambda_1^+$$

Division[15] by λ_1^+ and $(1-\tau)^{-1} \cdot (\varepsilon_1^+ + \nu_1^+)^{1-p}$ as well as multiplication by μ yields

$$\begin{aligned}&\frac{(\varepsilon_1^+ + \nu_1^+)}{(\varepsilon_1^+ + \nu_1^+)^{1-p}} \cdot (1-\tau) + \frac{(-\varepsilon_1^- + \nu_1^-)}{(-\varepsilon_1^- + \nu_1^-)^{1-p}} \cdot \tau \\ &+ \cdots \\ &+ \frac{(\varepsilon_n^+ + \nu_n^+)}{(\varepsilon_n^+ + \nu_n^+)^{1-p}} \cdot (1-\tau) + \frac{(-\varepsilon_n^- + \nu_n^-)}{(-\varepsilon_n^- + \nu_n^-)^{1-p}} \cdot \tau = \mu\end{aligned}$$

i.e.,

(A1.40)

$$\begin{aligned}&(\varepsilon_1^+ + \nu_1^+)^p \cdot (1-\tau) + (-\varepsilon_1^- + \nu_1^-)^p \cdot \tau \\ &+ \cdots \\ &+(\varepsilon_n^+ + \nu_n^+)^p \cdot (1-\tau) + (-\varepsilon_n^- + \nu_n^-)^p \cdot \tau = \mu\end{aligned}$$

[15] In the case that asset 1 is not bought, λ_1^+ will be zero. In that case, however, all equations will be expressed as a function of λ_1^- and the computations will follow the exactly same path.

(A1.40) contains the unobservable variables ν_i^+ and ν_i^-. However, according to (2.53) ε_i^+ is nonnegative and ε_i^- nonpositive, meaning that $-\varepsilon_i^-$ is nonnegative. Moreover, ν_i^+ and ν_i^- are nonnegative in the dual program (A1.34). Therefore, it holds

$$\varepsilon_i^+ + \nu_i^+ \geq \varepsilon_i^+ \text{ and } -\varepsilon_i^- + \nu_i^- \geq -\varepsilon_i^-$$

and, hence,

(A1.41)

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} \varepsilon_i^{+p} + \tau \cdot \sum_{i=i}^{n} (-\varepsilon_i^-)^p\right]^{\frac{1}{p}} \leq \mu$$

Eventually, to avoid writing ε_i^+ for the undervalued, but $-\varepsilon_i^-$ for the overvalued companies, absolute values might be more convenient—note both ε_i^+ and $-\varepsilon_i^-$ are positive:

(A1.42)

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} |\varepsilon_i^+|^p + \tau \cdot \sum_{i=i}^{n} |\varepsilon_i^-|^p\right]^{\frac{1}{p}} \leq \mu$$

(A1.42) is a constraint useful for integrating into the dual program (A1.34) because it no longer depends on λ_i^+ and λ_i^- as well as ν_i^+ and ν_i^-.

Appendix 1.4.5 Dual Program: Final Version

Substituting (A1.42) for the necessary conditions regarding λ_i^+ and λ_i^- in the first form of dual program (A1.34), it is finally gained

(A1.43)

$$\min_{\mu,\beta_1,\dots,\beta_m} \mu \cdot x - \sum_{j=1}^{m} \beta_j \cdot a_j$$

s.t.

$$\left[(1-\tau)\cdot\sum_{i=1}^{n}|\varepsilon_i^+|^p+\tau\cdot\sum_{i=i}^{n}|\varepsilon_i^-|^p\right]^{\frac{1}{p}}\leq\mu$$

$$\mu\geq 0$$
$$\beta_1\geq 0,\dots,\beta_m\geq 0$$

where

$$\varepsilon_i^+=\sum_{j=1}^{m}A_{i,j}^*\beta_j-y_i^*\geq 0$$
$$\varepsilon_i^-=\sum_{j=1}^{m}A_{i,j}^*\beta_j-y_i^*<0$$

(A1.43) is the desired final form of the dual program.

Appendix 1.4.6 Dual Program for the Special Case: p=1

The constraint regarding portfolio holdings in the buyer's primal program

(A1.28)

$$\left[(1-\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^+)^{\frac{p}{p-1}}+(\tau)^{\frac{1}{1-p}}\cdot\sum_{i=1}^{n}(\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}}\leq x$$

makes it immediately clear that the case $p=1$ cannot be integrated into (A1.28) but needs a treatment of its own. In fact, since not even the primal program can be adapted to this special case, the derivation of the dual program must start complete anew.

Appendix 1.4.6.1 Primal Program

The primal program in standard form (1st step) of this special case of the optimize-the-price-approach reads

(A1.44)

$$\min_{\lambda_1^+,\lambda_1^-,\dots,\lambda_n^+,\lambda_n^-} P_1 \cdot (\lambda_1^+ - \lambda_1^-) + \cdots + P_n \cdot (\lambda_n^+ - \lambda_n^-)$$

s.t.

<u>Constraints on individual portfolio holdings</u>

$$\lambda_1^+ \leq x_1^+$$
$$\vdots$$
$$\lambda_n^+ \leq x_n^+$$

and

$$\lambda_1^- \leq x_1^-$$
$$\vdots$$
$$\lambda_n^- \leq x_n^-$$

<u>Constraints on accounting figures</u>

$$-A_{1,1} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,1} \cdot (\lambda_n^+ - \lambda_n^-) \leq -a_1$$
$$\vdots$$
$$-A_{1,m} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,m} \cdot (\lambda_n^+ - \lambda_n^-) \leq -a_m$$

$$-\lambda_1^+ \leq 0$$
$$-\lambda_1^- \leq 0$$
$$\vdots$$
$$-\lambda_n^+ \leq 0$$
$$-\lambda_n^- \leq 0$$

Appendix 1.4.6.2 Preparing for Dualization

The Lagrange function of the primal problem (A1.44) reads (2nd step)

(A1.45)

$$\begin{aligned} L = {} & P_1 \cdot (\lambda_1^+ - \lambda_1^-) + \cdots + P_n \cdot (\lambda_n^+ - \lambda_n^-) \\ & + \mu_1^+ \cdot (\lambda_1^+ - x_1^+) + \mu_1^- \cdot (\lambda_1^- - x_1^-) \\ & + \cdots \end{aligned}$$

$$
\begin{aligned}
&+\mu_n^+ \cdot (\lambda_n^+ - x_n^+) + \mu_n^- \cdot (\lambda_n^- - x_n^-) \\
&+\beta_1 \cdot \left(-A_{1,1} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,1} \cdot (\lambda_n^+ - \lambda_n^-) + a_1\right) \\
&+\cdots \\
&+\beta_m \cdot \left(-A_{1,m} \cdot (\lambda_1^+ - \lambda_1^-) - \cdots - A_{n,m} \cdot (\lambda_n^+ - \lambda_n^-) + a_m\right) \\
&+\nu_1^+ \cdot (-\lambda_1^+) + \nu_1^- \cdot (-\lambda_1^-) + \cdots + \nu_n^+ \cdot (-\lambda_n^+) + \nu_n^- \cdot (-\lambda_n^-)
\end{aligned}
$$

Forming necessary conditions (3rd step) delivers

(A1.46)

$$
\begin{aligned}
&\frac{\partial L}{\partial \lambda_1^+} = P_1 + \mu_1^+ - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m} - \nu_1^+ = 0 \\
&\frac{\partial L}{\partial \lambda_1^-} = -P_1 + \mu_1^- + \beta_1 \cdot A_{1,1} + \cdots + \beta_m \cdot A_{1,m} - \nu_1^- = 0 \\
&\qquad \vdots \\
&\frac{\partial L}{\partial \lambda_n^+} = P_n + \mu_n^+ - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m} - \nu_n^+ = 0 \\
&\frac{\partial L}{\partial \lambda_n^-} = -P_n + \mu_n^- + \beta_1 \cdot A_{n,1} + \cdots + \beta_m \cdot A_{n,m} - \nu_n^- = 0
\end{aligned}
$$

Collecting decision variables λ_i^+ and λ_i^- in the Lagrange function (A1.45) delivers (4th step)

(A1.47)

$$
\begin{aligned}
L = &\left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right) \cdot (\lambda_1^+ - \lambda_1^-) \\
&+\cdots \\
&+\left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right) \cdot (\lambda_n^+ - \lambda_n^-) \\
&-\nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\
&+\mu_1^+ \cdot \lambda_1^+ + \mu_1^- \cdot \lambda_1^- + \cdots + \mu_n^+ \cdot \lambda_n^+ + \mu_n^- \cdot \lambda_n^- \\
&+\sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}} \\
&-\sum_{i=1}^{n} \mu_i^+ \cdot x_i^+ - \sum_{i=1}^{n} \mu_i^- \cdot x_i^-
\end{aligned}
$$

To prepare for inserting the necessary conditions into the Lagrange function (5th step), an intermediate transformation is recommended. Each necessary condition with respect to λ_i^+ and λ_i^- is multiplied by its λ_i^+ and λ_i^- respectively. Then, all these multiplied necessary conditions are added. In other words, the multiplied and aggregated necessary condition reads

(A1.48)

$$
\begin{aligned}
0 = &\left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right) \cdot (\lambda_1^+ - \lambda_1^-) \\
&+ \cdots \\
&+\left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right) \cdot (\lambda_n^+ - \lambda_n^-) \\
&-\nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\
&+\mu_1^+ \cdot \lambda_1^+ + \mu_1^- \cdot \lambda_1^- + \cdots + \mu_n^+ \cdot \lambda_n^+ + \mu_n^- \cdot \lambda_n^-
\end{aligned}
$$

Plugging the aggregated necessary conditions for λ (A1.48) into Lagrange function (A1.47) yields (5th step)

$$
\begin{aligned}
L = &\left(P_1 - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m}\right) \cdot (\lambda_1^+ - \lambda_1^-) \\
&\underbrace{\begin{aligned}&+ \cdots \\
&+\left(P_n - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m}\right) \cdot (\lambda_n^+ - \lambda_n^-) \\
&-\nu_1^+ \cdot \lambda_1^+ - \nu_1^- \cdot \lambda_1^- - \cdots - \nu_n^+ \cdot \lambda_n^+ - \nu_n^- \cdot \lambda_n^- \\
&+\mu_1^+ \cdot \lambda_1^+ + \mu_1^- \cdot \lambda_1^- + \cdots + \mu_n^+ \cdot \lambda_n^+ + \mu_n^- \cdot \lambda_n^-\end{aligned}}_{=0 \; because \; of \; (\text{A1.48})} \\
&+\sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}} \\
&-\sum_{i=1}^{n} \mu_i^+ \cdot x_i^+ - \sum_{i=1}^{n} \mu_i^- \cdot x_i^-
\end{aligned}
$$

Consequently, it is obtained

(A1.49)

$$
-\sum_{i=1}^{n} \mu_i^+ \cdot x_i^+ - \sum_{i=1}^{n} \mu_i^- \cdot x_i^- + \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}}
$$

Appendix 1.4.6.3 Dual Program: First Version

Adapting the general dual program (A1.3) to the optimize-the-price-approach environment (6th step) and translating the maximization problem into a minimization problem (7th step) leads to

(A1.50)

$$\min_{\substack{\mu_1^+,\mu_1^-,\dots,\mu_n^+,\mu_n^-,\\ \beta_1,\dots,\beta_m}} \sum_{i=1}^{n} \mu_i^+ \cdot x_i^+ + \sum_{i=1}^{n} \mu_i^- \cdot x_i^- - \sum_{j=1}^{m} \beta_j \cdot a_{\mathrm{j}}$$

s.t.

$$\begin{aligned} P_1 + \mu_1^+ - \beta_1 \cdot A_{1,1} - \cdots - \beta_m \cdot A_{1,m} - \nu_1^+ &= 0 \\ -P_1 + \mu_1^- + \beta_1 \cdot A_{1,1} + \cdots + \beta_m \cdot A_{1,m} - \nu_1^- &= 0 \\ &\vdots \\ P_n + \mu_n^+ - \beta_1 \cdot A_{n,1} - \cdots - \beta_m \cdot A_{n,m} - \nu_n^+ &= 0 \\ -P_n + \mu_n^- + \beta_1 \cdot A_{n,1} + \cdots + \beta_m \cdot A_{n,m} - \nu_n^- &= 0 \end{aligned}$$

$$\begin{aligned} \mu_1^+ \geq 0, \dots, \mu_n^+ \geq 0 \\ \mu_1^- \geq 0, \dots, \mu_n^- \geq 0 \\ \nu_1^+ \geq 0, \dots, \nu_n^+ \geq 0 \\ \nu_1^- \geq 0, \dots, \nu_n^- \geq 0 \\ \beta_1 \geq 0, \dots, \beta_m \geq 0 \end{aligned}$$

Appendix 1.4.6.4 Dual Program: Final Version (8th step)

The necessary conditions of the dual program: first version (A1.50) are a function of the Lagrange multipliers of portfolio holdings ν_i^+ and ν_i^- why the dual program (A1.50) has no ready interpretation as error minimization problem/cannot be related to the superordinate categories regression approaches and method of multiples. Therefore, ν_i^+ and ν_i^- must be removed from the constraints of dual program (A1.50).

Using the definition of ε_i^+ and ε_i^- from (2.53), the necessary conditions of (A1.50) can be re-formulated as

$$\begin{aligned} \varepsilon_1^+ + \nu_1^+ &= \mu_1^+ \\ -\varepsilon_1^- + \nu_1^- &= \mu_1^- \\ &\vdots \\ \varepsilon_n^+ + \nu_n^+ &= \mu_n^+ \\ -\varepsilon_n^- + \nu_n^- &= \mu_n^- \end{aligned}$$

According to (2.53) ε_i^+ is nonnegative and ε_i^- nonpositive, meaning that $-\varepsilon_i^-$ is nonnegative. Moreover, ν_i^+ and ν_i^- are nonnegative in the dual program (A1.34). Therefore, it holds

$$\varepsilon_i^+ + \nu_i^+ \geq \varepsilon_i^+ \text{ and } -\varepsilon_i^- + \nu_i^- \geq -\varepsilon_i^-$$

and, hence,

$$\begin{aligned} \varepsilon_1^+ &\leq \mu_1^+ \\ -\varepsilon_1^- &\leq \mu_1^- \\ &\vdots \\ \varepsilon_n^+ &\leq \mu_n^+ \\ -\varepsilon_n^- &\leq \mu_n^- \end{aligned}$$

This means, the dual program: first version (A1.50) simplifies to

(A1.51)

$$\min_{\substack{\mu_1^+,\mu_1^-,\dots,\mu_n^+,\mu_n^-,\\ \beta_1,\dots,\beta_m}} \sum_{i=1}^{n} \mu_i^+ \cdot x_i^+ + \sum_{i=1}^{n} \mu_i^- \cdot x_i^- - \sum_{j=1}^{m} \beta_j \cdot a_\mathrm{j}$$

s.t.

$$\begin{aligned} \varepsilon_1^+ &\leq \mu_1^+ \\ -\varepsilon_1^- &\leq \mu_1^- \\ &\vdots \\ \varepsilon_n^+ &\leq \mu_n^+ \\ -\varepsilon_n^- &\leq \mu_n^- \end{aligned}$$

$$\mu_1^+ \geq 0, \dots, \mu_n^+ \geq 0$$
$$\mu_n^- \geq 0, \dots, \mu_n^- \geq 0$$
$$\beta_1 \geq 0, \dots, \beta_m \geq 0$$

(A1.51) is the desired final form of the dual program.

Appendix 1.4.7 Dual Program for the Seller

The primal program of the seller's optimize-the-price-approach reads (see Chapter V, Formulas (5.6) to (5.8))

(5.6)

$$\max_{\lambda_1^+,\lambda_1^-,\ldots,\lambda_n^+,\lambda_n^-} P_1 \cdot (\lambda_1^+ - \lambda_1^-) + \cdots + P_n \cdot (\lambda_n^+ - \lambda_n^-)$$

s.t.

(5.7)

$$\begin{aligned} A_{1,1}(\lambda_1^+ - \lambda_1^-) + \cdots + A_{n,1}(\lambda_n^+ - \lambda_n^-) &\le a_1 \\ &\vdots \\ A_{1,m}(\lambda_1^+ - \lambda_1^-) + \ldots + A_{n,m}(\lambda_n^+ - \lambda_n^-) &\le a_m \end{aligned}$$

(5.8)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^+)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n} (\lambda_i^-)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \le x$$

$$\begin{aligned} \lambda_1^+ &\ge 0 \\ \lambda_1^- &\ge 0 \\ &\vdots \\ \lambda_n^+ &\ge 0 \\ \lambda_n^- &\ge 0 \end{aligned}$$

and in standard form

(A1.52)

$$\min_{N_{1,t}^+,N_{1,t}^-,\ldots,N_{n,t}^+,N_{n,t}^-} -P_{1,t} \cdot \left(N_{1,t}^+ - N_{1,t}^-\right) - \cdots - P_{n,t} \cdot \left(N_{n,t}^+ - N_{n,t}^-\right)$$

s.t.

(A1.53)

$$\begin{aligned} N_{1,t}^+ A_{1,1,t} - N_{1,t}^- A_{1,1,t} + \cdots + N_{n,t}^+ A_{n,1,t} - N_{n,t}^- A_{n,1,t} &\le a_1 \\ &\vdots \end{aligned}$$

$$N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} \leq a_m$$

(A1.54)

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$N_{1,t}^{+} \geq 0$$
$$N_{1,t}^{-} \geq 0$$
$$\vdots$$
$$N_{n,t}^{+} \geq 0$$
$$N_{n,t}^{-} \geq 0$$

If the primal program (A1.52) to (A1.54) is compared to the buyer's primal program (A1.28) where price $P_{i,t}$ is used instead of y_i^*, $N_{i,t}^{+}$ instead of λ_i^{+}, $N_{i,t}^{-}$ instead of λ_i^{-}, and $A_{1,j,t}$ instead of $A_{i,j}^{*}$

(A1.28)

$$\min_{N_{1,t}^{+},N_{1,t}^{-},\ldots,N_{n,t}^{+},N_{n,t}^{-}} P_{1,t} \cdot \left(N_{1,t}^{+} - N_{1,t}^{-}\right) + \cdots + P_{n,t} \cdot \left(N_{n,t}^{+} - N_{n,t}^{-}\right)$$

s.t.

$$N_{1,t}^{+}A_{1,1,t} - N_{1,t}^{-}A_{1,1,t} + \cdots + N_{n,t}^{+}A_{n,1,t} - N_{n,t}^{-}A_{n,1,t} \geq a_1$$
$$\vdots$$
$$N_{1,t}^{+}A_{1,m,t} - N_{1,t}^{-}A_{1,m,t} + \cdots + N_{n,t}^{+}A_{n,m,t} - N_{n,t}^{-}A_{n,m,t} \geq a_m$$

$$\left[(1-\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n}\left(N_{i,t}^{+}\right)^{\frac{p}{p-1}} + (\tau)^{\frac{1}{1-p}} \cdot \sum_{i=1}^{n}\left(N_{i,t}^{-}\right)^{\frac{p}{p-1}}\right]^{\frac{p-1}{p}} \leq x$$

$$-N_{1,t}^{+} \leq 0$$
$$-N_{1,t}^{-} \leq 0$$
$$\vdots$$
$$-N_{n,t}^{+} \leq 0$$
$$-N_{n,t}^{-} \leq 0$$

it becomes clear that a simple variable substitution achieves the transformation of the seller's into the buyer's primal program, namely use

$-P_{i,t}$ in the buyer's problem to obtain (A1.52) of the seller's problem

$-A_{i,j,t}$ in the buyer's problem to obtain (A1.53) of the seller's problem

$-a_j$ in the buyer's problem to obtain (A1.54) of the seller's problem

Applying these variable substitutions in (A1.43), the seller's dual program: final version reads

(A1.55)

$$\min_{\mu,\beta_1,\dots,\beta_m} \mu \cdot x + \sum_{j=1}^{m} \beta_j \cdot a_j$$

s.t.

$$\left[(1-\tau) \cdot \sum_{i=1}^{n} |\varepsilon_i^+|^p + \tau \cdot \sum_{i=i}^{n} |\varepsilon_i^-|^p\right]^{\frac{1}{p}} \leq \mu$$

$$\mu \geq 0$$
$$\beta_1 \geq 0, \dots, \beta_m \geq 0$$

where

$$\varepsilon_i^+ = -\sum_{j=1}^{m} A_{i,j}\beta_j + P_i \geq 0$$
$$\varepsilon_i^- = -\sum_{j=1}^{m} A_{i,j}\beta_j + P_i < 0$$

For the special case $p = 1$ it is obtained from (A1.51) by using the above variable substitutions

(A1.56)

$$\min_{\substack{\mu_1^+,\mu_1^-,\dots,\mu_n^+,\mu_n^-,\\ \beta_1,\dots,\beta_m}} \sum_{i=1}^{n} \mu_i^+ \cdot x_i^+ + \sum_{i=1}^{n} \mu_i^- \cdot x_i^- + \sum_{j=1}^{m} \beta_j \cdot a_\mathrm{j}$$

s.t.

$$\begin{aligned} \varepsilon_1^+ &\leq \mu_1^+ \\ -\varepsilon_1^- &\leq \mu_1^- \\ &\vdots \\ \varepsilon_n^+ &\leq \mu_n^+ \\ -\varepsilon_n^- &\leq \mu_n^- \end{aligned}$$

$$\begin{aligned} &\mu_1^+ \geq 0, \dots, \mu_n^+ \geq 0 \\ &\mu_n^- \geq 0, \dots, \mu_n^- \geq 0 \\ &\beta_1 \geq 0, \dots, \beta_m \geq 0 \end{aligned}$$

where

$$\varepsilon_i^+ = -\sum_{j=1}^{m} A_{i,j}\beta_j + P_i \geq 0$$

$$\varepsilon_i^- = -\sum_{j=1}^{m} A_{i,j}\beta_j + P_i < 0$$

Appendix 1.5 Strong Duality

Strong duality means that primal and dual program possess the same value of the objective function in the optimum, i.e., the duality gap is equal to zero (see, e.g., Boyd/Vandenberghe (2009), p. 226). From an economic perspective strong duality is desirable: Empirical asset pricing models can be connected with their dual program since the value of both objective functions coincide in the optimum.

According to Boyd/Vandenberghe (2009), p. 226 two steps are required to show strong duality: (i) the primal problem is convex meaning that usually but not always strong duality is given; (ii) so-called constraint qualifications must be given. One example of a constraint qualification is Slater's condition.

Appendix 1.5.1 Regression Approach as Primal Program

The constraint k ($k \in \{1, \dots, n\}$) of the primal program (A1.4) reads

$$\beta_0 + \sum_{j=1}^{m} A^*_{k,j}\beta_j - y^*_k - \mu^+_k \leq 0$$

Constraints in inequality form can always be expressed with the help of constraints in equality form and a slack variable (see Boyd/Vandenberghe (2009), pp. 131 f.):

$$\beta_0 + \sum_{j=1}^{m} A^*_{k,j}\beta_j - y^*_k - \mu^+_k + s^+_k = 0$$

where s^+_k denotes this slack variable.

Boyd/Vandenberghe (2009), p. 227 point out that the Slater condition reduces to the requirement of feasibility of the problem when the constraints are all linear equalities and the domain of the objective function is open. Since the minimum error—regression approaches minimize L_p-norms of errors—is less than infinity, the primal problem is indeed feasible.

Only the different sign of the value of the objective function of the primal and the dual problem must be taken into consideration: The dual program transforms the maximization problem into a minimization problem by multiplying by -1 (see 7th step, Appendix 1.1.2).

This is plausible since regression approaches consider valuation errors, but their dual programs consider "arbitrage profits" which means negative prices or positive payments. Thus, the valuation error can be interpreted as a profit.

Appendix 1.5.2 Optimize-the-Price Approach as Primal Program

The constraint h ($h \in \{1, \dots, m\}$) of the primal program (A1.28) reads

$$-A_{1,h} \cdot (\lambda_h^+ - \lambda_h^-) - \cdots - A_{n,h} \cdot (\lambda_h^+ - \lambda_h^-) \leq -a_h$$

Constraints in inequality form can always be expressed with the help of constraints in equality form and a slack variable (see Boyd/Vandenberghe (2009), pp. 131 f.):

$$-A_{1,h} \cdot (\lambda_h^+ - \lambda_h^-) - \cdots - A_{n,h} \cdot (\lambda_h^+ - \lambda_h^-) + s_h^+ = -a_h$$

where s_h^+ denotes this slack variable.

Boyd/Vandenberghe (2009), p. 227 point out that the Slater condition reduces to the requirement of feasibility of the problem when the constraints are all linear equalities and the domain of the objective function is open. Since the minimum price is greater than $-\infty$—otherwise an arbitrage profit would be possible—, the primal problem is indeed feasible.

Again the difference in sign of the primal/dual objective function is due to the 7th step (transformation of a maximization into a minimization problem).

Appendix 2 The Area Under the Cumulative Density Compared to the Area Under the Dirac Distribution Function

Chapter III, Section 2.1.3.3 introduces a criterion to characterize the similarity between different regression approaches: It compares the cumulative density function of differences (indicates, e.g., how close WLS and OLS regressions are) to the case where no differences exist because two approaches are identical. The latter case can be described with the help of the Dirac distribution. The less both functions diverge, i.e., the smaller the area between the cumulative distribution function of differences and the Dirac distribution is, the more similar the two approaches are.—The comparison of areas reads formally

(A2.1)

$$area\ (3.3) = \underbrace{\int_{-\infty}^{0} F(t)dt}_{region\ of\ negative\ differences} + \underbrace{\int_{0}^{\infty} \big(1 - F(t)\big)dt}_{region\ of\ positve\ differences}$$

If, e.g., WLS regression is compared to OLS regression, for some companies WLS regression might result in smaller estimated values than OLS regression—region of negative differences—and for some companies in greater values—region of positive differences. The Dirac distribution on the other hand captures the idea that there are no differences in value, i.e., all companies have zero valuation differences. That is the reason why in (A2.1) zero is the number that separates both regions.

The purpose of this appendix is to prove that this comparison of Areas (A2.1) can be expressed as expected value of the absolute value of the differences, in the case of Chapter III the Ratio (3.1) is used. To that end, Hajek (2015), p. 20 and Rao (2012) prove helpful.

The expected value of the absolute value of the Ratio (3.1) can be written as

(A2.2)

$$E\{|ratio_{C_i,j}|\} = E\{ratio^-_{C_i,j} + ratio^+_{C_i,j}\} = E\{ratio^-_{C_i,j}\} + E\{ratio^+_{C_i,j}\}$$

$ratio^-_{C_i,j} = \max\{-ratio_{C_i,j}; 0\}$ contains all negative values including zero, $ratio^+_{C_i,j} = \max\{ratio_{C_i,j}; 0\}$ contains all positive ratios.

To be able to differentiate between positive and negative ratios, $ratio^-$ and $ratio^+$ are expressed using indicator functions:

(A2.3)

$$ratio^-_{C_i,j} = \int_{-\infty}^{0} 1_{ratio_{C_i,j} \leq t}(t)dt$$

(A2.4)

$$ratio^+_{C_i,j} = \int_{0}^{\infty} 1_{ratio_{C_i,j} > t}(t)dt$$

The intuition behind this representation of $ratio^-_{C_i,j}$ and $ratio^+_{C_i,j}$ is the following:

$\boldsymbol{ratio^-_{C_i,j}}$:

$$1_{ratio_{C_i,j} \leq t}(t) = \begin{cases} 0 & \text{if } ratio_{C_i,j} > t \\ 1 & \text{if } ratio_{C_i,j} \leq t \end{cases}$$

Therefore, the integral reads

$$\begin{aligned} ratio^-_{C_i,j} &= \int_{-\infty}^{0} 1_{ratio_{C_i,j} \leq t}(t)dt \\ &= \int_{-\infty}^{ratio_{C_i,j}} 0\,dt + \int_{ratio_{C_i,j}}^{0} 1\,dt \\ &= 0 + t|^{0}_{ratio_{C_i,j}} \\ &= -ratio_{C_i,j} \end{aligned}$$

$\boldsymbol{ratio^+_{C_i,j}}$:

$$1_{ratio_{C_i,j} > t}(t) = \begin{cases} 0 & \text{if } ratio_{C_i,j} \leq t \\ 1 & \text{if } ratio_{C_i,j} > t \end{cases}$$

Therefore, the integral reads

$$
\begin{aligned}
ratio^{+}_{C_i,j} &= \int_0^{\infty} 1_{ratio_{C_i,j}>t}(t)dt \\
&= \int_0^{ratio_{C_i,j}} 1\,dt + \int_{ratio_{C_i,j}}^{\infty} 0\,dt \\
&= t\Big|_0^{ratio_{C_i,j}} + 0 \\
&= ratio_{C_i,j}
\end{aligned}
$$

Computing the expected value (A2.2) using the ratio specifications (A2.3) and (A2.4) means

$$E\{|ratio_{C_i,j}|\} = E\left\{\int_{-\infty}^{0} 1_{ratio_{C_i,j}\leq t}(t)dt\right\} + E\left\{\int_0^{\infty} 1_{ratio_{C_i,j}>t}(t)dt\right\}$$

Since the expected value of a sum (or of an integral) is equal to the sum (integral) of the expected values, one obtains

$$E\{|ratio_{C_i,j}|\} = \int_{-\infty}^{0} E\left\{1_{ratio_{C_i,j}\leq t}(t)\right\}dt + \int_0^{\infty} E\left\{1_{ratio_{C_i,j}>t}(t)\right\}dt$$

The expected value of the indicator function corresponds to a probability, i.e.,

$$E\{|ratio_{C_i,j}|\} = \int_{-\infty}^{0} Pr\big(ratio_{C_i,j} \leq t\big)dt + \int_0^{\infty} Pr\big(ratio_{C_i,j} > t\big)dt$$

where Pr denotes probability.

Since

$$Pr\big(ratio_{C_i,j} \leq t\big) = F(t) \text{ and } Pr\big(ratio_{C_i,j} > t\big) = 1 - F(t)$$

where $F(.)$ denotes the distribution function

it is finally obtained

(A2.5)

$$E\{|ratio_{C_i,j}|\} = \int_{-\infty}^{0} F(t)dt + \int_0^{\infty} \big(1 - F(t)\big)dt$$

Appendix 3 Definition of Variables

In the following table all the variables collected from Thompson Reuters Worldscope are listed together with their definition (see Thompson Reuters (2015)):

Variable	Abbreviation	Item No.	Definition
General information about the company			
Worldscope Permanent I.D.	WPID	WC06105	Represents a permanent identifier assigned to a company or security on the database.
Company Name	NAME	WC06001	Represents the legal name of the company as reported in the 10-K for U.S. companies and the annual report for non-U.S. companies.
Nation	NAT	WC06026	Represents the country in which the corporate office of a company is located.
ICB Code	ICB	WC07040	Represents an industry code within the Industrial Classification Benchmark (ICB) which was implemented as a result of a merger of the industrial classification of Dow Jones and FTSE. This benchmark allows for the comparison of companies through four hierarchical levels of industry classification. The ICB Code provided is the subsector code, the lowest level in the hierarchical structure.
Market Capitalization			
Market Capitalization – Fiscal Period End	P	WC08002	Market Price – Fiscal Period End * Common Shares Outstanding For companies with more than one type of common/ordinary share, market capitalization represents the total market value of the company.
Income Statement			
Net Sales Or Revenues	SA	WC01001	Represent gross sales and other operating revenue less discounts, returns and allowances.
Gross Income	GI	WC01100	Represents the difference between sales or revenues and cost of goods sold and depreciation/depletion, and amortization.
Earnings Before Interest, Taxes & Depreciation (EBITDA)	EBITDA	WC18198	Represent the earnings of a company before interest expense, income taxes and depreciation. It is calculated by taking the pre-tax income and adding back interest expense on debt and depreciation, depletion and amortization and subtracting interest capitalized.

Variable	Abbreviation	Item No.	Definition
Earnings Before Interest And Taxes (EBIT)	EBIT	WC18191	Represent the earnings of a company before interest expense and income taxes. It is calculated by taking the pre-tax income and adding back interest expense on debt and subtracting interest capitalized.
Earnings Before Taxes	EBT	WC01401	Pre-tax Income: Represents all income/loss before any federal, state or local taxes. Extraordinary items reported net of taxes are excluded.
Earnings	E	WC01751	Net Income Used To Calculate Earnings Per Share: Represents the net income the company uses to calculate its earnings per share. It is before extraordinary items.
		Balance Sheet	
Total Assets	TA	WC02999	Represent the sum of total current assets, long term receivables, investment in unconsolidated subsidiaries, other investments, net property plant and equipment and other assets.
Book Value Of Common Equity	B	WC03501	Represents common shareholders' investment in a company.
Invested Capital	IC (=TA-CE)	WC02999 WC02001	Total Assets - Cash & Short Term Investments
		Cash Flow Statement	
Operating Cash Flow	OCF	WC04860	Net Cash Flow – Operating Activities: Represent the net cash receipts and disbursements resulting from the operations of the company. It is the sum of Funds from Operations, Funds From/Used for Other Operating Activities and Extraordinary Items.
Ordinary Cash Dividends	D	WC18192	Dividends Provided For Or Paid – Common: Represents the total value of the common dividends declared for the year.

Table Appendix 3: Definition of Variables

Appendix 4 Overview of the Empirical Asset Pricing Literature

Empirical asset pricing is understood in this thesis as an approach that uses solely empirical analysis to derive asset prices. Therefore it should be distinguished from papers that (i) calibrate theoretical models to empirical data, most notably the literature concerned with the empirical testing of the CAPM or APT (Koijen/ Van Nieuwerburgh (2011) for calibration, Fama/MacBeth (1973) for CAPM, and Chen/Roll/Ross (1986) for APT testing); (ii) perform cross-sectional tests of return predictability or analyze the anomalies, i.e., deal with market efficiency (e.g., Fama (1991) and Fama (1998)).

We believe that there are two main strands of the empirical assets pricing literature that are, up to date, barely connected: value relevance studies in accounting and factor models/predictability of stock returns in finance.

One last introductory remark: Overview of the empirical asset pricing literature means extensive overview of the literature, but not complete analysis of all papers.—There are too many papers to be able to claim that we could completely capture the literature on empirical asset pricing: It took Kothari (2001) 127 pages to summarize the value relevance literature from the 1970s to the year 2000, Holthausen/Watts (2001) cover the same time period on 72 pages; the value relevance literature from 1990 to 2005 is contained in Mölls/Strauß (2007) on 42 pages. Finally, Harvey/Liu/Zhu (2016) provide on 63 pages an (highly) aggregated overview of factor models/predictability of stock returns: The literature discusses 316 predictors for asset returns.

Therefore, this appendix (i) puts a focus on new papers after 2010; (ii) regards statistical methods (e.g., ordinary, weighted, or generalized least squares) as equally important as factors; (iii) cites the "classical" papers in the field with publication years after 1990.

This means that papers will be listed in this appendix if they have identical factors, but different statistical methods or different factors, but identical statistical methods. Papers that analyze different markets (e.g., U.S. versus emerging markets) using identical factors and statistical methods will be ignored. Moreover, “non-classical” papers that employ factors that are a subset of factors examined in later papers are not contained either.—Only the later papers will be listed.

Value Relevance of Accounting Figures			
Paper	Dependent variable	Explanatory variables	Type of regression
		Publication year > 2010	
Baboukardos (2018)	Stock price	• Intercept • Earnings per share • Equity book value per share • Total score of the Environmental Performance Pillar of the Thomson Reuters ASSET4 database • Loss • Dummy indicating whether environmental information is contained in balance sheet • Industry dummy • Year dummy	Ordinary least squares regression
Barth/Li/McClure (2018)	Stock price	• Earnings per share • Equity book value per share • Research and development expense per share • Recognized intangible assets per share • Advertising expense per share • Cash, cash equivalents, and short-term investments per share • One-year revenue growth • Operating cash flow per share • Revenue per share • Special items per share • Other comprehensive income per share • Declared dividends to common shareholders per share • Capital expenditures per share • Cost of goods sold per share • Selling, general, and administrative expense per share • Total assets per share • Indicator variables for the Fama-French ten industry groups	Classification and Regression Trees (CART) estimation function; CART is a non-parametric estimation approach that does not require the researcher to specify the relation’s functional form.

Yan (2018a)	Stock price	• Intercept • Earnings per share • Book value per share • Difference between Chinese-GAAP and IFRS book value • Difference between Chinese-GAAP and IFRS earnings • Dummy variable capturing the pre convergence period in Chinese accounting standards	Ordinary least squares regression
Yan (2018b)	Stock price	• Intercept • Earnings per share • Book value per share • Difference between Japanese-GAAP and IFRS book value • Difference between Japanese-GAAP and IFRS earnings • R&D expenses • Goodwill • Difference between Japanese-GAAP and IFRS R&D expenses • Difference between Japanese-GAAP and IFRS goodwill	Ordinary least squares regression

Givoly/Hayn/Katz (2017)	Stock price	• Intercept • Earnings per share • Book value per share	Ordinary least squares regression
	Yield-to-maturity on the bond and the yield-to-maturity of the matched Treasury note	• Return on assets • $EBITDA_t/Liab_{t-1}$ • $Liab_t/Assets_t$ • Indicator variable of loss • Interaction terms between loss and the other three variables • Estimated probability of default	
	Stock returns	• Intercept • Firm j's income before extraordinary items in year t deflated by the total market value of equity at the end of year t – 1 • Unexpected net income defined as the analysts' earnings forecast error for year t.	
	Difference between bond return and the return on the duration-matched Treasury note	• Change of return on assets • Change of $EBITDA_t/Liab_{t-1}$ • Change of $Liab_t/Assets_t$ • Change of indicator variable of loss • Change of interaction terms between loss and the other three variables • Change of estimated probability of default	

Outa/Ozili/Eisenberg (2017)	Stock price	• Intercept • Net income per share • Book value per share • Extra accounting information known by the market but not yet integrated into the accounting system • Dummy for accounting convergence • Beta factor, size, debt to equity ratio	Random effects generalized least squares regression
Herath/Richardson/ Roubi/Tippett (2015)	Market value of equity	• Intercept • Earnings • Book value • Quadratic and cubic earnings • Quadratic and cubic book value • Cross products of earnings and book value (of linear and higher order)	Principal component analysis
Aleksanyan/Karim (2013)	Market value of common stock	• Intercept • Book value of common equity • Earnings available to common stockholders • Common dividends • Research and development costs • Advertising expenses	Ordinary least squares regression
Jiang/Stark (2013)	Market value of common stock	• Intercept • Book value of common equity • Net income • Research and development expenditures • Change in sales • Cash balance • Capital contributions • Capital contributions in the prior year • Change in long term debt between the current year and the prior year • Dummy for firm’s R&D intensity • Dummy for payment of dividends	Ordinary least squares regression

AbuGhazaleh/Al-Hares/Haddad (2012)	Stock price	• Intercept • Book value of common equity minus carrying value of goodwill • Pre-tax profit plus reported goodwill impairment loss • Carrying value of goodwill plus reported goodwill impairment loss • Reported goodwill impairment loss	Ordinary least squares regression
Barth/Landsman/Lang/ Williams (2012)	Stock price	• Intercept • Net earnings per share before extraordinary items • Book value per share	Ordinary least squares regression
	Stock return	• Intercept • Earnings per share divided by beginning of period stock price • Change in earnings per share divided by beginning of period stock price • Dummy for negative earnings per share	
Tsalavoutas/André/Evans (2012)	Stock price	• Intercept • Net earnings after taxes per share • Book value per share	Ordinary least squares regression, Weighted least squares regression
	Stock return	• Intercept • Net earnings after taxes per share • Change in net earnings after taxes per share	
Balachandran/Mohanram (2011)	Stock price	• Intercept • Earnings per share • Book value per share • Industry dummy • Dummy for negative earnings per share	Ordinary least squares regression
	Stock return	• Intercept • Earnings per share divided by beginning of period stock price • Change in earnings per share divided by beginning of period stock price • Industry dummy • Dummy for negative earnings per share	

Publication year 2000 to 2010			
Horton/Serafeim (2010)	Stock price	• Intercept • Earnings per share • Book value per share • Difference between IFRS and UK-GAAP book value • Difference between IFRS and UK-GAAP earnings • R&D expenses • Goodwill • Difference between Japanese-GAAP and IFRS R&D expenses • Difference between Japanese-GAAP and IFRS goodwill	Ordinary least squares regression
Easton/Monahan/Vasvari (2009)	Bond return	• Intercept • Earnings divided by market value • Bad versus good news • Investment versus speculative grade rating	Ordinary least squares regression
Kousenidis/Ladas/Negakis (2009)	Stock return	• Intercept • Earnings per share divided by beginning of period stock price • Change in earnings per share divided by beginning of period stock price	Ordinary least squares regression, Fixed/Random cross-sectional effects ordinary least squares regression
Hung/Subramanyam (2007)	Stock price	• Intercept • Book value of equity per share • Net income per share • Inverse Mills Ratio in the Heckman two-stage regression model	Ordinary least squares regression

Cazavan-Jeny/Jeanjean (2006)	Stock price	• Intercept • Book value of equity per share • Reported annual earnings • Adjusted book value per share, i.e., net of capitalized R&D • Adjusted earnings per share, i.e., before R&D expense and amortization of capitalized R&D • Annual amount of net capitalized R&D per share • Annual amount of expensed R&D per share • Year dummy • Industry dummy	Ordinary least squares regression
	Stock return	• Intercept • Reported annual earnings per share • Change in earnings per share in one year • Adjusted earnings per share, i.e., before R&D expense and amortization • Change in adjusted earnings per share in one year • Change in capitalized R&D per share • Change in annual amount of expensed R&D per share • Year dummy • Industry dummy	
Barth/Beaver/Hand/ Landsman (2005)	Market value of equity	• Intercept • Abnormal earnings, defined as earnings minus the normal return on equity book value • Book value • Other information • Annual change in receivables • Annual change in inventory • Annual change in payables • Depreciation and amortization expense • All explanatory variables follow vector autoregressive processes	Ordinary least squares regression

Bartov/Goldberg/Kim (2005)	Stock return	• Intercept • Income before extraordinary items divided by the market value of equity at the beginning of the year • Dummy for U.S. GAAP • Dummy for IAS	Ordinary least squares regression
Kallapur/Kwan (2004)	Market value of equity	• Intercept • Book value of equity • Net income • Brand assets	Ordinary least squares regression
Sami/Zhou (2004)	Stock price	• Intercept • Book value per share • Earnings per share • Other non-accounting information per share	Generalized least squares regression
Core/Guay/Van Buskirk (2003)	Market value of equity	• Intercept • Book value of equity • Net income • R&D expenditures • Advertising expenditures • Capital expenditures • 1 year change in constant-dollar sales	Ordinary least squares regression
	Deflated market value of equity	• Intercept divided by book value of equity • Constant • Net income divided by book value of equity • R&D expenditures divided by book value of equity • Advertising expenditures divided by book value of equity • Capital expenditures divided by book value of equity	

Easton/Sommers (2003)	Market value of equity	• Intercept • Book value of common equity • Net income	Ordinary least squares regression, Weighted least squares regression
	Deflated market value of equity	• Intercept divided by market capitalization • Book value of common equity divided by market capitalization • Net income divided by market capitalization	

Publication year 1990 to 2000			
Henning/Lewis/Shaw (2000)	Market value of equity	• Intercept • Earnings • Book value • Goodwill • Liabilities • Excess of the fair values of acquired net assets over their net book values • Difference between the fair value of assets recognized and the preacquisition market value of the target (GC) • Synergy gains • Excess of purchased goodwill over GC and synergy gains	Ordinary least squares regression
	Stock return	• Intercept • Earnings • Amortization of goodwill • Amortization of the difference between the fair value of assets recognized and the preacquisition market value of the target (GC) • Amortization of synergy gains • Amortization of excess of purchased goodwill over GC and synergy gains	
Brown/Lo/Lys (1999)	Stock price	• Intercept • Earnings per share • Book value per share	Ordinary least squares regression, Weighted least squares regression
	Deflated stock price: $\frac{P_{i,t}}{P_{i,t-1}}$	• Intercept • Time • Earnings per share normalized by $P_{i,t-1}$ • Book value per share normalized by $P_{i,t-1}$	
Francis/Schipper (1999)	Stock return	• Intercept • Earnings per share • Book value per share • Change in earnings per share	Ordinary least squares regression

Harris/Muller (1999)	Market value of equity	• Intercept • Earnings per share • Book value per share • Difference between US-GAAP and IAS book value • Difference between US-GAAP and IAS earnings • Number of shares outstanding	Ordinary least squares regression
	Stock return	• Intercept • Earnings per share divided by market capitalization at t-1 • Difference between US-GAAP and IAS earnings divided by market capitalization at t-1 • Change of the difference between US-GAAP and IAS earnings divided by market capitalization at t-1	
Simko (1999)	Stock price	• Intercept • Net book value of financial instruments • Net book value of nonfinancial instrument assets • Proxies for unobservable market expectations of future abnormal earnings • Lagged value of abnormal earnings • Net cumulative unrecognized holding gains on financial instrument assets, liabilities, and derivative contracts	Ordinary least squares regression
Kothari/Zimmermann (1995)	Stock price, Stock return, Price differences of stocks, Prices normalized by earnings	• Intercept • Earnings per share	Ordinary least squares regression
Easton/Harris (1991)	Stock price	• Intercept • Earnings per share or book value per share divided by beginning of period stock price • Change in earnings per share or book value per share divided by beginning of period stock price	Ordinary least squares regression

Factor Models/Predictability of Stock Returns			
Paper	Dependent variable	Explanatory variables	Type of regression
Publication year > 2010			
Gu/Kelly/Xiu (2018)	Risk premium of stock return	• Intercept • Dividend-price ratio • Earnings-price ratio • Book-to-market ratio • Net equity expansion • Treasury bill rate • Term spread • Default spread • Stock variance	• Minimizing mean squared predictions error • Models + (ordinary and weighted) least squares regression + generalized linear models with penalization + dimension reduction via principal component regression (PCR) and partial least squares (PLS) + regression trees (including boosted trees and random forests) + neural networks

Fama/French (2018)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio • Difference in returns on a portfolio of stocks with robust versus with weak operating profitability • Difference in returns on a portfolio of stocks with conservative and aggressive investments • Difference in returns on a portfolio with up movements versus low movements (momentum factor)	Ordinary least squares regression
Penman/Reggiani/ Richards/Tuna (2018)	Stock return	• Intercept • Earnings divided by stock price • Book value divided by stock price • Net debt divided by stock price	Ordinary least squares regression
Fama/French (2015)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio • Difference between the returns on a portfolio of low investment stocks and the returns on a portfolio of high investment stocks where investment is equal to annual change in total assets divided by 1-year-lagged total assets • Difference between the returns on a portfolio of high profitability (return on equity) and the return on a portfolio of low profitability stocks	Ordinary least squares regression

Hou/Xue/Zhang (2015)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio • Difference in returns on a portfolio of stocks with robust versus with weak operating profitability • Difference in returns on a portfolio of stocks with conservative and aggressive investments	Ordinary least squares regression
Lewellen (2015)	Stock return	• Intercept • Log market value of equity at the end of the prior month • Log book value of equity minus log market value of equity at the end of the prior month • Stock return from month −12 to month −2; Log growth in split-adjusted shares outstanding from month −36 to month −1 • Change in non-cash net working capital minus depreciation in the prior fiscal year • Income before extraordinary items divided by average total assets in the prior fiscal year • Log growth in total assets in the prior fiscal year • Dividends per share over the prior 12 months divided by price at the end of the prior month • Log stock return from month −36 to month −13 • Log growth in split-adjusted shares outstanding from month −12 to month −1 • Market beta estimated from weekly returns from month −36 to month −1 • Monthly standard deviation, estimated from daily returns from month −12 to month −1 • Average monthly turnover (shares traded/shares outstanding) from month −12 to month −1 • Short-term plus long-term debt divided by market value at the end of the prior month • Sales in the prior fiscal year divided by market value at the end of the prior month	Ordinary least squares regression

Defond/Zhang (2014)	Excess return of bond	• Intercept • Return on the 30-year Treasury Bond minus the return on the one-month Treasury Bill • Value-weighted return of all corporate bonds in the Datastream database with a maturity greater than 10 years minus the return on the 30-year Treasury Bond • Excess return on the market • Fama-French size factor, Small-Minus-Big Return • Fama-French market-to-book factor, High-Minus-Low Return	Ordinary least squares regression
Novy-Marx (2013)	Excess return of stock	• Intercept • Gross profitability scaled by total assets • Earnings scaled by book equity • Free cash flow scaled by book equity • Book-to-market • Size • Past performance	Ordinary least squares regression
Rapach/Zhou (2013)	Log return on the S&P 500 (including dividends) minus the log return on a risk-free bill	• Intercept • Log dividend-price ratio • Log dividend yield • Log earnings-price ratio • Log dividend-payout ratio • Stock variance • Book-to-market ratio • Net equity expansion • Treasury bill rate • Long-term government bond yield • Return on long-term government bonds • Long-term yield minus the Treasury bill rate • Default yield spread • Long-term corporate bond return minus the long-term government bond return • Inflation	Ordinary least squares regression

Allen/Singh/Powell (2011)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio	Quantile regression
Hu/Karolyi/Kho (2011)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks (local and foreign market portfolio) • Cash flow-to-price • Difference in returns on a portfolio with up movements versus low movements (momentum factor)	Ordinary least squares regression

Publication year 2000 to 2010			
Lewellen/Nagel/Shanken (2010)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio	Generalized least squares regression
Welch/Goyal (2008)	Log return on the S&P 500 (including dividends) minus the log return on a risk-free bill	• Intercept • Log dividend-price ratio • Log dividend yield • Log earnings-price ratio • Log dividend-payout ratio • Stock variance • Book-to-market ratio • Net equity expansion • Treasury bill rate • Long-term government bond yield • Return on long-term government bonds • Long-term yield minus the Treasury bill rate • Default yield spread • Long-term corporate bond return minus the long-term government bond return • Inflation • Investment-to-capital ratio	Ordinary least squares regression
Francis/LaFond/Olsson/ Schipper (2005)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio • Accounting quality	Ordinary least squares regression

Pástor/Stambaugh (2003)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio • Liquidity factor	Ordinary least squares regression

Publication year 1990 to 2000			
Carhart (1997)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio • Difference in returns on a portfolio with high versus low prior year return	Ordinary least squares regression
Fama/French (1993)	Excess return of stock	• Intercept • Excess return on the market portfolio of stocks • Difference in returns on portfolio of small versus big stocks • Difference in returns on a portfolio with high book-to-market versus low book-to-market ratio	Ordinary least squares regression
	Excess return of bond	• Intercept • Monthly percent long-term government bond return minus one-month Treasury bill rate • Return on a proxy for the market portfolio of corporate bonds minus percent long-term government bond return	

Appendix 5 Empirical Results

Appendix 5.1 Figures in Connection with Chapter III

Appendix 5.1.1 Factors when Controlled for Regressions Section 4.2.2.2

To measure the interplay between factors and statistical methods, Figure 3.2 is analyzed for each regression separately. Recall in this connection that M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept).

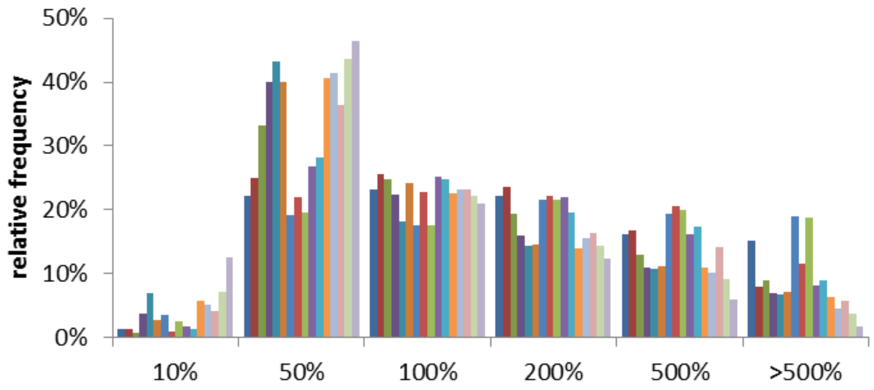


Figure Appendix 5.1.1a: |Ratio (3.1)| for method WLS with reference OLS

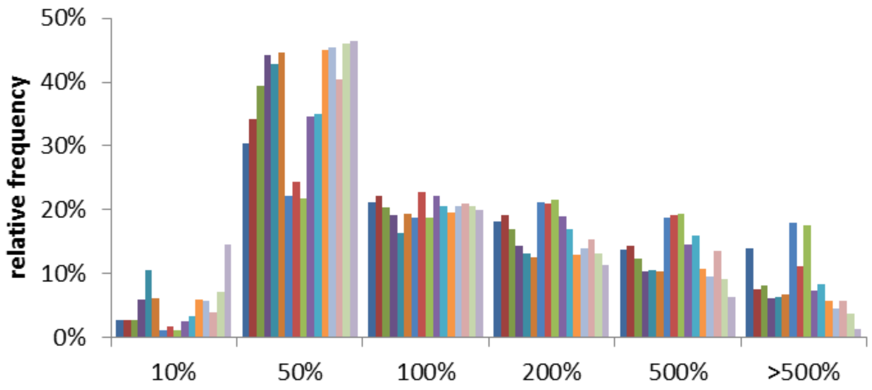


Figure Appendix 5.1.1b: |Ratio (3.1)| for method Quantile (0.25) with reference OLS

M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 M13 M14 M15 M16

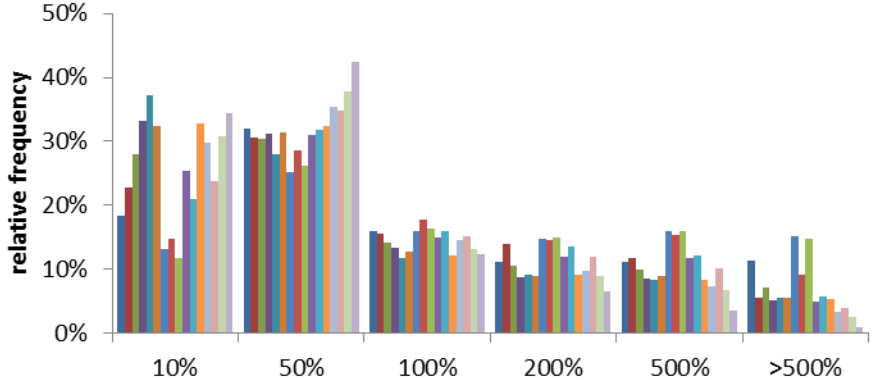


Figure Appendix 5.1.1c: |Ratio (3.1)| for method Quantile (0.50) with reference OLS

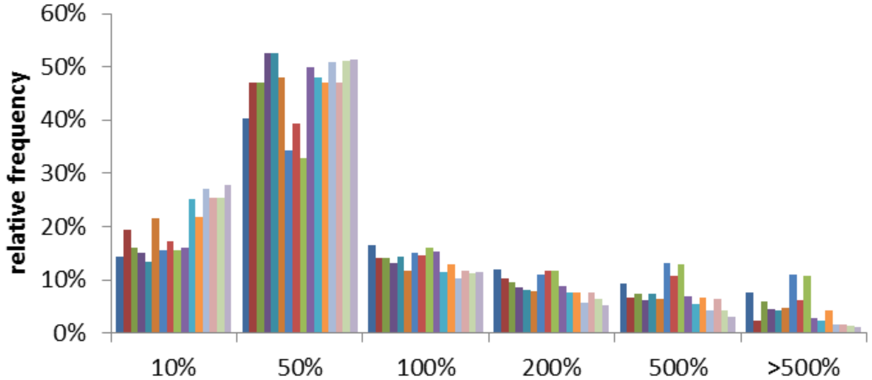


Figure Appendix 5.1.1d: |Ratio (3.1)| for method Quantile (0.75) with reference OLS

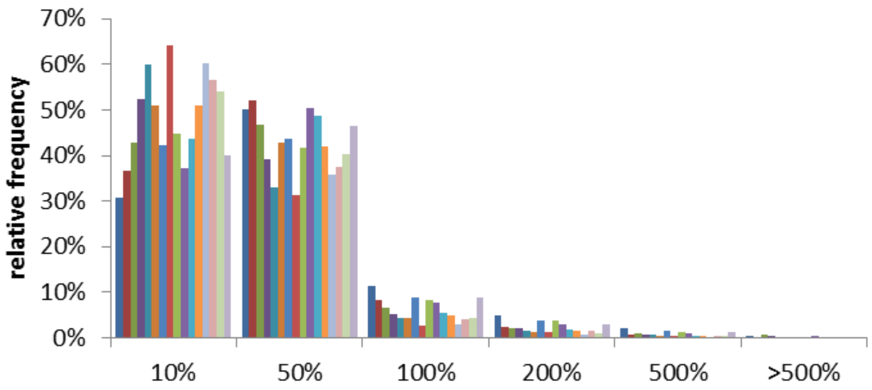


Figure Appendix 5.1.1e: |Ratio (3.1)| for method Quantile (0.25) with reference WLS

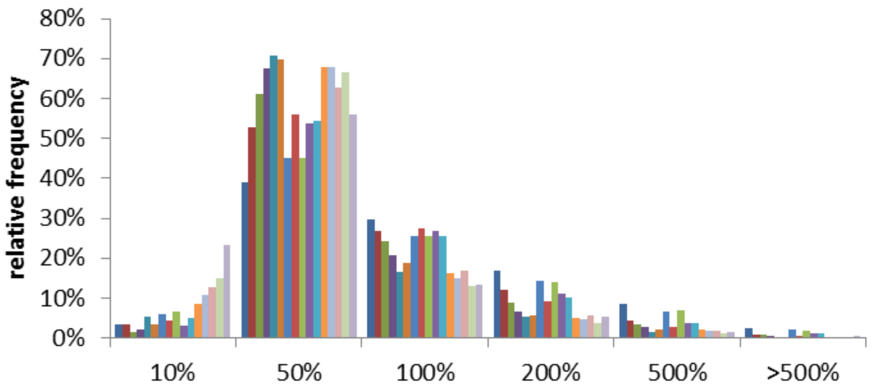


Figure Appendix 5.1.1f: |Ratio (3.1)| for method Quantile (0.50) with reference WLS

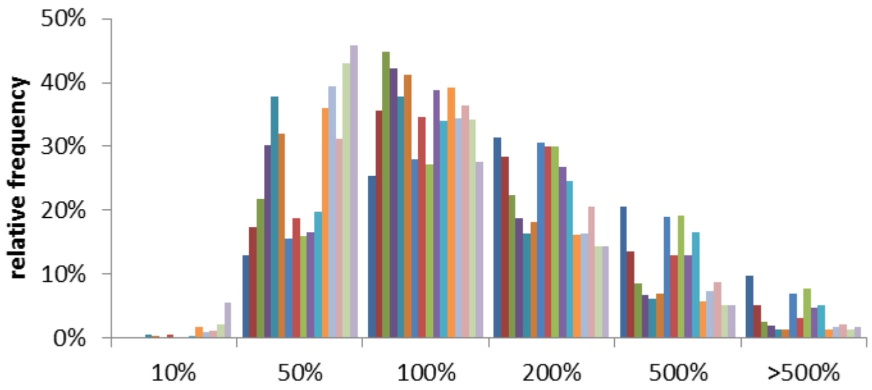


Figure Appendix 5.1.1g: |Ratio (3.1)| for method Quantile (0.75) with reference WLS

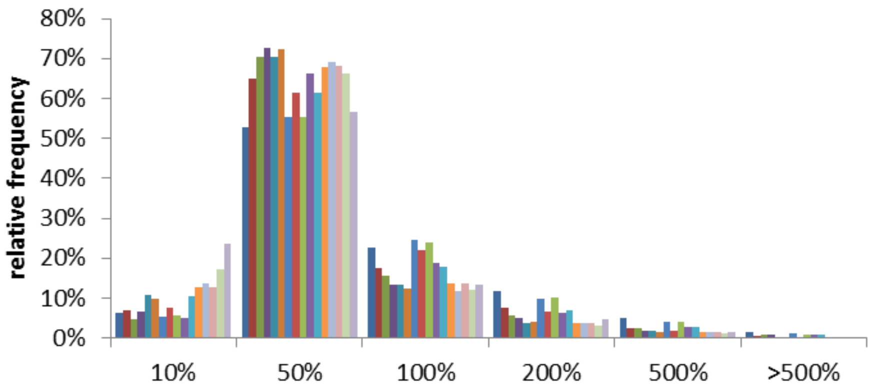


Figure Appendix 5.1.1h: |Ratio (3.1)| for method Quantile (0.50) with reference Quantile (0.25)

Figure Appendix 5.1.1i: |Ratio (3.1)| for method Quantile (0.75) with reference Quantile (0.25)

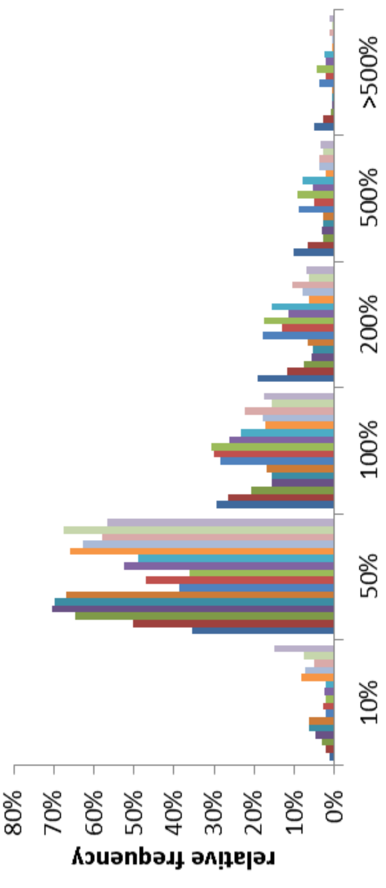
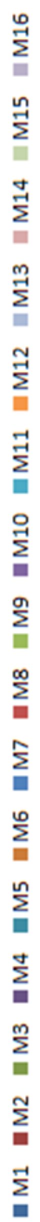


Figure Appendix 5.1.1j: |Ratio (3.1)| for method Quantile (0.75) with reference Quantile (0.50)



Appendix 5.1.2 Regressions when Controlled for Factors (Section 4.2.3.2)

To measure the interplay between factors and regressions, Figure 3.6 is analyzed for each factor separately. Recall in this connection that M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept).

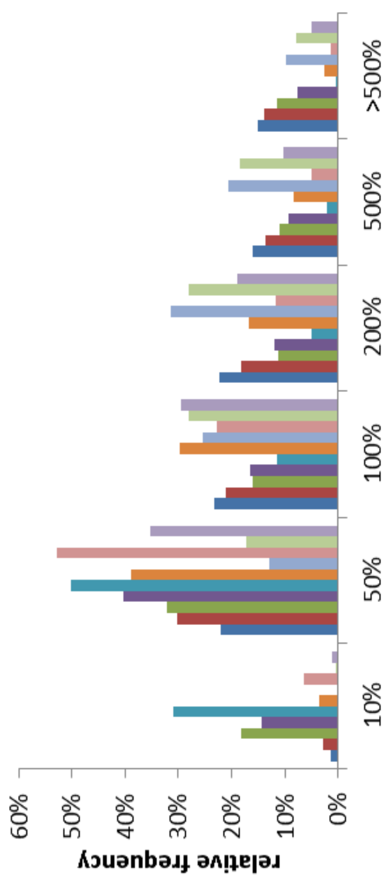


Figure Appendix 5.1.2a: |Ratio (3.1)| for model M1

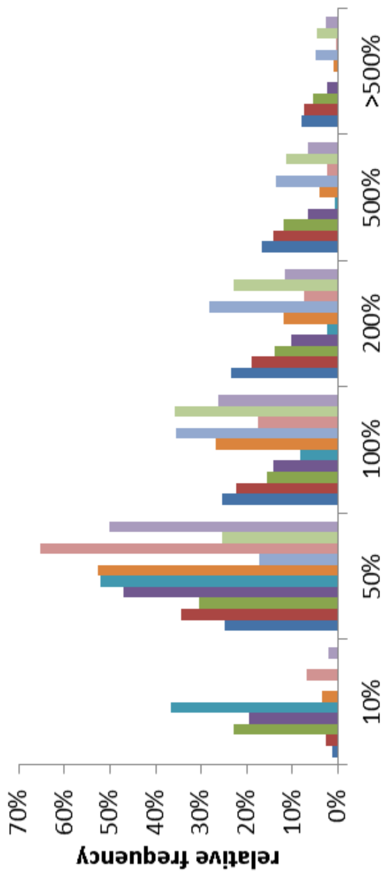


Figure Appendix 5.1.2b: |Ratio (3.1)| for model M2

OLS_WLS
OLS_Quantile(0.25)
OLS_Quantile(0.50)
OLS_Quantile(0.75)
WLS_Quantile(0.25)
WLS_Quantile(0.50)
WLS_Quantile(0.75)
Quantile(0.25)_Quantile(0.50)
Quantile(0.25)_Quantile(0.75)
Quantile(0.50)_Quantile(0.75)

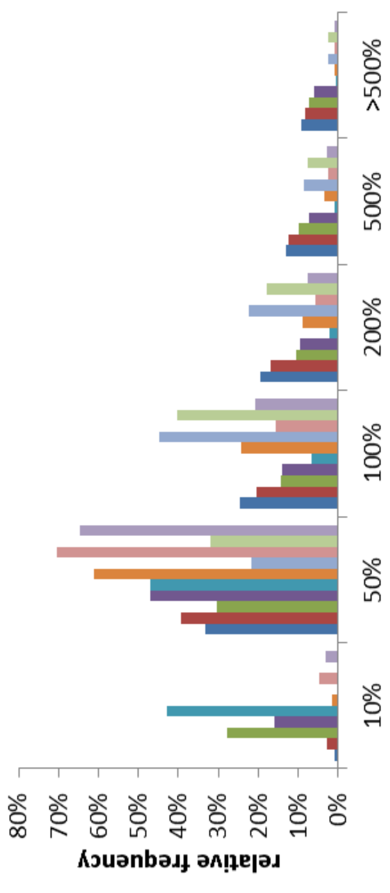


Figure Appendix 5.1.2c: |Ratio (3.1)| for model M3

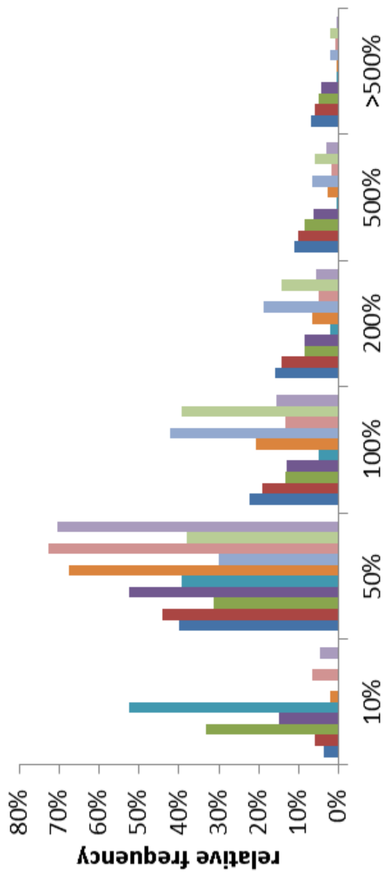


Figure Appendix 5.1.2d: |Ratio (3.1)| for model M4

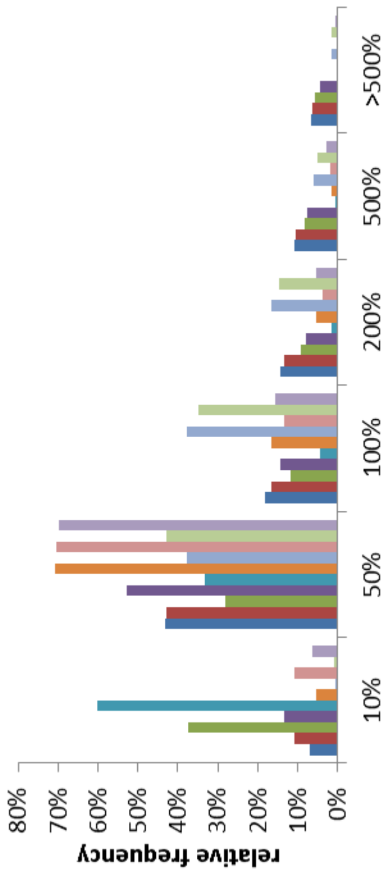


Figure Appendix 5.1.2e: |Ratio (3.1)| for model M5

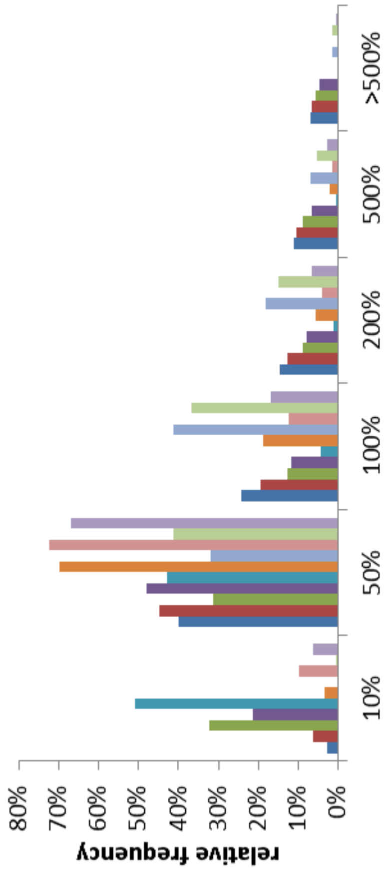


Figure Appendix 5.1.2f: |Ratio (3.1)| for model M6

OLS_WLS ■ OLS_Quantile(0.25) ■ OLS_Quantile(0.50) ■ OLS_Quantile(0.75) ■ WLS_Quantile(0.25)
■ WLS_Quantile(0.50) ■ WLS_Quantile(0.75) ■ Quantile(0.25)_Quantile(0.50) ■ Quantile(0.25)_Quantile(0.75) ■ Quantile(0.50)_Quantile(0.75)

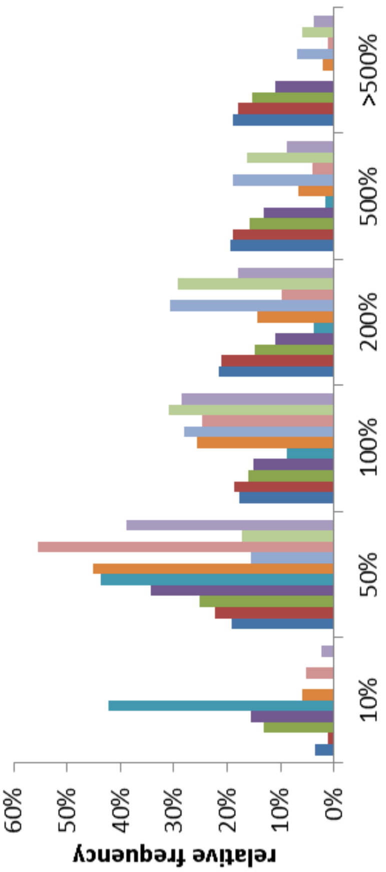


Figure Appendix 5.1.2g: |Ratio (3.1)| for model M7

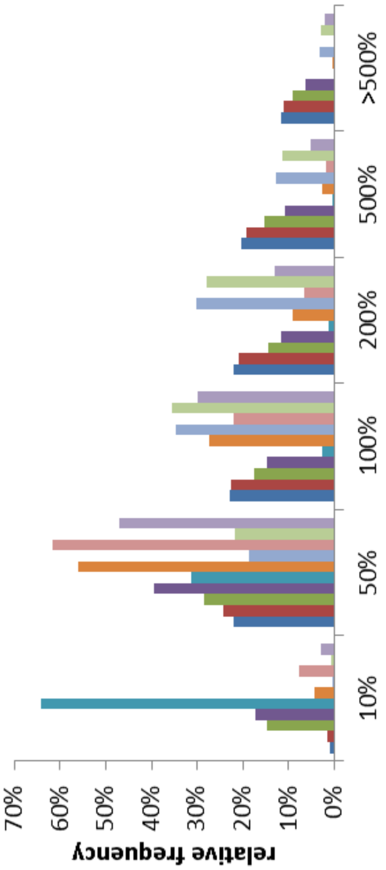


Figure Appendix 5.1.2h: |Ratio (3.1)| for model M8

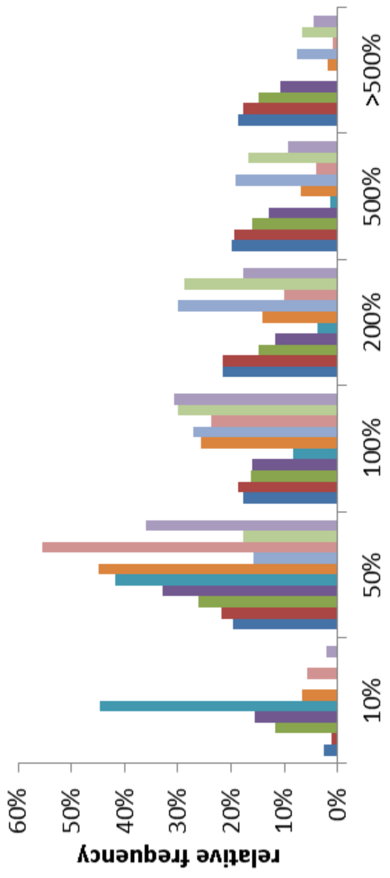


Figure Appendix 5.1.2i: |Ratio (3.1)| for model M9

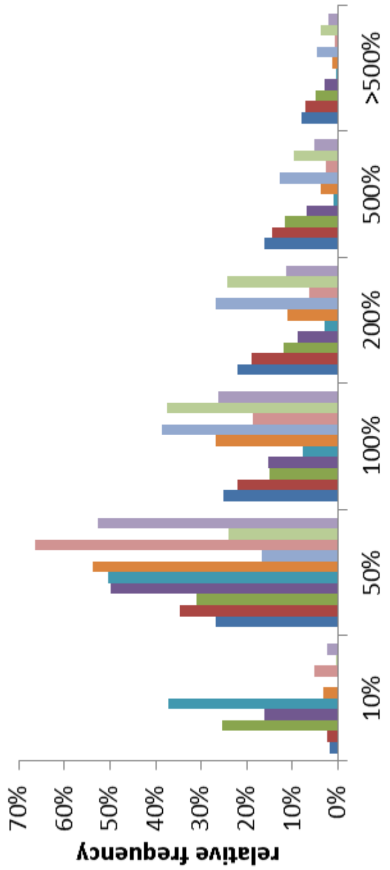


Figure Appendix 5.1.2j: |Ratio (3.1)| for model M10

■ OLS_WLS ■ OLS_Quantile(0.25) ■ OLS_Quantile(0.50) ■ OLS_Quantile(0.75) ■ WLS_Quantile(0.25)
■ WLS_Quantile(0.50) ■ WLS_Quantile(0.75) ■ Quantile(0.25)_Quantile(0.50) ■ Quantile(0.25)_Quantile(0.75) ■ Quantile(0.50)_Quantile(0.75)

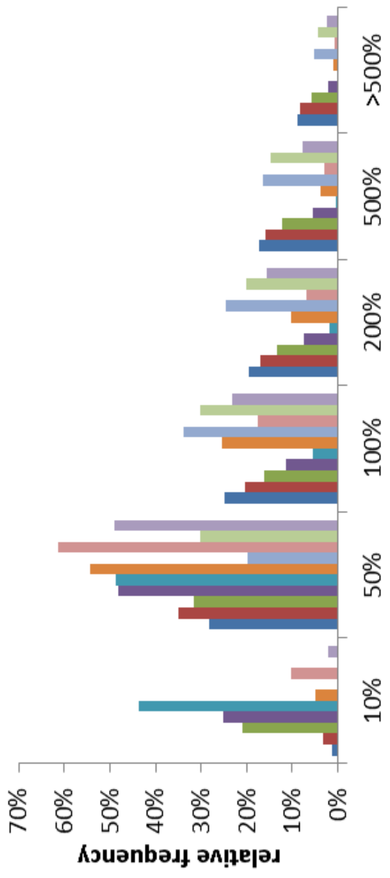


Figure Appendix 5.1.2k: |Ratio (3.1)| for model M11

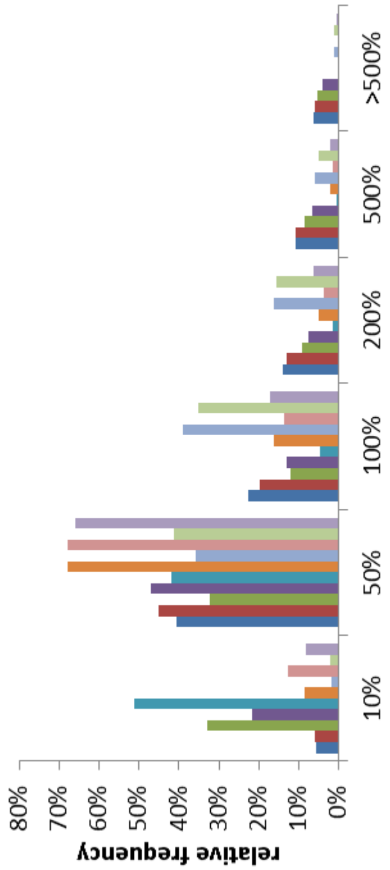


Figure Appendix 5.1.2l: |Ratio (3.1)| for model M12

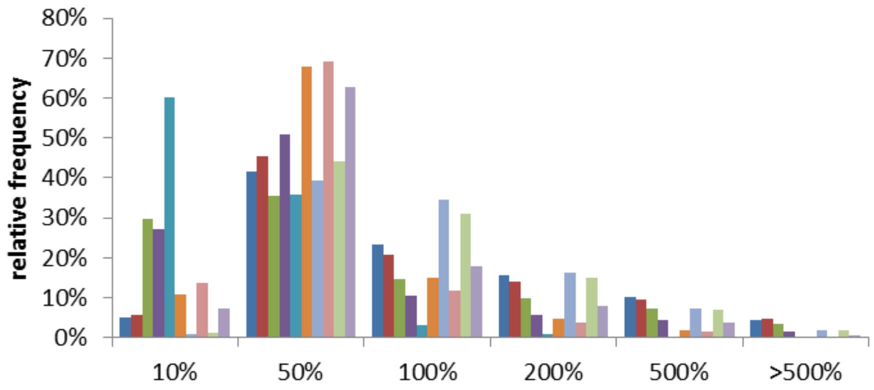


Figure Appendix 5.1.2m: |Ratio (3.1)| for model M13

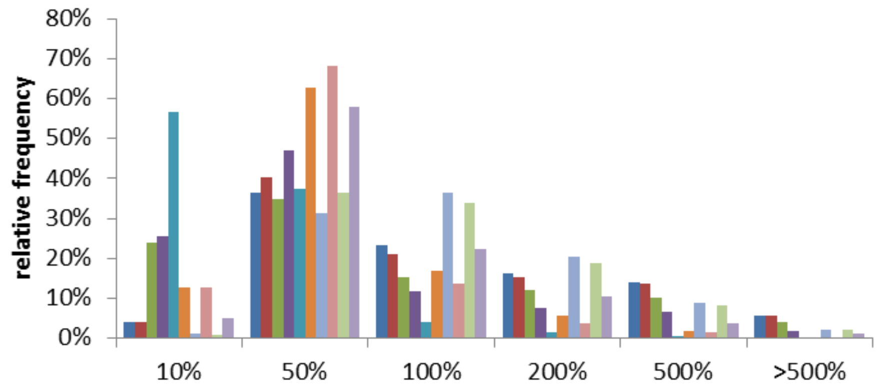
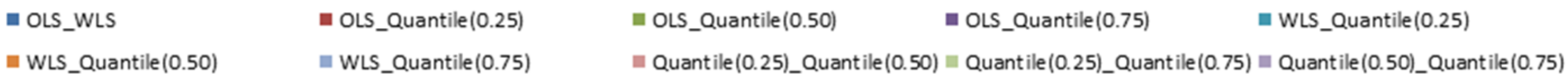


Figure Appendix 5.1.2n: |Ratio (3.1)| for model M14



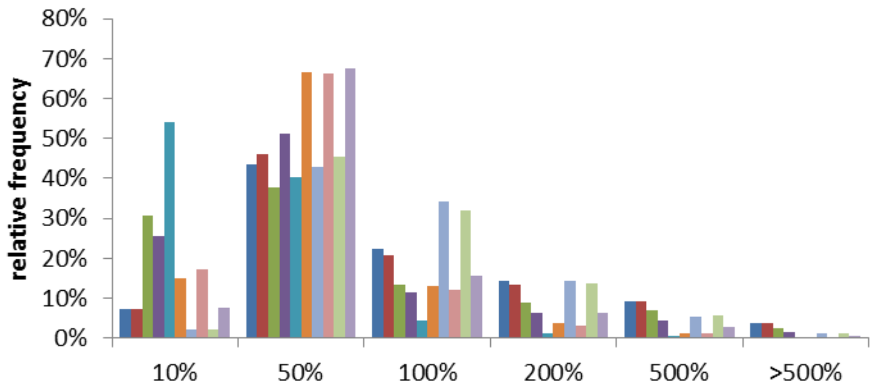


Figure Appendix 5.1.2o: |Ratio (3.1)| for model M15

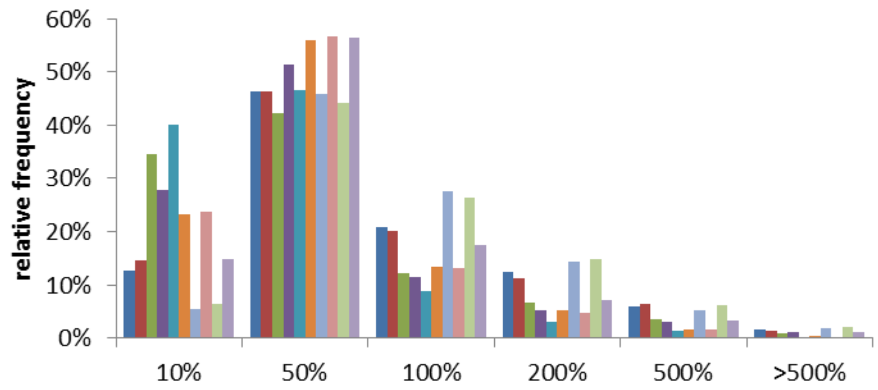


Figure Appendix 5.1.2p: |Ratio (3.1)| for model M16

Appendix 5.1.3 Statistical Methods that Generate High or Low Prices (Section 4.2.3.3)

To measure whether some statistical methods result in higher or lower prices than other statistical methods, Ratio (3.1) is broken down by statistical methods for each factor. Formally, the figures from Appendix 5.1 are split into positive and negative Ratio (3.1) using the following classes:

$$0\% < \text{Ratio (3.1)} \leq 10\% \quad \text{and} \quad -10\% < \text{Ratio (3.1)} \leq 0\%$$

$$10\% < \text{Ratio (3.1)} \leq 50\% \quad \text{and} \quad -50\% < \text{Ratio (3.1)} \leq -10\%$$

$$50\% < \text{Ratio (3.1)} \leq 100\% \quad \text{and} \quad -100\% < \text{Ratio (3.1)} \leq -50\%$$

$$100\% < \text{Ratio (3.1)} \leq 200\% \quad \text{and} \quad -200\% < \text{Ratio (3.1)} \leq -100\%$$

$$200\% < \text{Ratio (3.1)} \leq 500\% \quad \text{and} \quad -500\% < \text{Ratio (3.1)} \leq -200\%$$

$$\text{Ratio (3.1)} > 500\% \quad \text{and} \quad \text{Ratio (3.1)} \leq -500\%$$

In other words, classes are defined inclusive of the upper bound and exclusive of the lower bound.

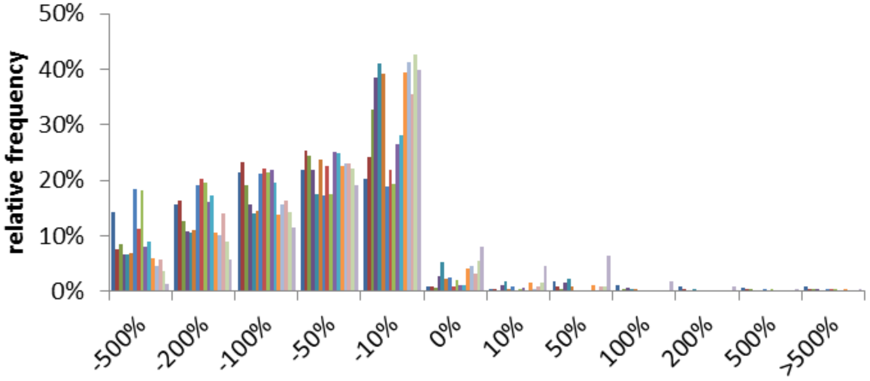


Figure Appendix 5.1.3a: Ratio (3.1) for method WLS with reference OLS

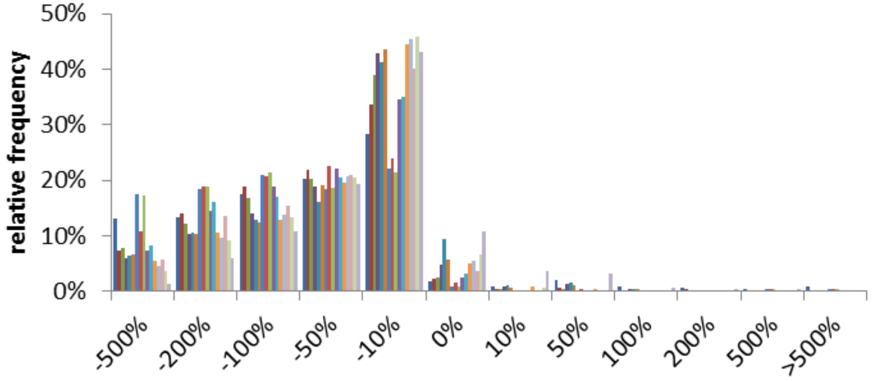


Figure Appendix 5.1.3b: Ratio (3.1) for method Quantile (0.25) with reference OLS

M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 M13 M14 M15 M16

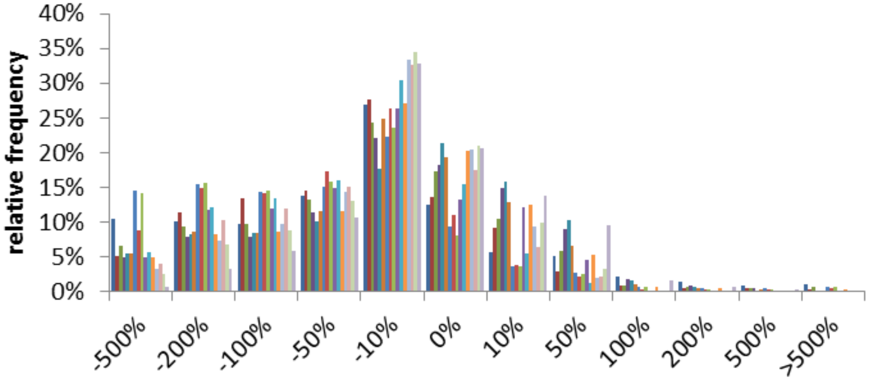


Figure Appendix 5.1.3c: Ratio (3.1) for method Quantile (0.50) with reference OLS

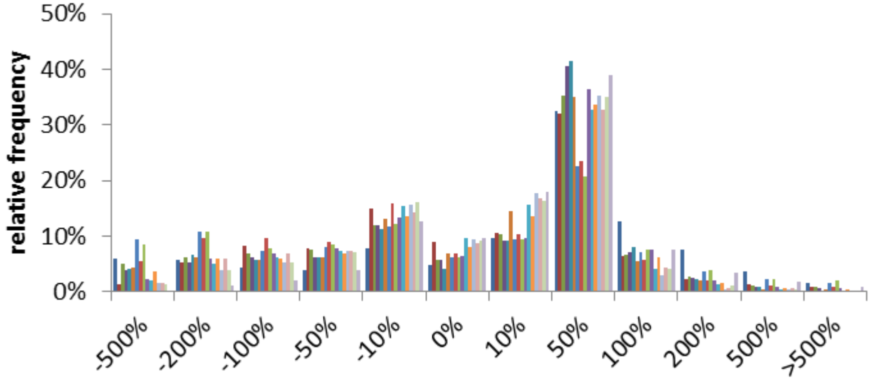


Figure Appendix 5.1.3d: Ratio (3.1) for method Quantile (0.75) with reference OLS

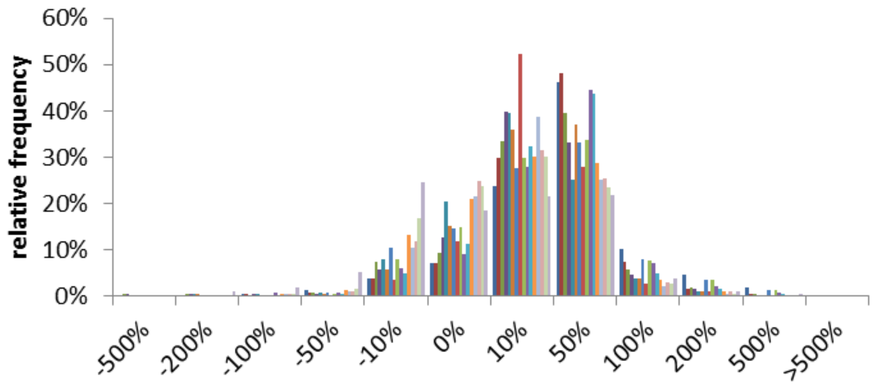


Figure Appendix 5.1.3e: Ratio (3.1) for method Quantile (0.25) with reference WLS

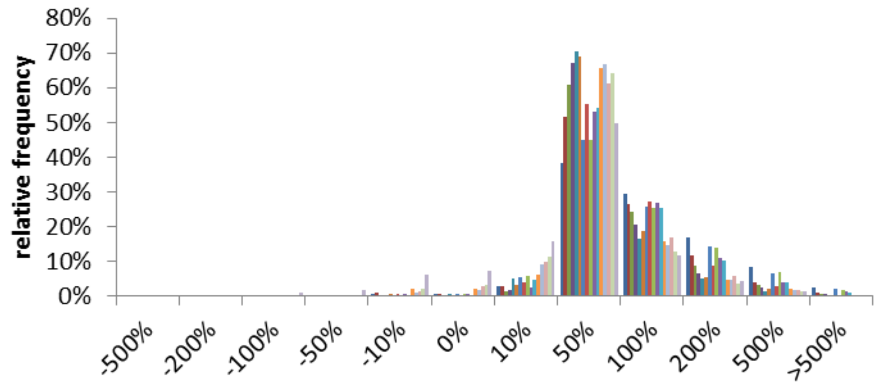


Figure Appendix 5.1.3f: Ratio (3.1) for method Quantile (0.50) with reference WLS

M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 M13 M14 M15 M16

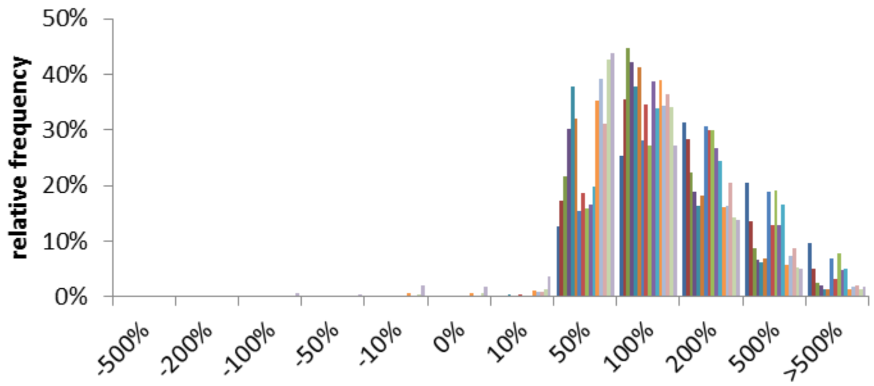


Figure Appendix 5.1.3g: Ratio (3.1) for method Quantile (0.75) with reference WLS

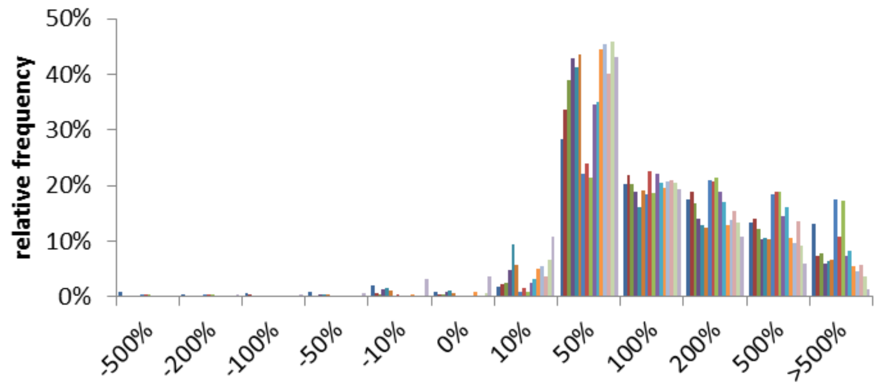


Figure Appendix 5.1.3h: Ratio (3.1) for method Quantile (0.50) with reference Quantile (0.25)

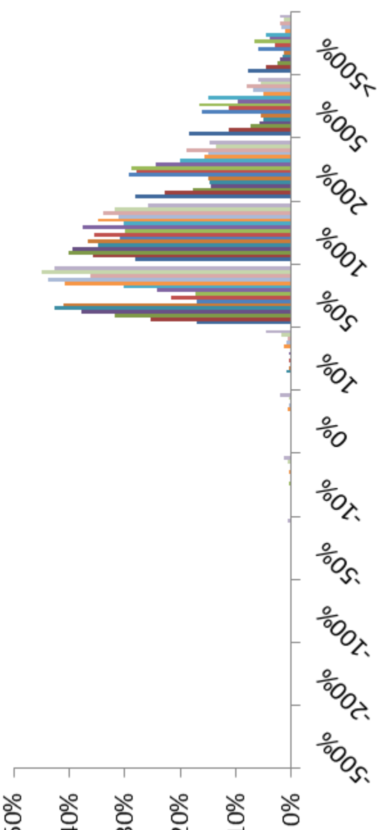


Figure Appendix 5.1.3i: Ratio (3.1) for method Quantile (0.75) with reference Quantile (0.25)

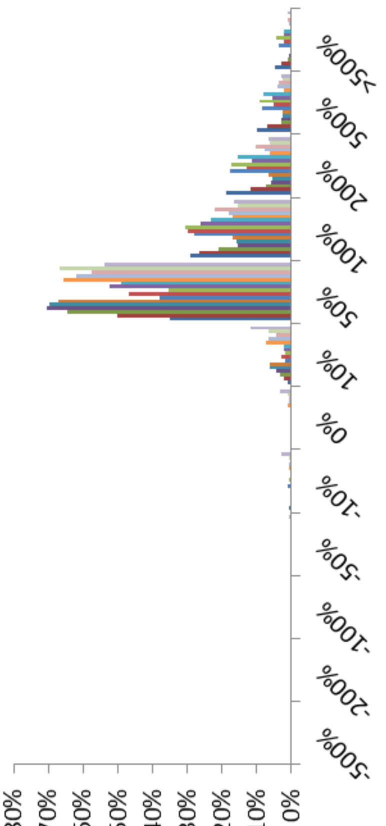
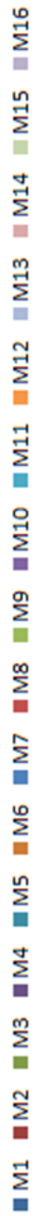


Figure Appendix 5.1.3j: Ratio (3.1) for method Quantile (0.75) with reference Quantile (0.50)



Appendix 5.1.4 Factors when Controlled for Regressions (Section 4.3.2.2)

To measure the interplay between factors and statistical methods, Figure 3.12 is analyzed for each regression separately. Recall in this connection that M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept).

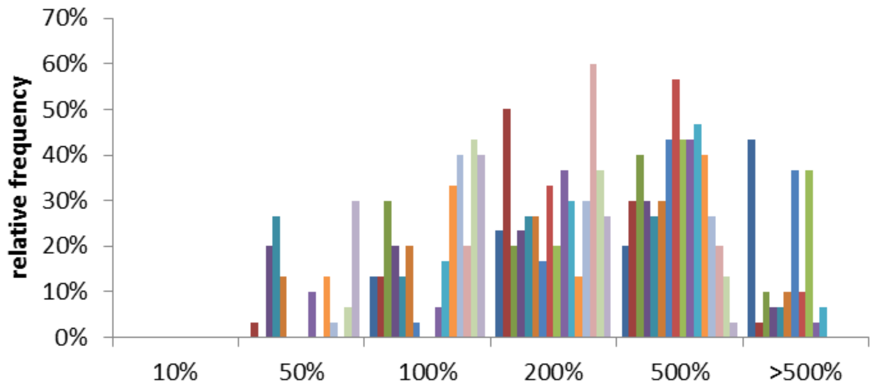


Figure Appendix 5.1.4a: Area (3.3) for method WLS with reference OLS

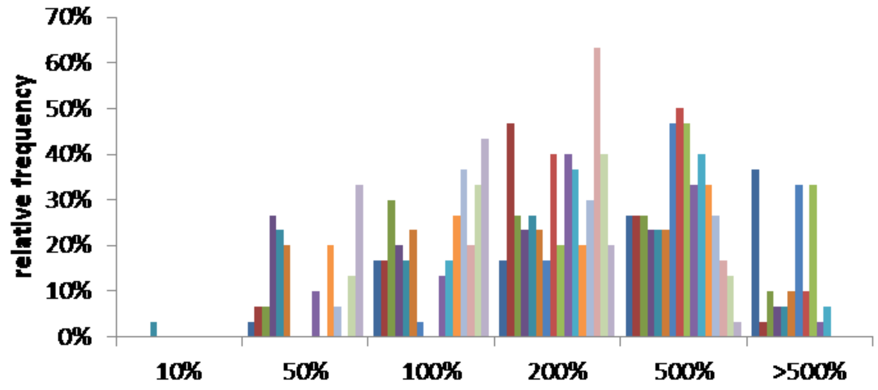


Figure Appendix 5.1.4b: Area (3.3) for method Quantile (0.25) with reference OLS

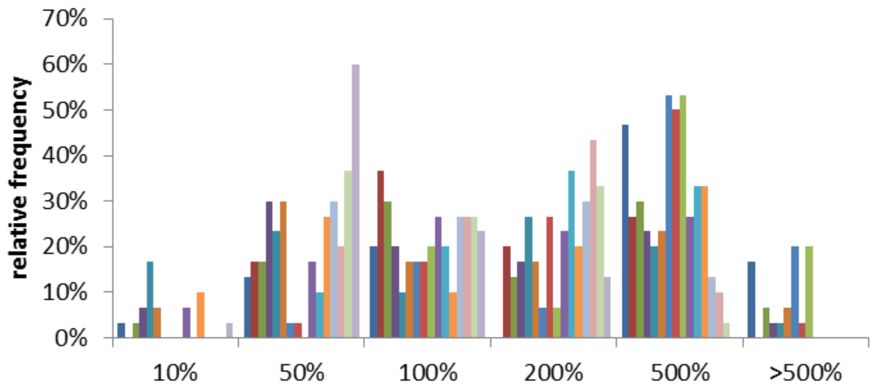


Figure Appendix 5.1.4c: Area (3.3) for method Quantile (0.50) with reference OLS

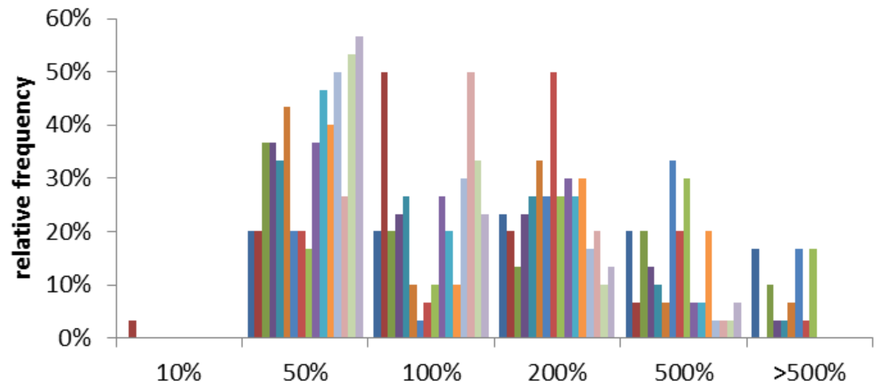


Figure Appendix 5.1.4d: Area (3.3) for method Quantile (0.75) with reference OLS

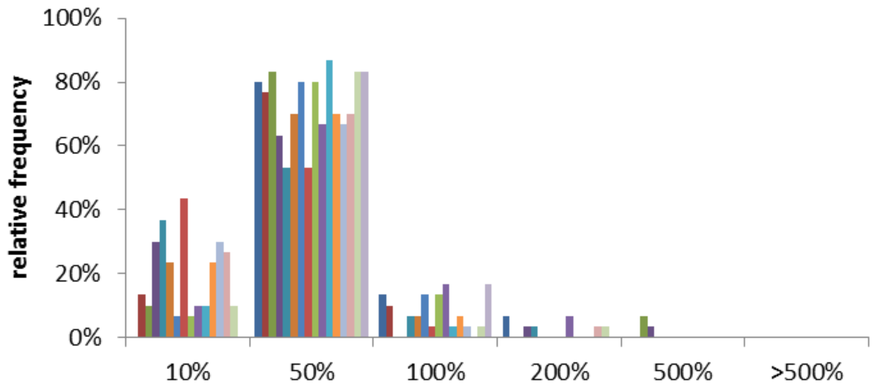


Figure Appendix 5.1.4e: Area (3.3) for method Quantile (0.25) with reference WLS

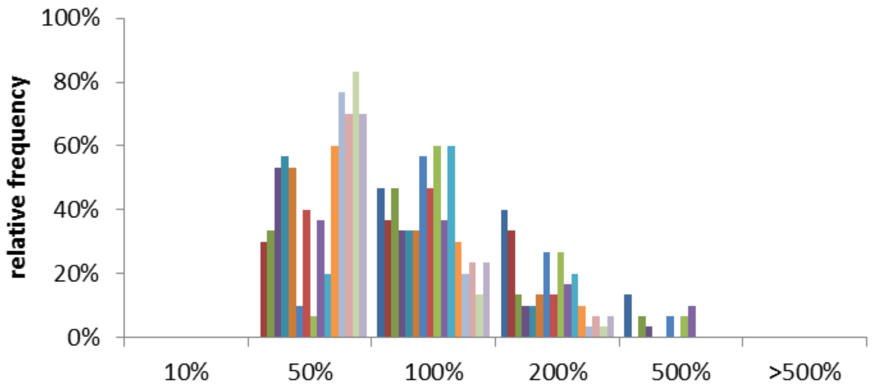


Figure Appendix 5.1.4f: Area (3.3) for method Quantile (0.50) with reference WLS

M1 M2 M3 M4 M5 M6 M7 M8 M9 M10 M11 M12 M13 M14 M15 M16

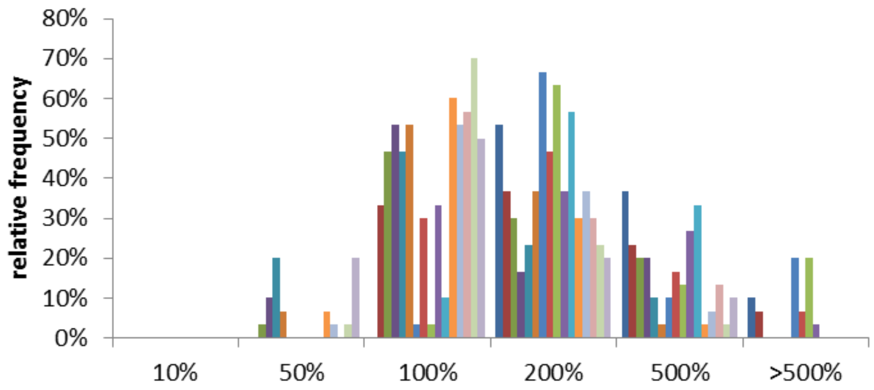


Figure Appendix 5.1.4g: Area (3.3) for method Quantile (0.75) with reference WLS

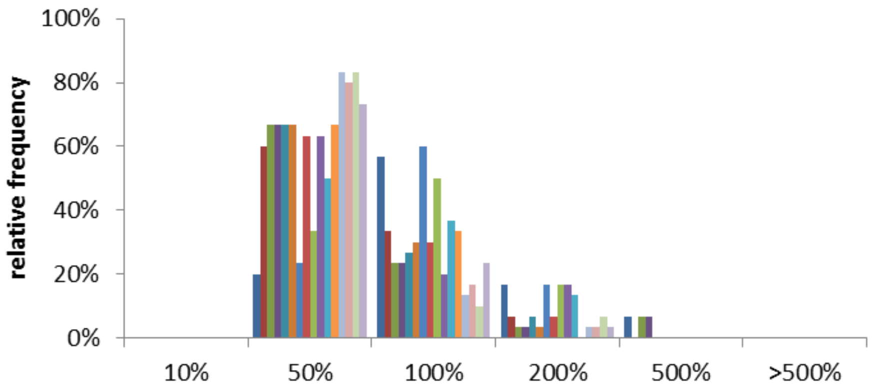


Figure Appendix 5.1.4h: Area (3.3) for method Quantile (0.50) with reference Quantile (0.25)

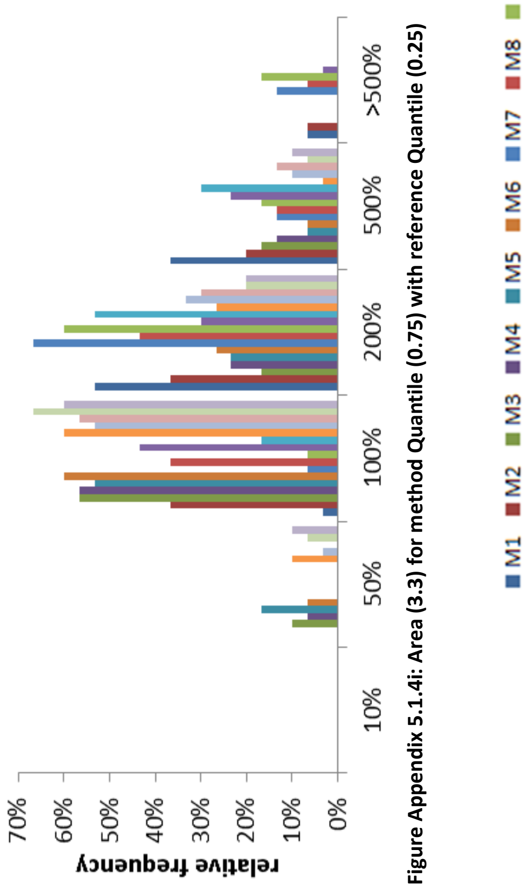


Figure Appendix 5.1.4i: Area (3.3) for method Quantile (0.75) with reference Quantile (0.25)

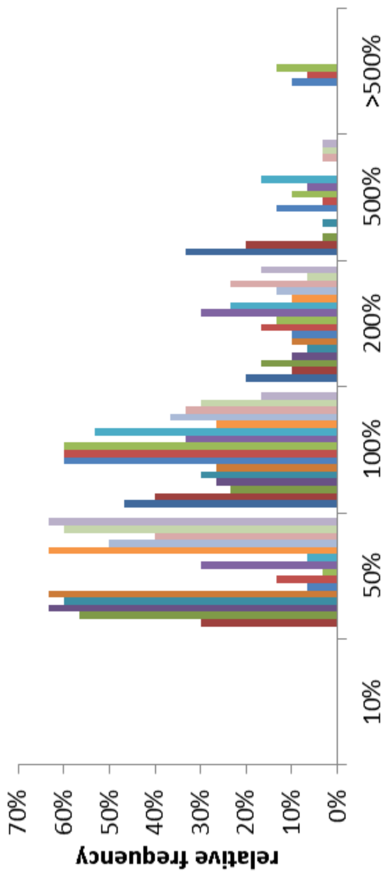


Figure Appendix 5.1.4j: Area (3.3) for method Quantile (0.75) with reference Quantile (0.50)

Appendix 5.1.5 Regressions when Controlled for Factors (Section 4.3.3.2)

To measure the interplay between factors and statistical methods, Figure 3.16 is analyzed for each factor separately. Recall in this connection that M1 to M11 are single-factor models (plus intercept); M12 to M14 are two-factor models (plus intercept); M15 is a three-factor model (plus intercept); M16 is an eleven-factor model (plus intercept).

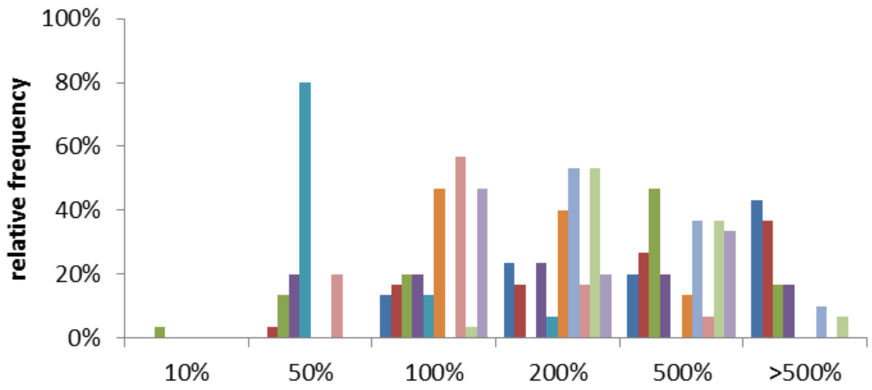


Figure Appendix 5.1.5a: Area (3.3) for model M1

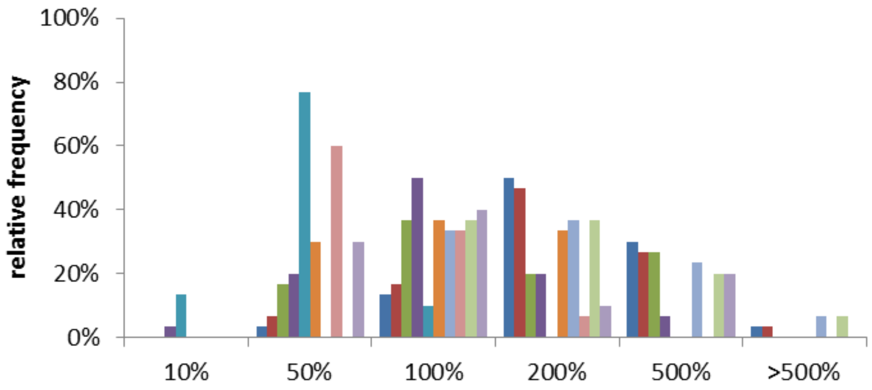


Figure Appendix 5.1.5b: Area (3.3) for model M2

OLS_WLS OLS_Quantile(0.25) OLS_Quantile(0.50) OLS_Quantile(0.75) WLS_Quantile(0.25)
WLS_Quantile(0.50) WLS_Quantile(0.75) Quantile(0.25)_Quantile(0.50) Quantile(0.25)_Quantile(0.75) Quantile(0.50)_Quantile(0.75)

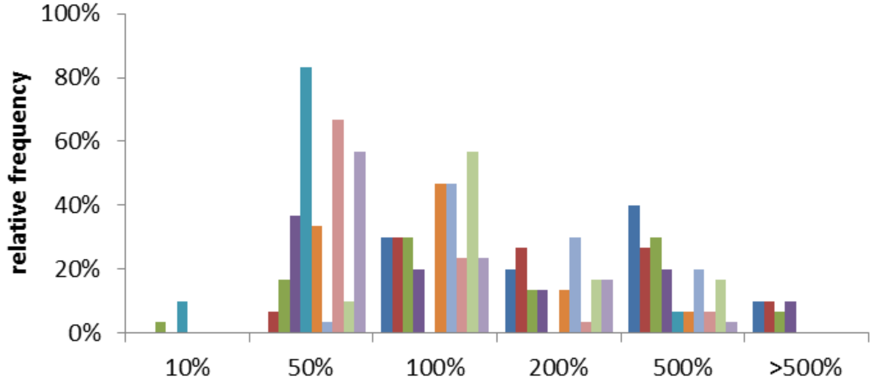


Figure Appendix 5.1.5c: Area (3.3) for model M3

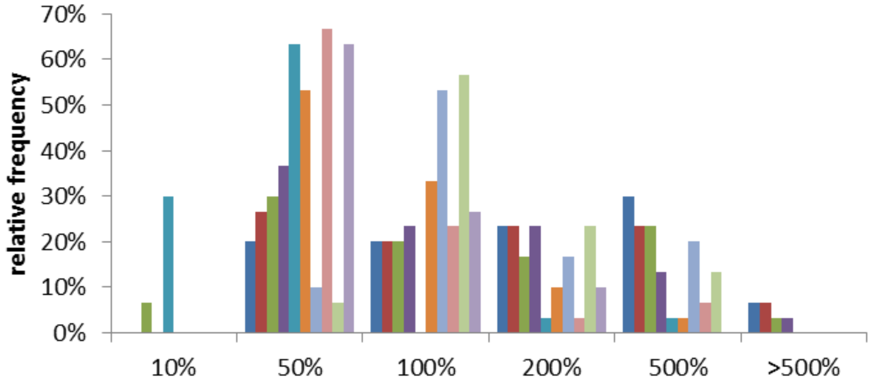


Figure Appendix 5.1.5d: Area (3.3) for model M4

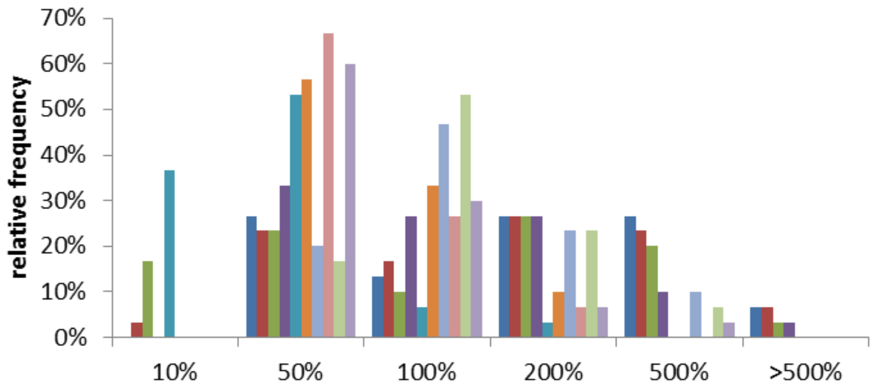


Figure Appendix 5.1.5e: Area (3.3) for model M5

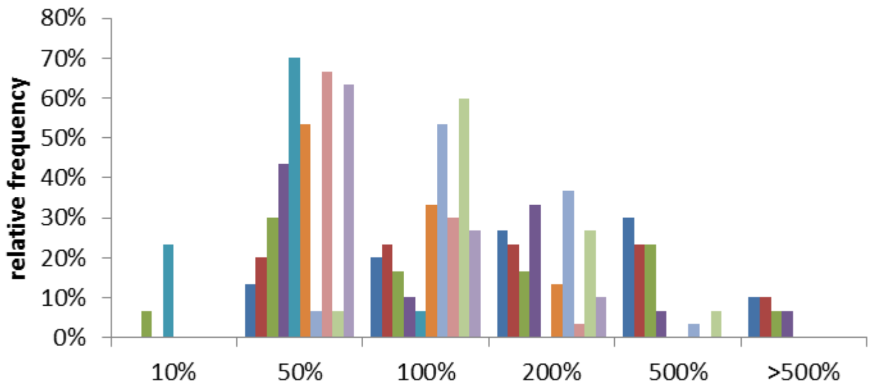
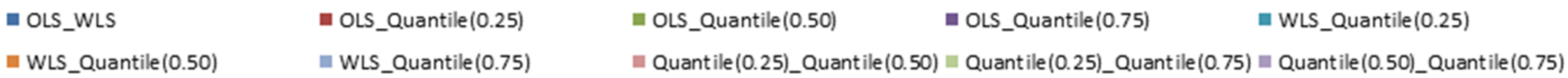


Figure Appendix 5.1.5f: Area (3.3) for model M6



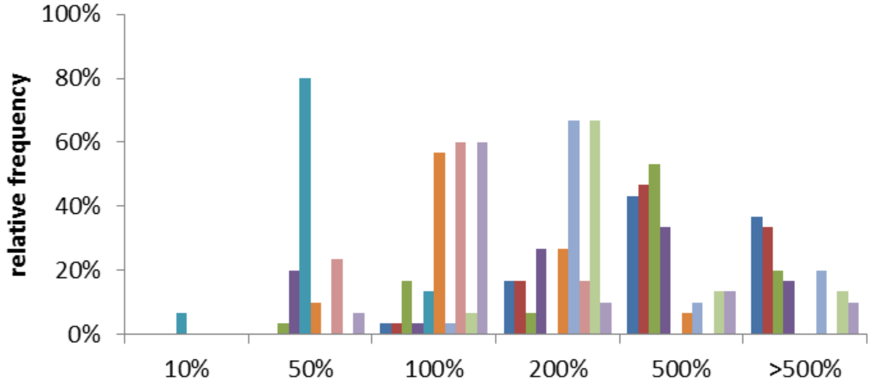


Figure Appendix 5.1.5g: Area (3.3) for model M7

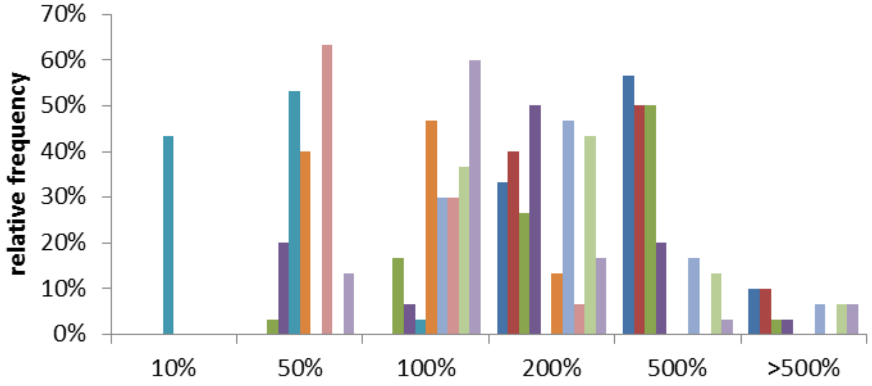


Figure Appendix 5.1.5h: Area (3.3) for model M8

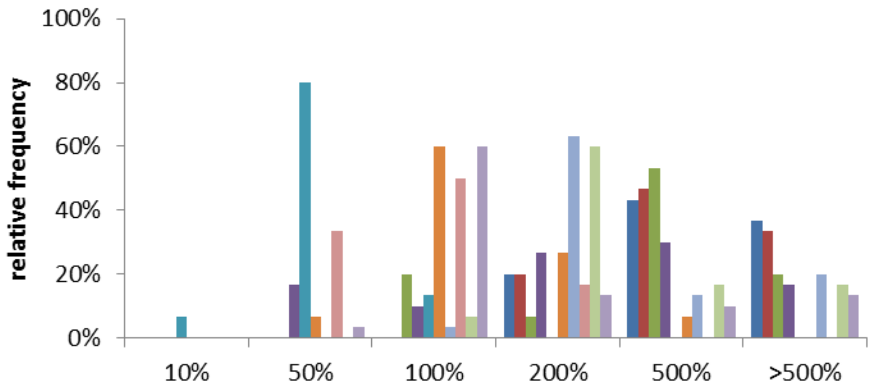


Figure Appendix 5.1.5i: Area (3.3) for model M9

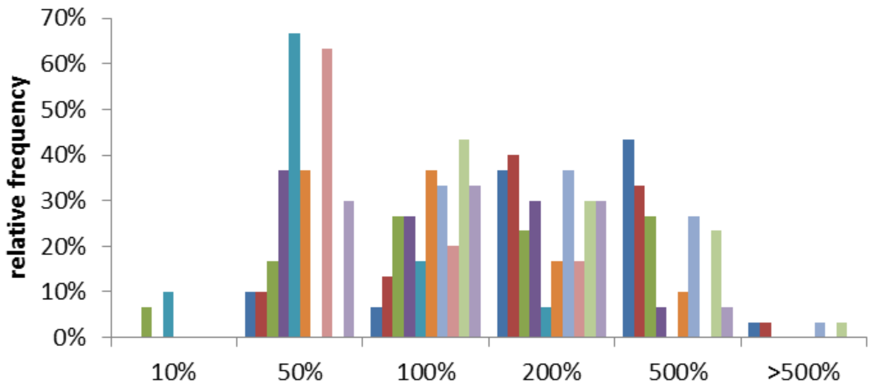
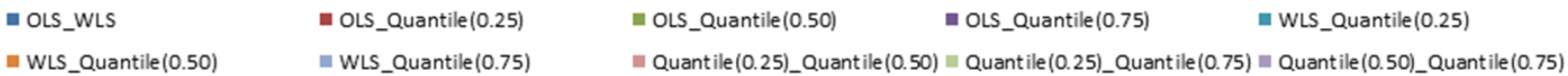


Figure Appendix 5.1.5j: Area (3.3) for model M10



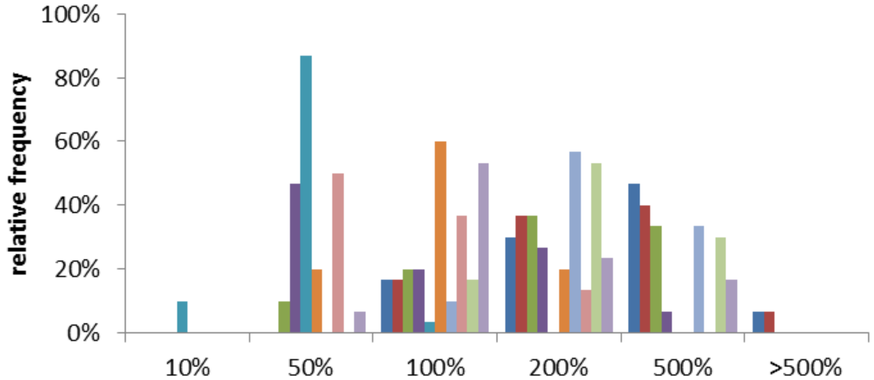


Figure Appendix 5.1.5k: Area (3.3) for model M11

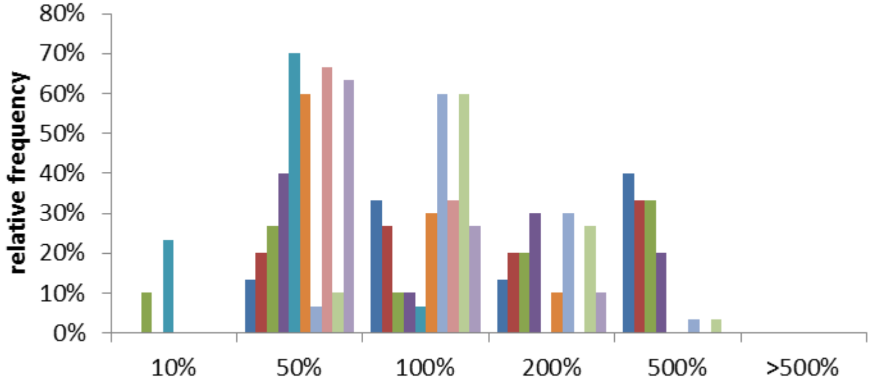


Figure Appendix 5.1.5l: Area (3.3) for model M12

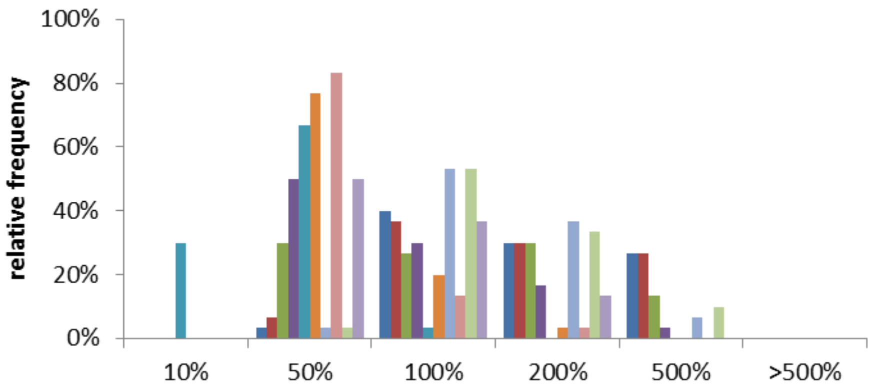


Figure Appendix 5.1.5m: Area (3.3) for model M13

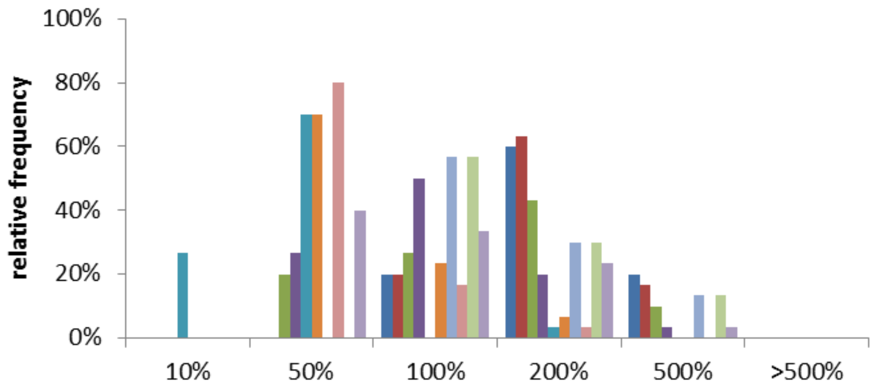
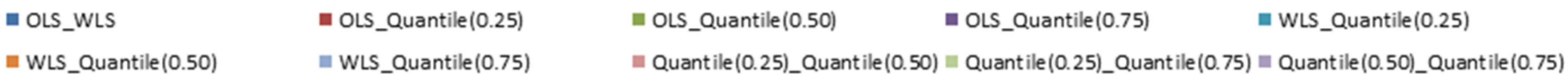


Figure Appendix 5.1.5n: Area (3.3) for model M14



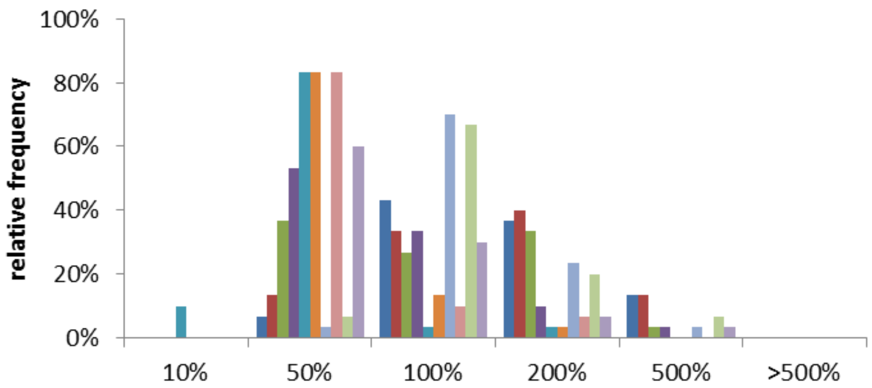


Figure Appendix 5.1.5o: Area (3.3) for model M15

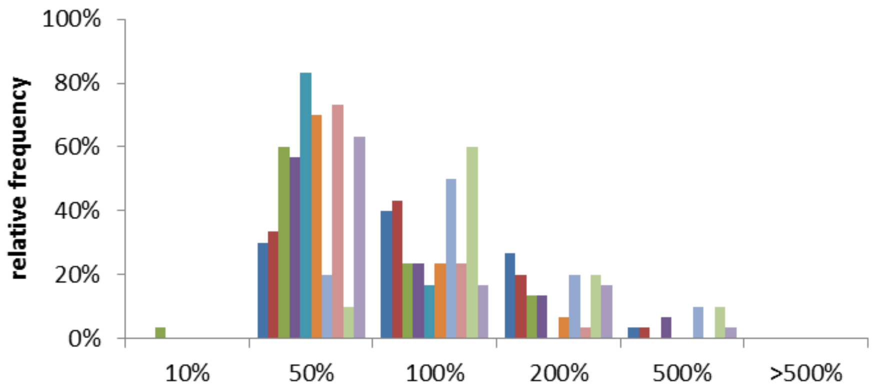


Figure Appendix 5.1.5p: Area (3.3) for model M16

Appendix 5.2 Figures in Connection with Chapter V Section 3.3

Appendix 5.2.1 Effect of Shape and Size of Constraints on Portfolio Holdings (Section 3.3.2)

Appendix 5.2.1.1 “Magnitude"

Appendix 5.2.1.1.1 The Role of Shape

– “Magnitude” of price differences of different shapes of constraints on portfolio holdings

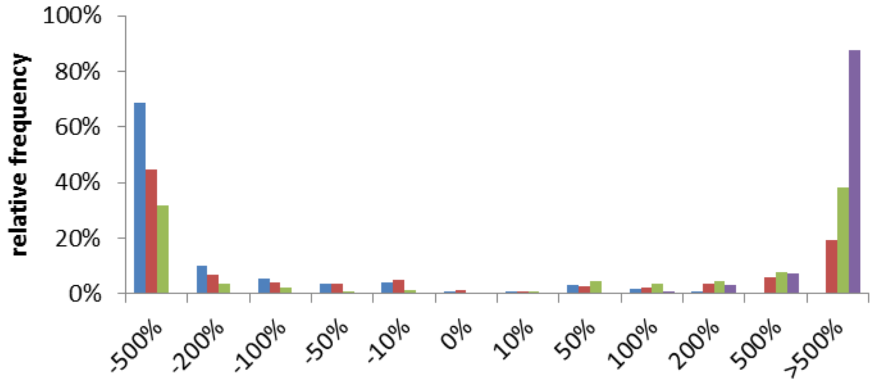


Figure Appendix 5.2.1.1.1a: Ratio (3.1) for x=0.5 and buyer

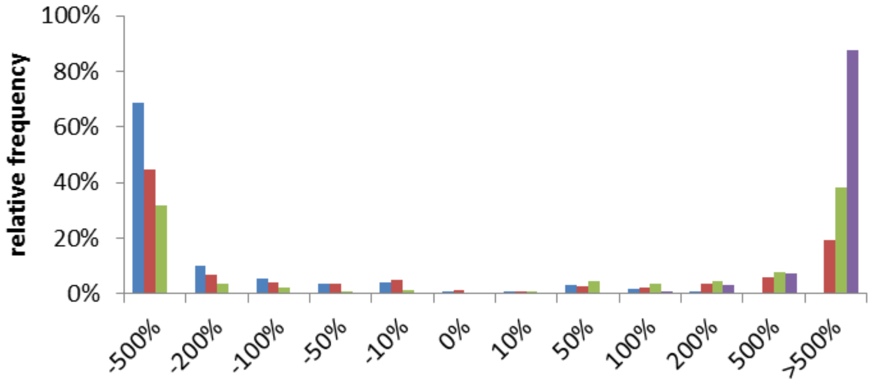


Figure Appendix 5.2.1.1.1b: Ratio (3.1) for x=1 and buyer

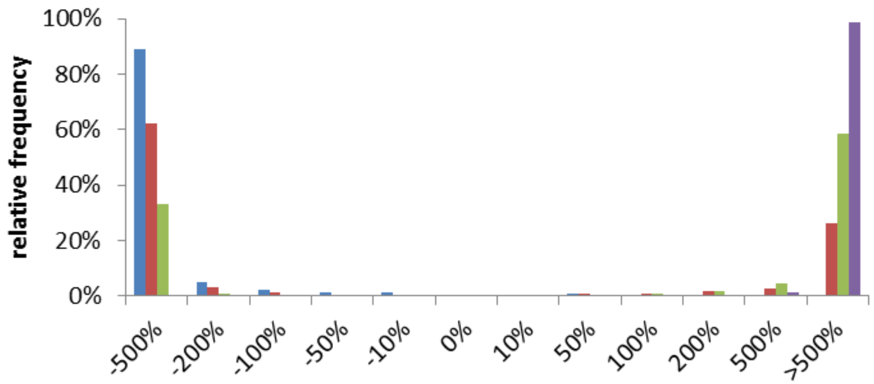


Figure Appendix 5.2.1.1.1c: Ratio (3.1) for x=2 and buyer

OTPB OTPB_L1 OTPB_L2 OTPB_LInf

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_L1) and seller (OTPS_L1)
L_2-norm constraint and value for buyer (OTPB_L2) and seller (OTPS_L2)
L_∞-norm constraint and value for buyer (OTPB_LInf) and seller (OTPS_LInf)

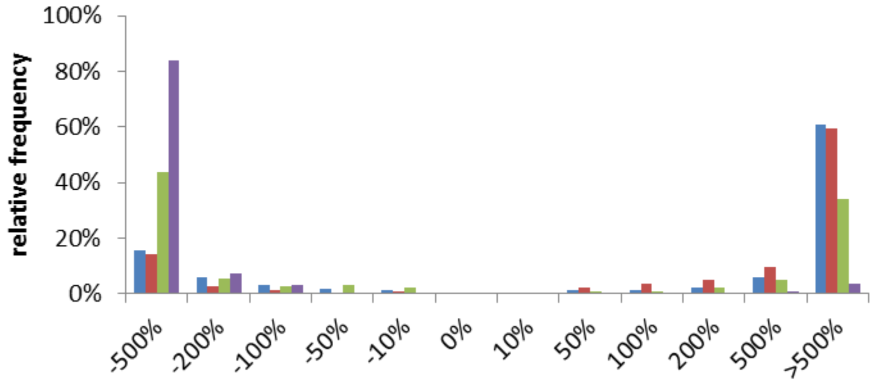


Figure Appendix 5.2.1.1.1d: Ratio (3.1) for x=0.5 and seller

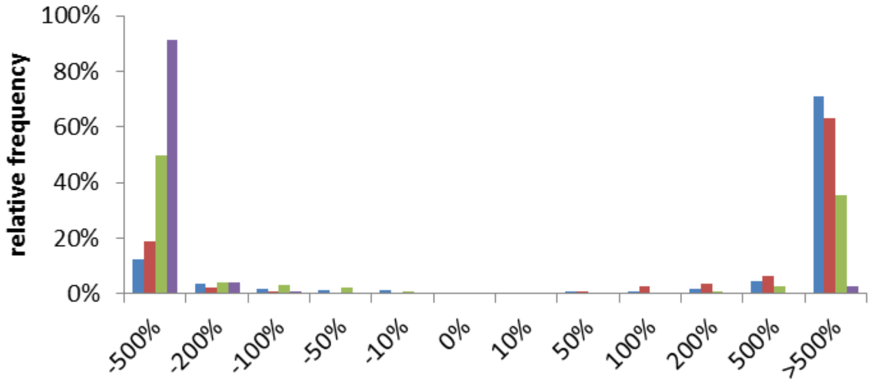


Figure Appendix 5.2.1.1.1e: Ratio (3.1) for x=1 and seller

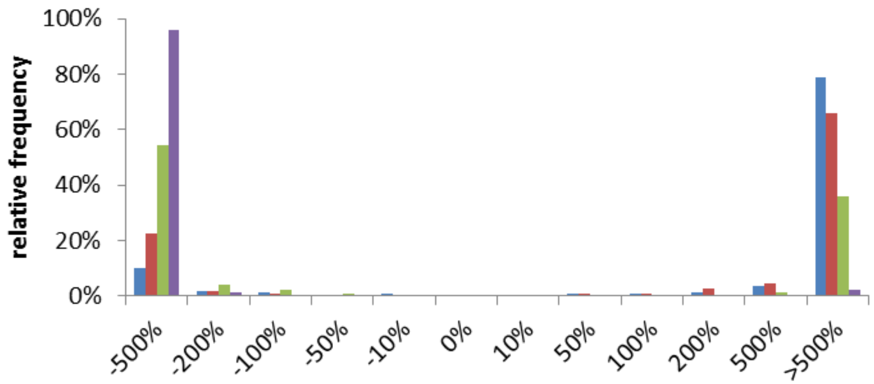


Figure Appendix 5.2.1.1.1f: Ratio (3.1) for x=2 and seller

OTPS OTPS_L1 OTPS_L2 OTPS_LInf

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_L1) and seller (OTPS_L1)
L_2-norm constraint and value for buyer (OTPB_L2) and seller (OTPS_L2)
L_∞-norm constraint and value for buyer (OTPB_LInf) and seller (OTPS_LInf)

– “Magnitude” of price differences of different shapes of constraints on portfolio holdings broken down by factors

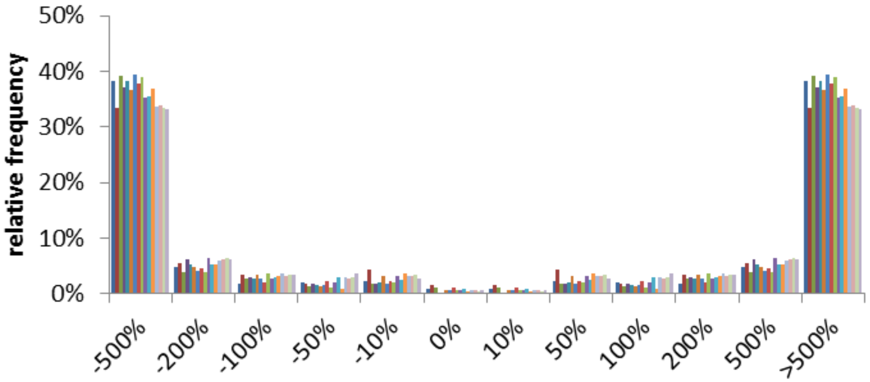


Figure Appendix 5.2.1.1.1g: Ratio (3.1) for x=0.5 and buyer broken down by factors

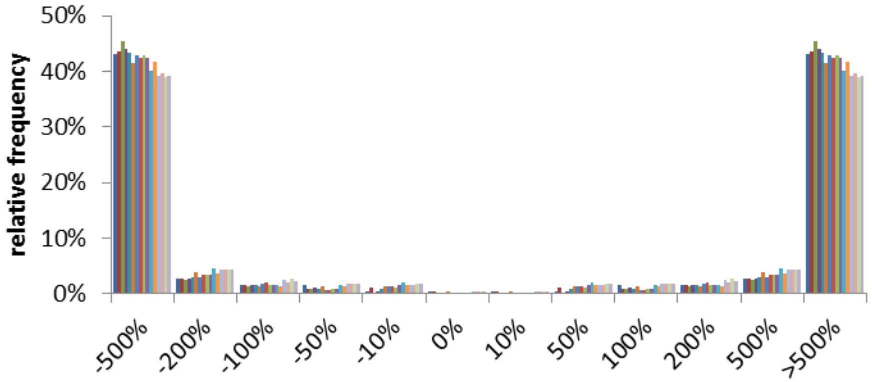


Figure Appendix 5.2.1.1.1h: Ratio (3.1) for x=1 and buyer broken down by factors

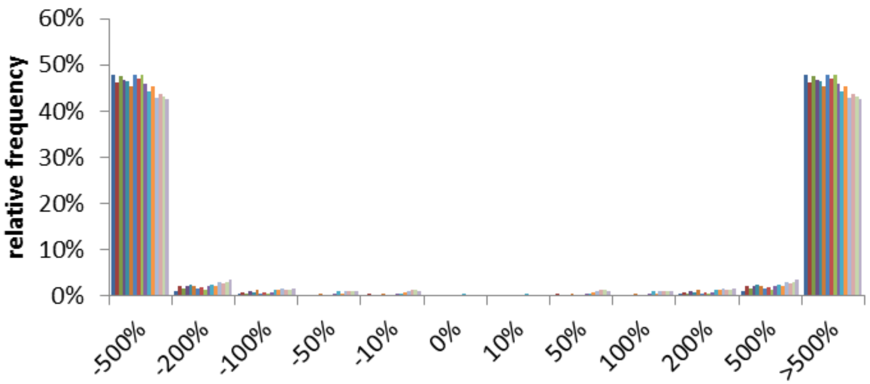


Figure Appendix 5.2.1.1.1i: Ratio (3.1) for x=2 and buyer broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

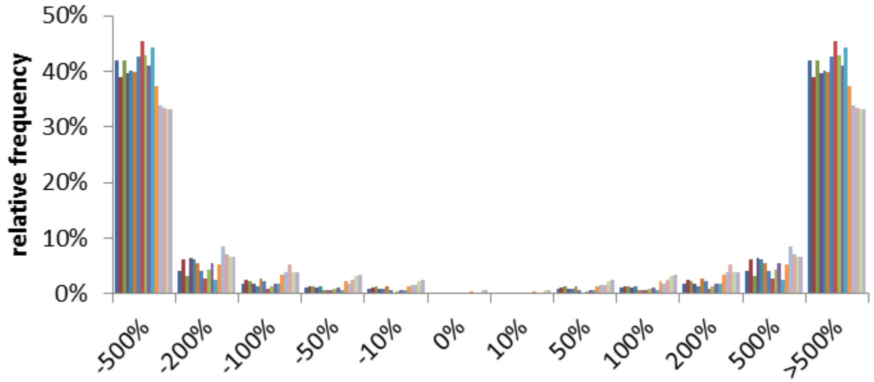


Figure Appendix 5.2.1.1.1j: Ratio (3.1) for x=0.5 and seller broken down by factors

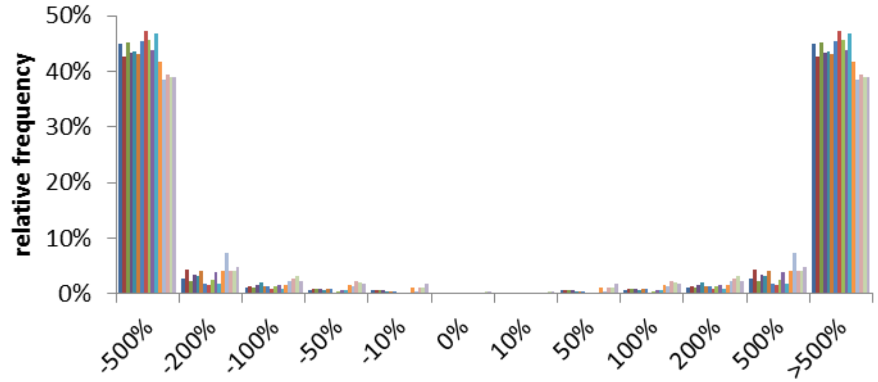


Figure Appendix 5.2.1.1.1k: Ratio (3.1) for x=1 and seller broken down by factors

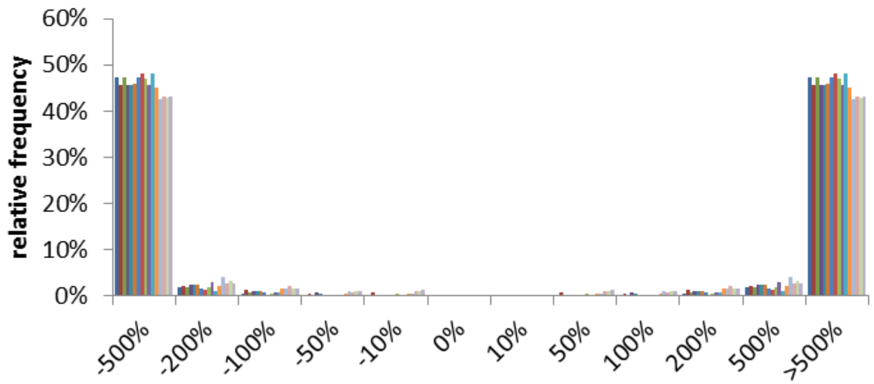


Figure Appendix 5.2.1.1.1l: Ratio (3.1) for x=2 and seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

Appendix 5.2.1.1.2 The Role of Size

– "Magnitude" of price differences of different sizes of constraints on portfolio holdings

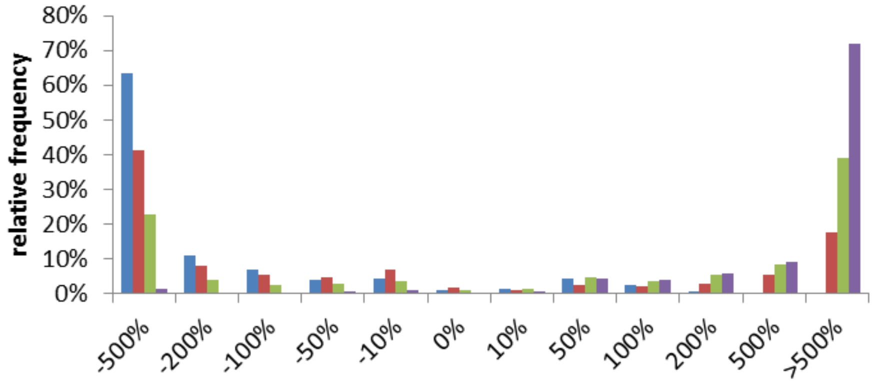


Figure Appendix 5.2.1.1.2a: Ratio (3.1) for L_1 and buyer

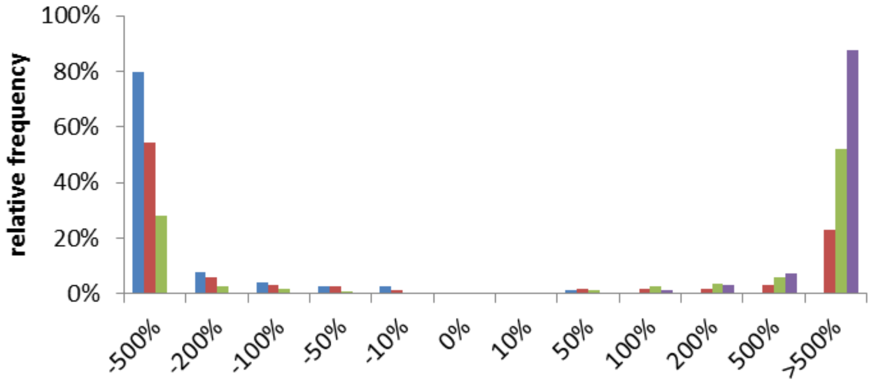


Figure Appendix 5.2.1.1.2b: Ratio (3.1) for L_2 and buyer

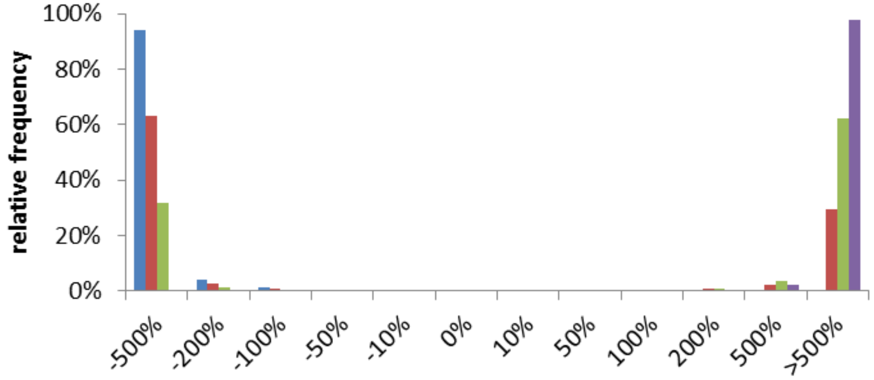


Figure Appendix 5.2.1.1.2c: Ratio (3.1) for L_∞ and buyer

OTPB OTPB_x0.5 OTPB_x1 OTPB_x2

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_x0.5) and seller (OTPS_x0.5)
L_2-norm constraint and value for buyer (OTPB_x1) and seller (OTPS_x1)
L_∞-norm constraint and value for buyer (OTPB_x2) and seller (OTPS_x2)

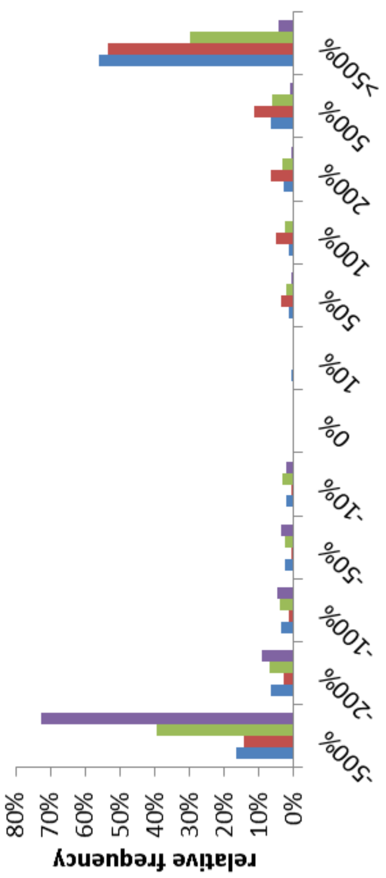


Figure Appendix 5.2.1.1.2d: Ratio (3.1) for L_1 and seller

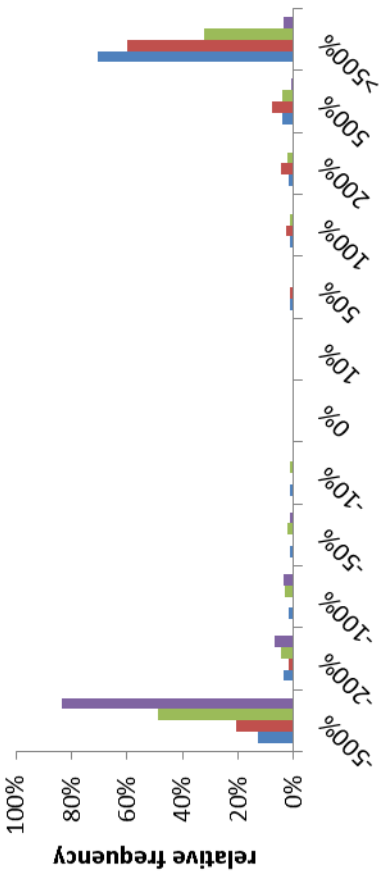


Figure Appendix 5.2.1.1.2e: Ratio (3.1) for L_2 and seller

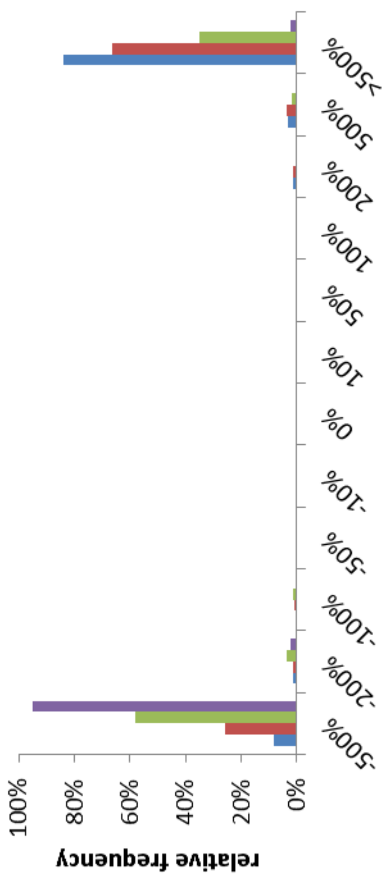


Figure Appendix 5.2.1.1.2f: Ratio (3.1) for L_∞ and seller

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_x0.5) and seller (OTPS_x0.5)
L_2-norm constraint and value for buyer (OTPB_x1) and seller (OTPS_x1)
L_∞-norm constraint and value for buyer (OTPB_x2) and seller (OTPS_x2)

– “Magnitude” of price differences of different sizes of constraints on portfolio holdings broken down by factors

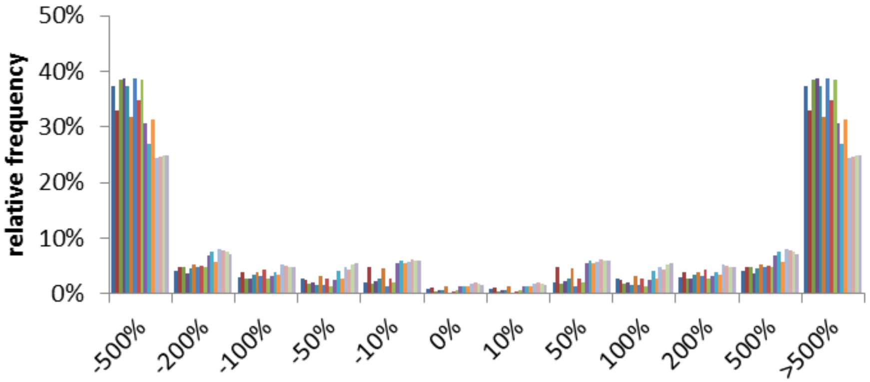


Figure Appendix 5.2.1.1.2g: Ratio (3.1) for L_1 and buyer broken down by factors

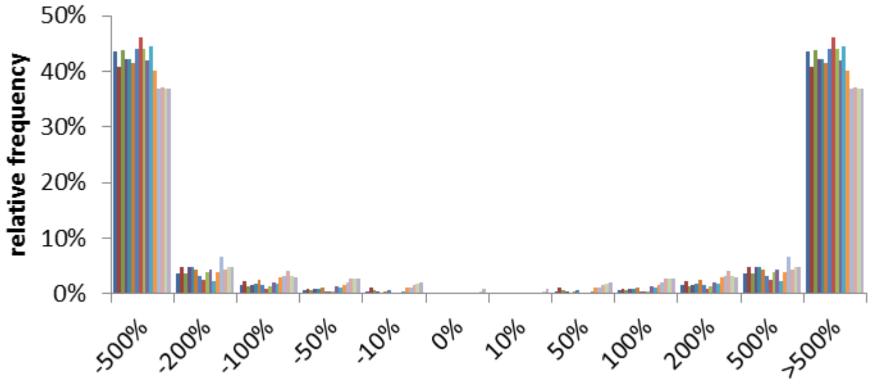


Figure Appendix 5.2.1.1.2h: Ratio (3.1) for L_2 and buyer broken down by factors

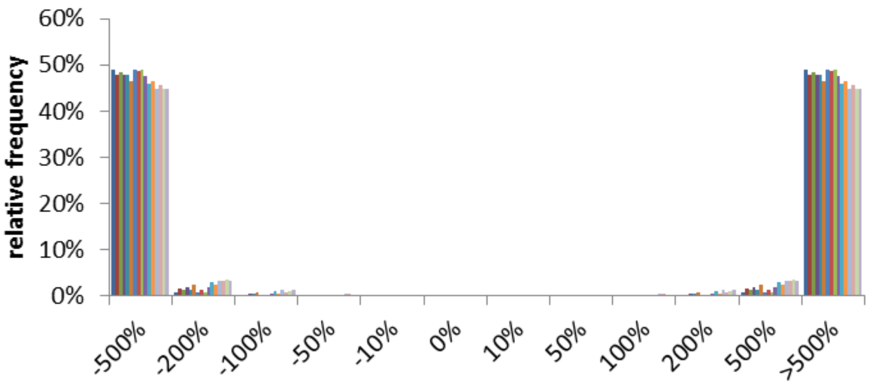


Figure Appendix 5.2.1.1.2i: Ratio (3.1) for L_∞ and buyer broken down by factors

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

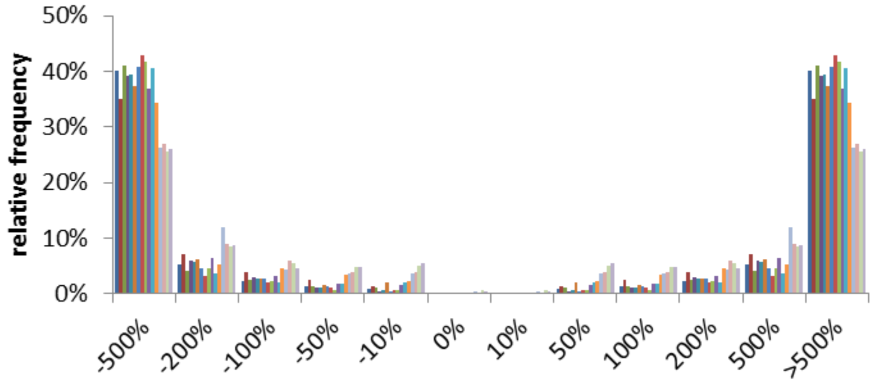


Figure Appendix 5.2.1.1.2j: Ratio (3.1) for L_1 and seller broken down by factors

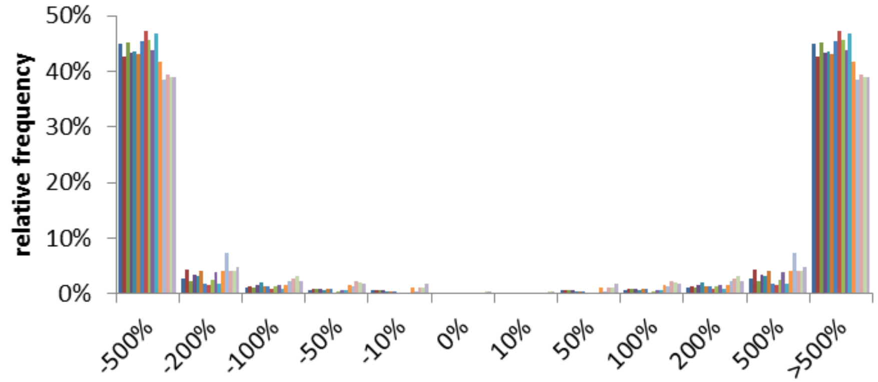


Figure Appendix 5.2.1.1.2k: Ratio (3.1) for L_2 and seller broken down by factors

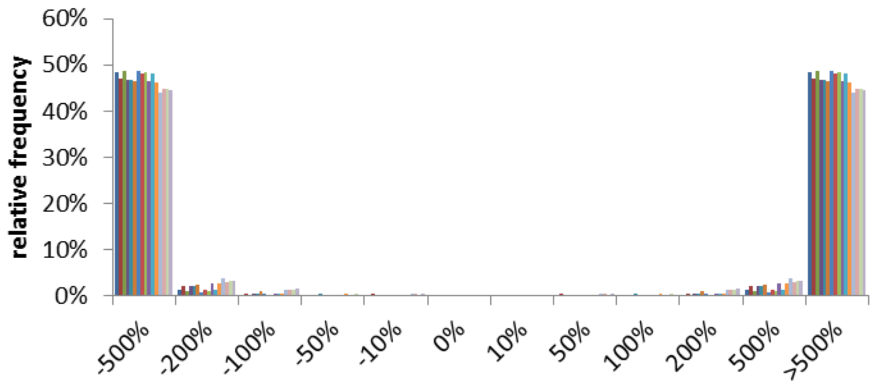


Figure Appendix 5.2.1.1.2l: Ratio (3.1) for L_∞ and seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

Appendix 5.2.1.1.3 Prices of Buyers versus Sellers

– “Magnitude” of price differences of buyers and seller for all shapes and sizes of constraints on portfolio holdings

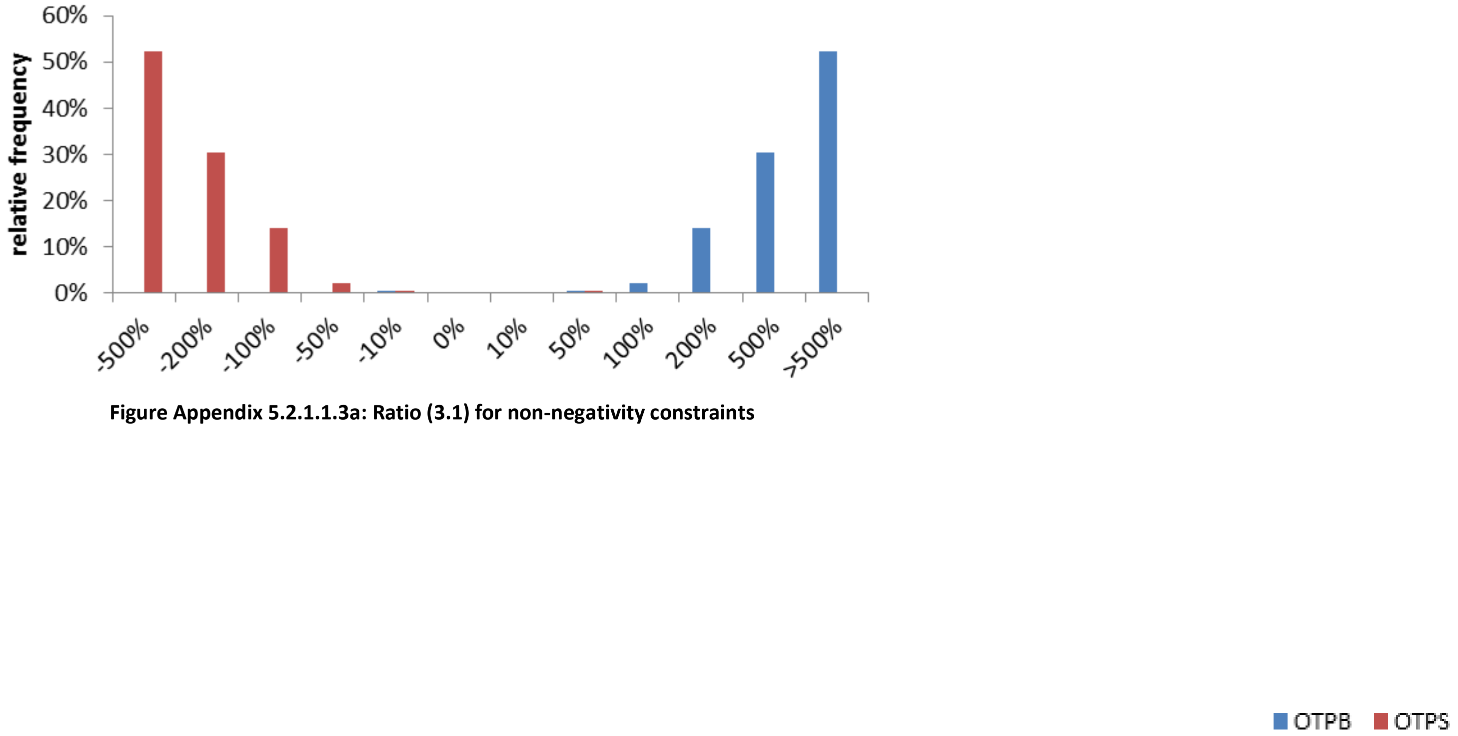


Figure Appendix 5.2.1.1.3a: Ratio (3.1) for non-negativity constraints

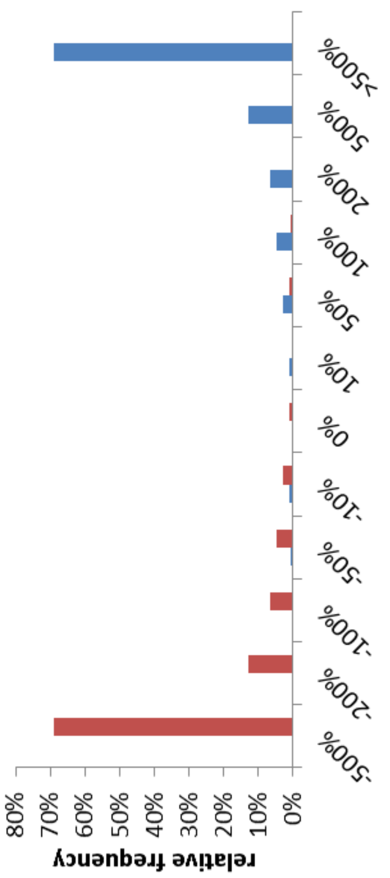


Figure Appendix 5.2.1.1.3b: Ratio (3.1) for L_1 and x=0.5

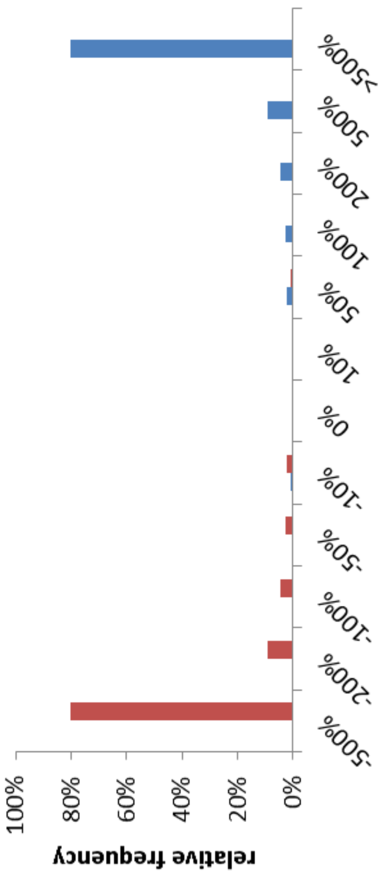
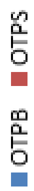


Figure Appendix 5.2.1.1.3c: Ratio (3.1) for L_1 and x=1



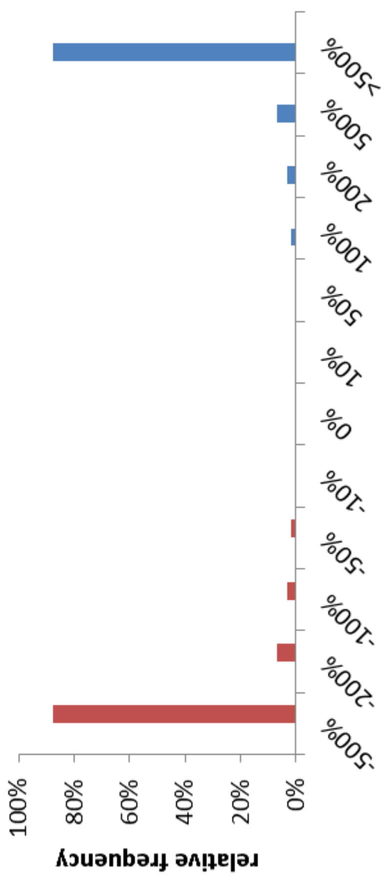


Figure Appendix 5.2.1.1.3d: Ratio (3.1) for L_1 and x=2

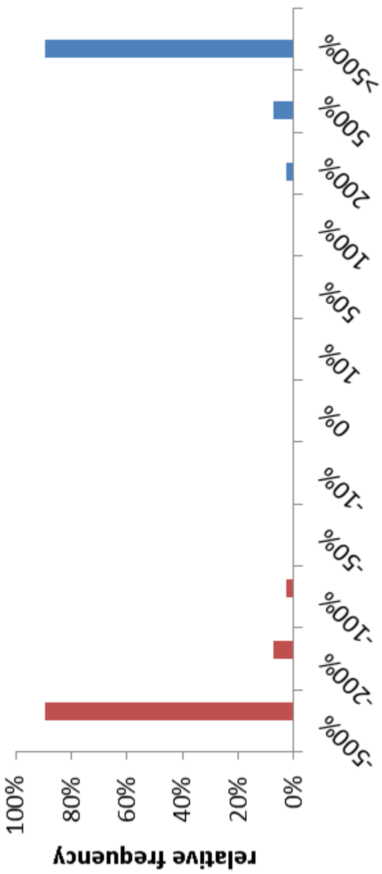
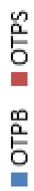


Figure Appendix 5.2.1.1.3f: Ratio (3.1) for L_2 and x=1



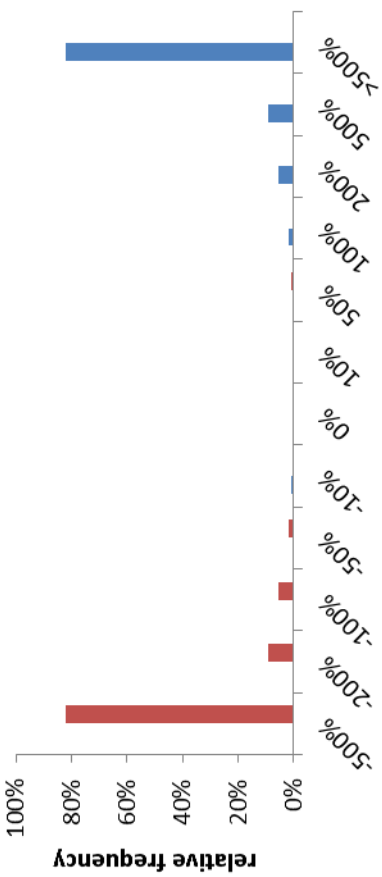


Figure Appendix 5.2.1.1.3e: Ratio (3.1) for L_2 and x=0.5

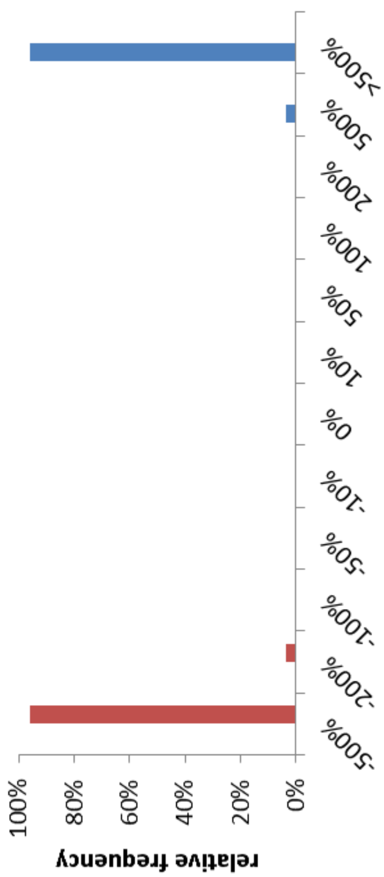


Figure Appendix 5.2.1.1.3g: Ratio (3.1) for L_2 and x=2

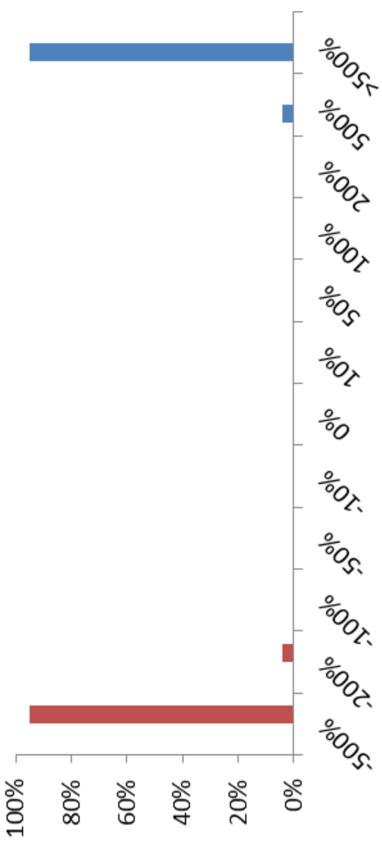


Figure Appendix 5.2.1.1.3h: Ratio (3.1) for L∞ and x=0.5

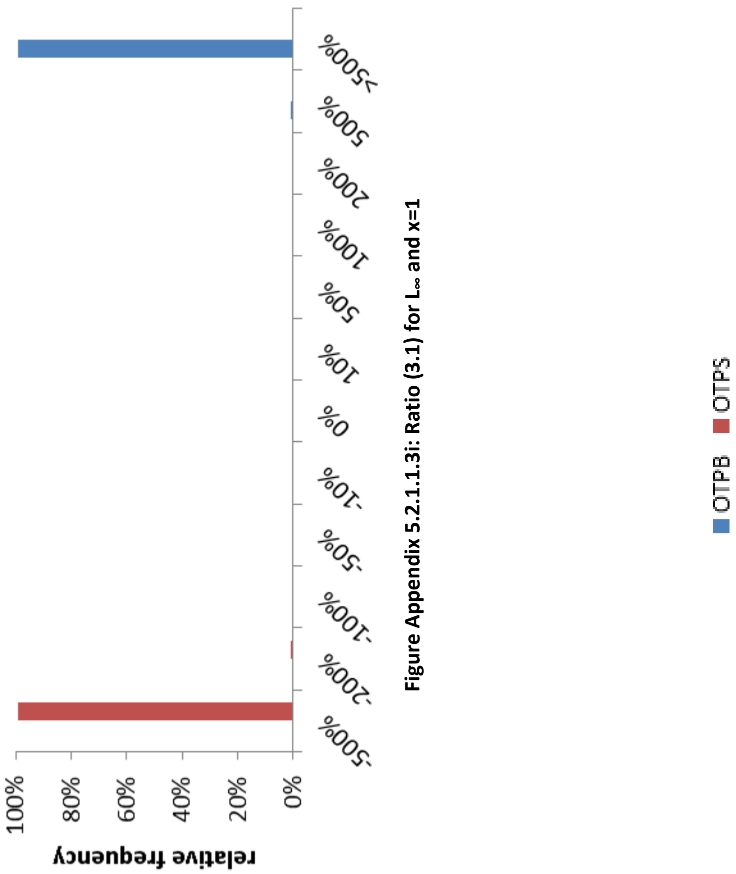


Figure Appendix 5.2.1.1.3i: Ratio (3.1) for L∞ and x=1

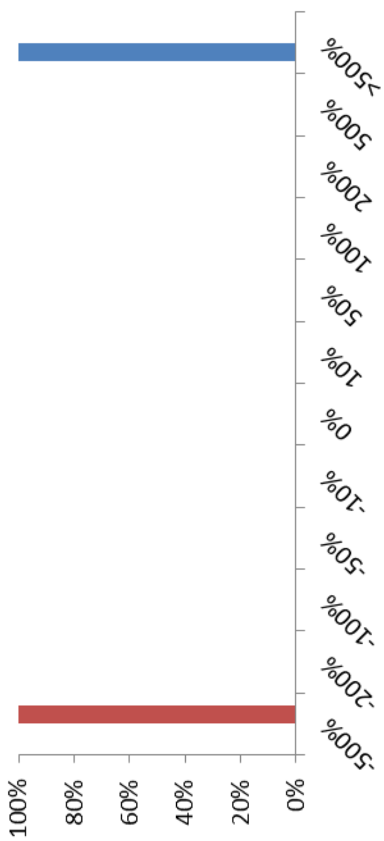


Figure Appendix 5.2.1.1.3j: Ratio (3.1) for L∞ and x=2

– "Magnitude" of price differences of buyers and seller for all shapes and sizes of constraints on portfolio holdings broken down by factors

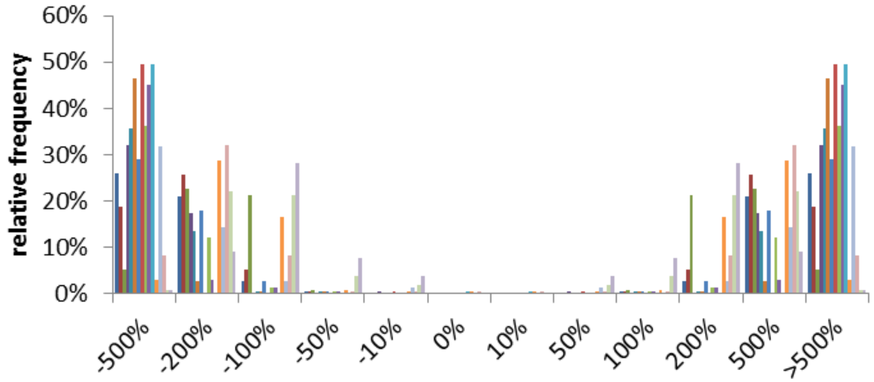


Figure Appendix 5.2.1.1.3k: Ratio (3.1) for non-negativity constraints broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

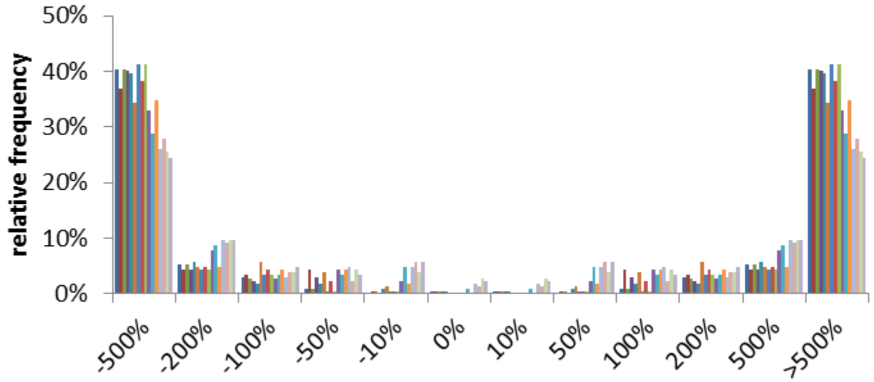


Figure Appendix 5.2.1.1.3l: Ratio (3.1) for L_1 and x=0.5 broken down by factors

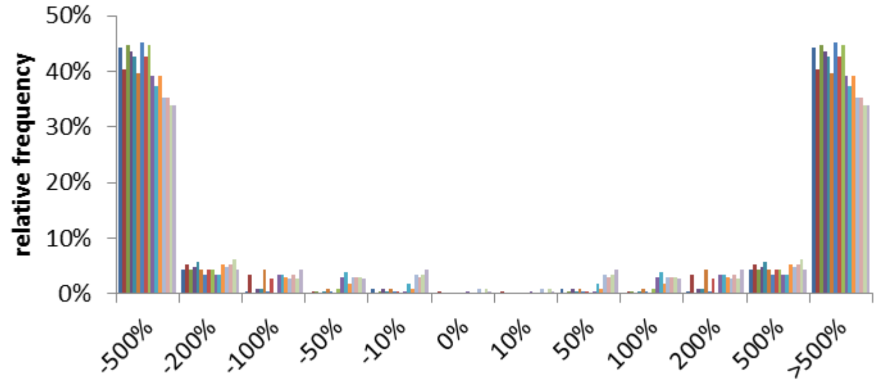


Figure Appendix 5.2.1.1.3m: Ratio (3.1) for L_1 and x=1 broken down by factors

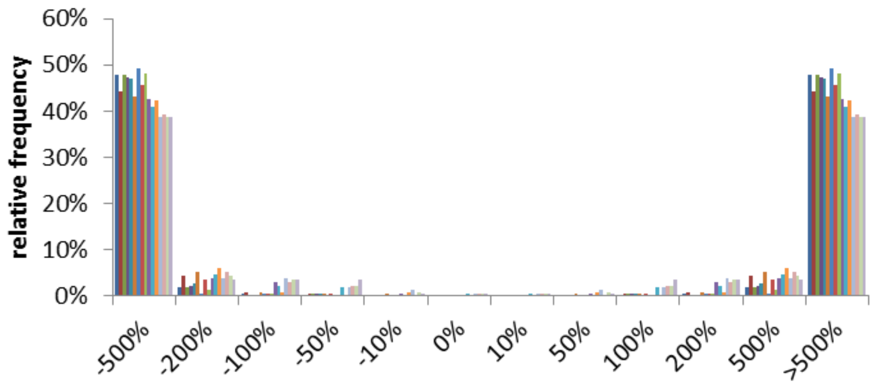


Figure Appendix 5.2.1.1.3n: Ratio (3.1) for L_1 and x=2 broken down by factors

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

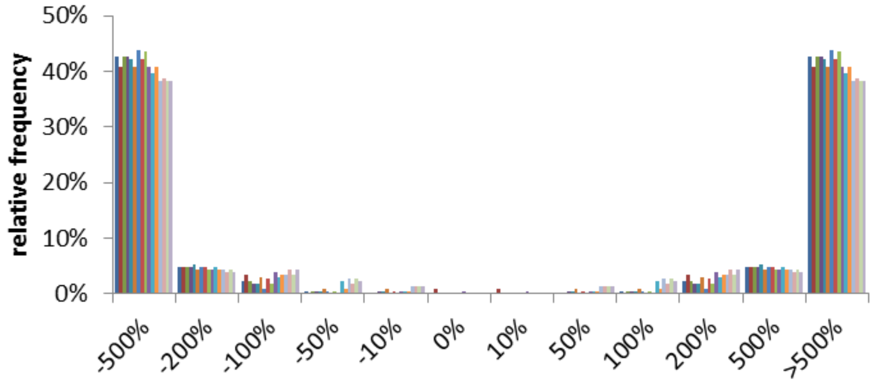


Figure Appendix 5.2.1.1.3o: Ratio (3.1) for L_2 and x=0.5 broken down by factors

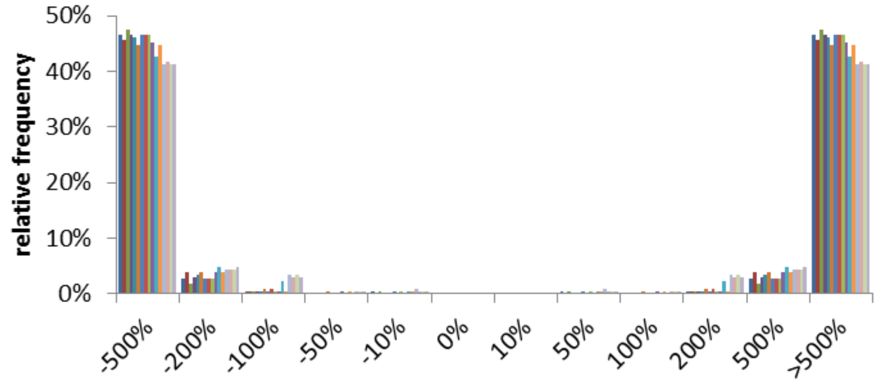


Figure Appendix 5.2.1.1.3p: Ratio (3.1) for L_2 and x=1 broken down by factors

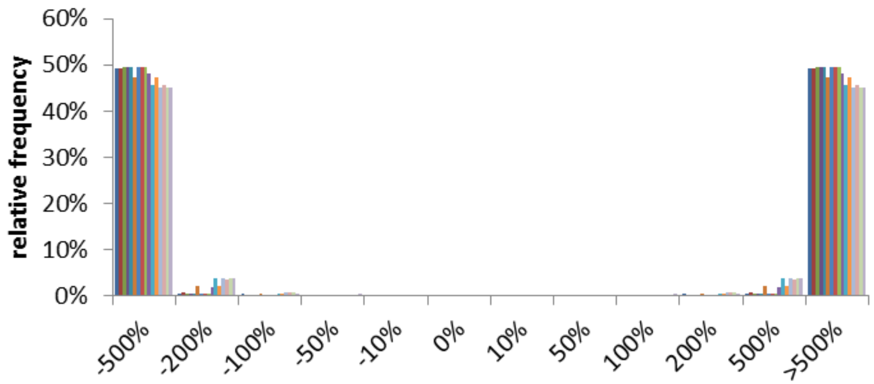


Figure Appendix 5.2.1.1.3q: Ratio (3.1) for L_2 and x=2 broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

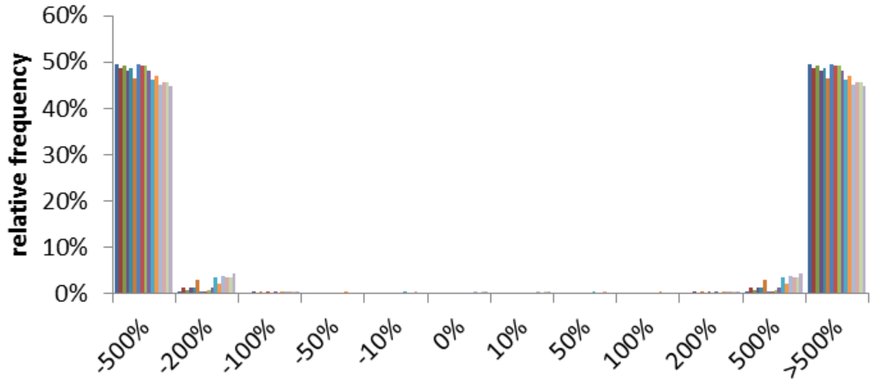


Figure Appendix 5.2.1.1.3r: Ratio (3.1) for L∞ and x=0.5 broken down by factors

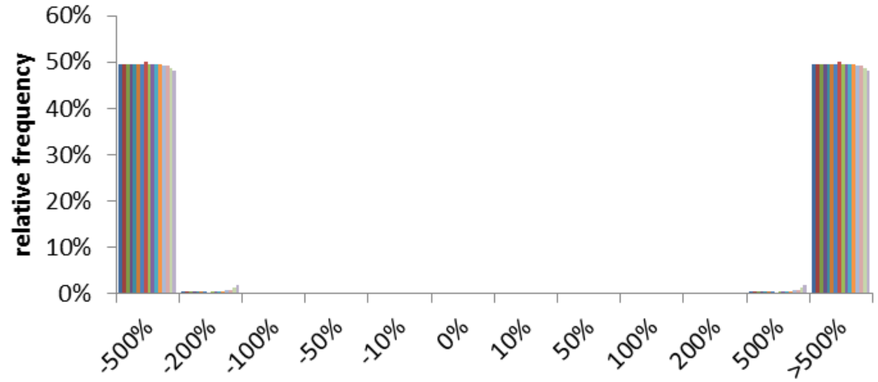


Figure Appendix 5.2.1.1.3s: Ratio (3.1) for L∞ and x=1 broken down by factors

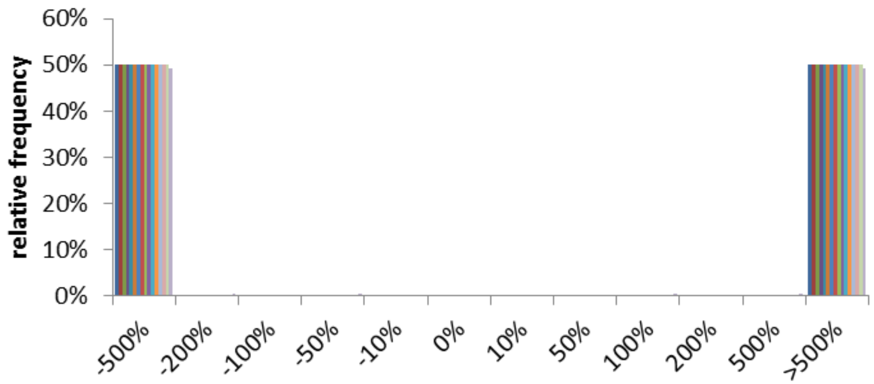


Figure Appendix 5.2.1.1.3t: Ratio (3.1) for L∞ and x=2 broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

Appendix 5.2.1.2 "Similarity"

Appendix 5.2.1.2.1 The Role of Shape

– “Similarity” of price differences of different shapes of constraints on portfolio holdings

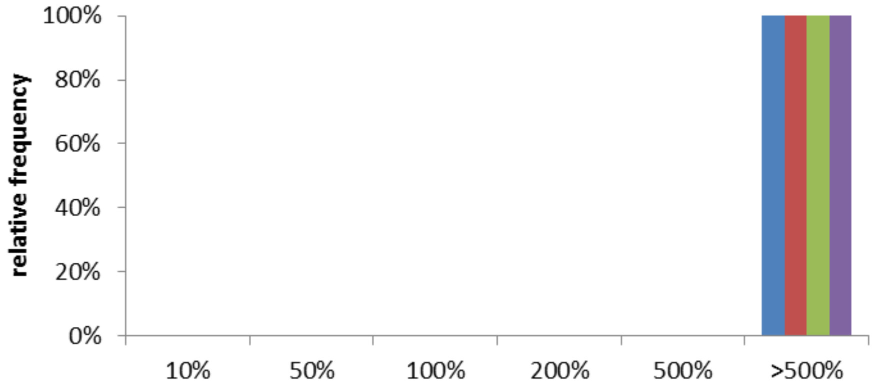


Figure Appendix 5.2.1.2.1a: Area (3.3) for x=0.5 and buyer

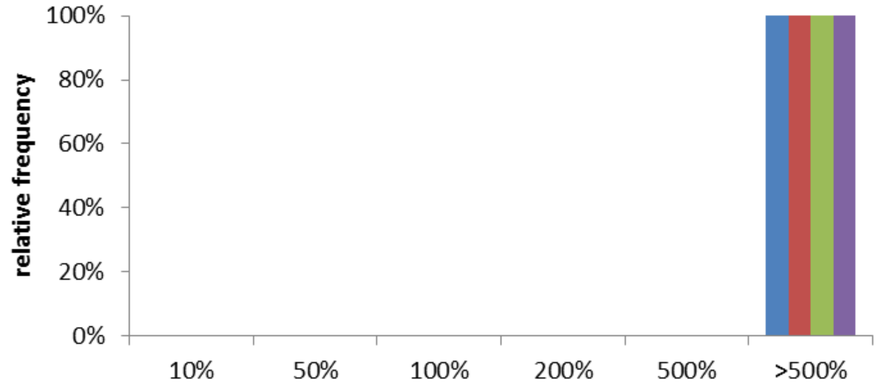


Figure Appendix 5.2.1.2.1b: Area (3.3) for x=1 and buyer

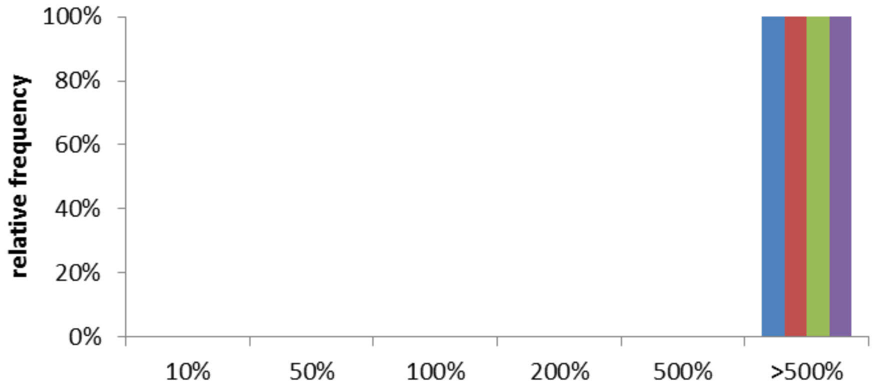


Figure Appendix 5.2.1.2.1c: Area (3.3) for x=2 and buyer

OTPB OTPB_L1 OTPB_L2 OTPB_LInf

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_L1) and seller (OTPS_L1)
L_2-norm constraint and value for buyer (OTPB_L2) and seller (OTPS_L2)
L_∞-norm constraint and value for buyer (OTPB_LInf) and seller (OTPS_LInf)

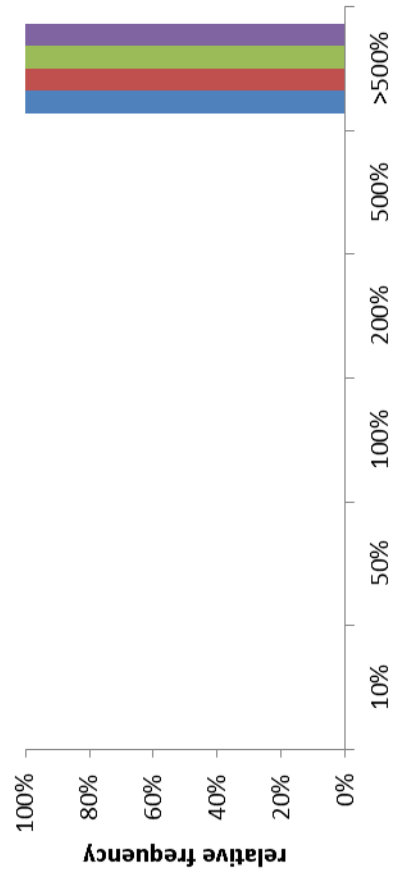


Figure Appendix 5.2.1.2.1d: Area (3.3) for x=0.5 and seller

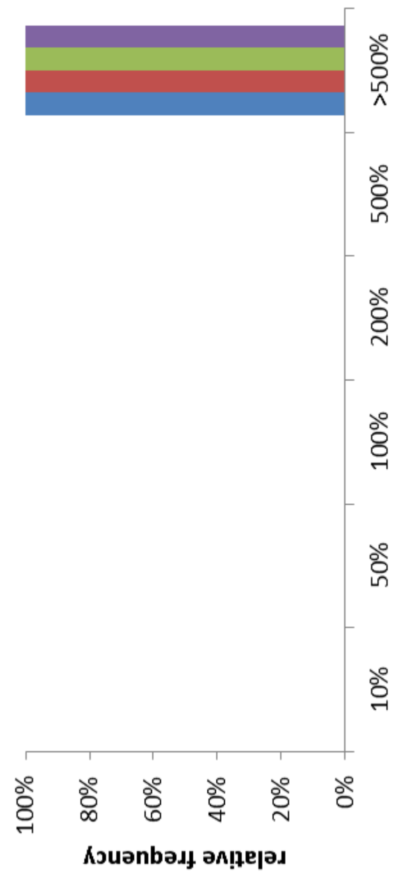


Figure Appendix 5.2.1.2.1e: Area (3.3) for x=1 and seller

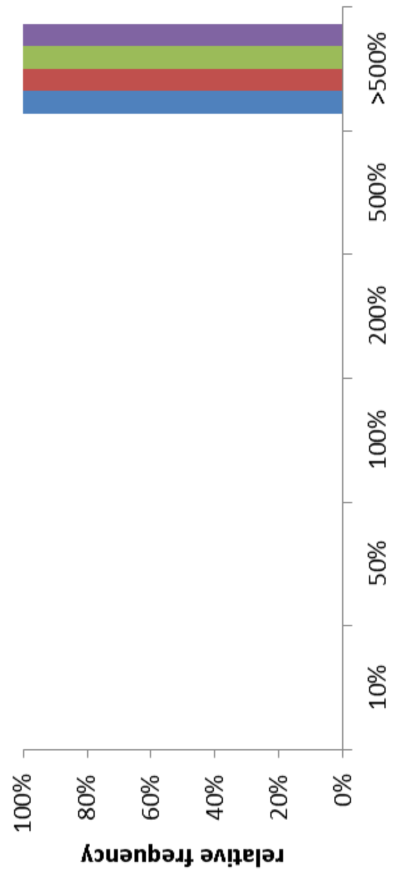


Figure Appendix 5.2.1.2.1f: Area (3.3) for x=2 and seller

■ OTPS ■ OTPS_L1 ■ OTPS_L2 ■ OTPS_LInf

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_L1) and seller (OTPS_L1)
L_2-norm constraint and value for buyer (OTPB_L2) and seller (OTPS_L2)
L_∞-norm constraint and value for buyer (OTPB_LInf) and seller (OTPS_LInf)

– "Similarity" of price differences of different shapes of constraints on portfolio holdings broken down by factors

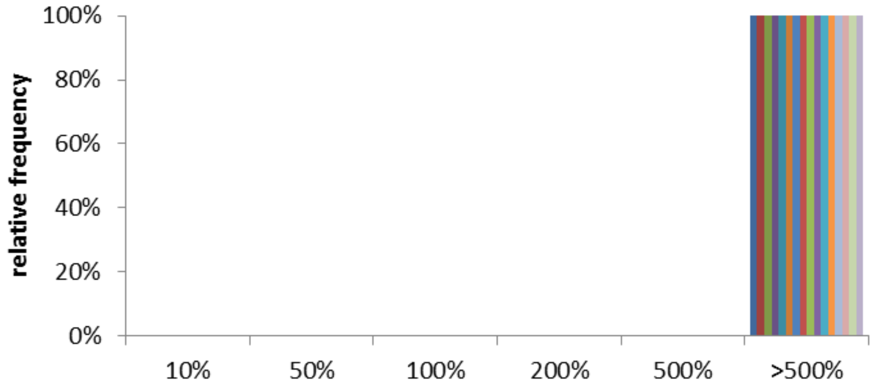


Figure Appendix 5.2.1.2.1g: Area (3.3) for x=0.5 and buyer broken down by factors

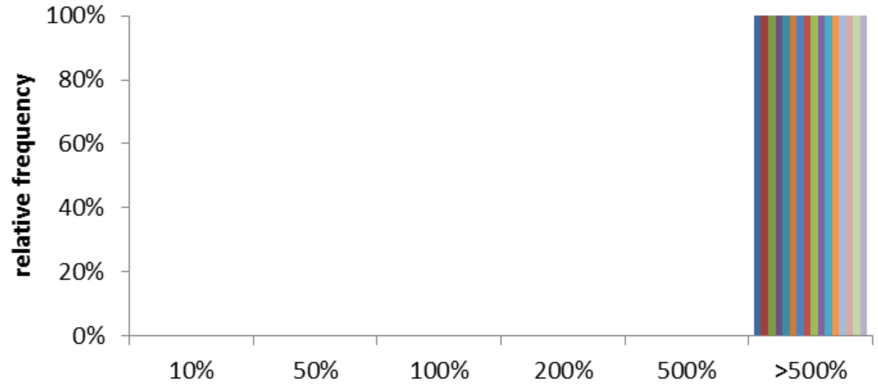


Figure Appendix 5.2.1.2.1h: Area (3.3) for x=1 and buyer broken down by factors

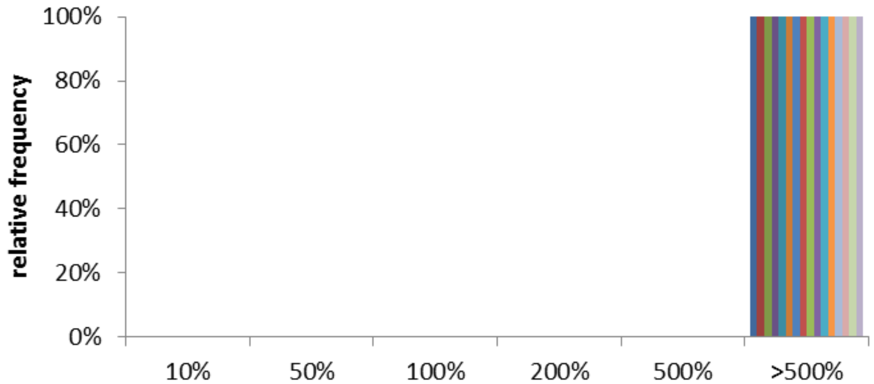


Figure Appendix 5.2.1.2.1i: Area (3.3) for x=2 and buyer broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

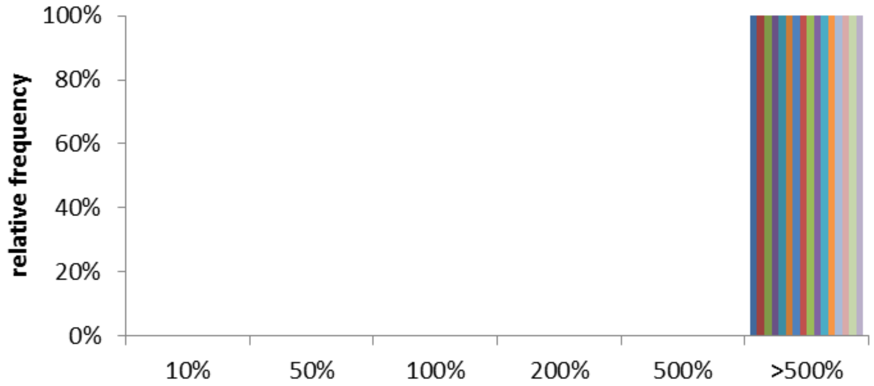


Figure Appendix 5.2.1.2.1j: Area (3.3) for x=0.5 and seller broken down by factors

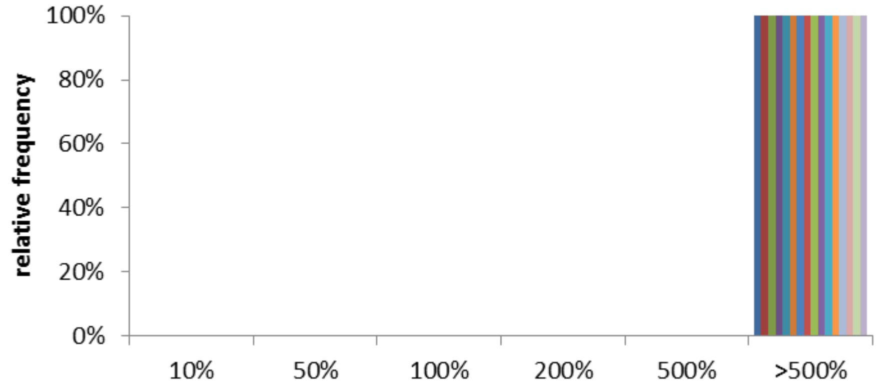


Figure Appendix 5.2.1.2.1k: Area (3.3) for x=1 and seller broken down by factors

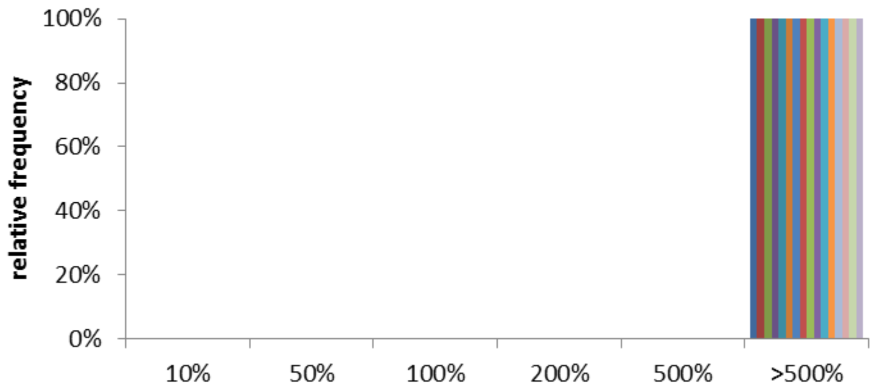


Figure Appendix 5.2.1.2.1l: Area (3.3) for x=2 and seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

Appendix 5.2.1.2.2 The Role of Size

– “Similarity” of price differences of different sizes of constraints on portfolio holdings

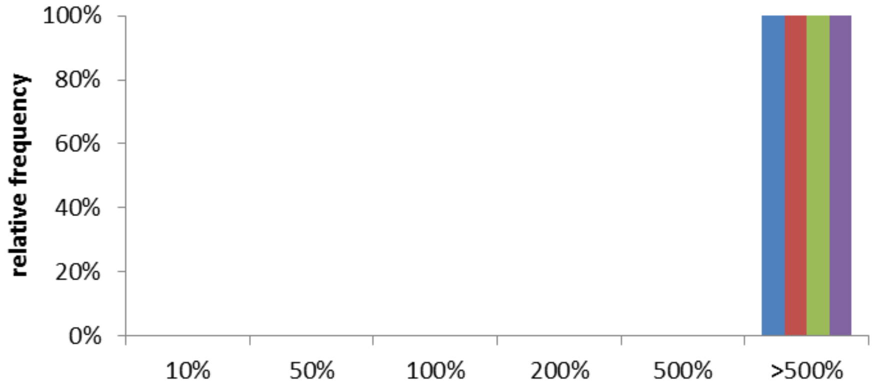


Figure Appendix 5.2.2.1.2a: Area (3.3) for L_1 and buyer

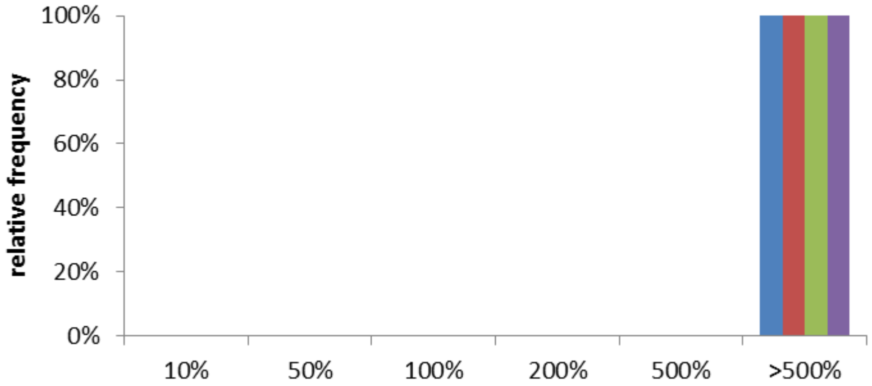


Figure Appendix 5.2.1.2.2b: Area (3.3) for L_2 and buyer

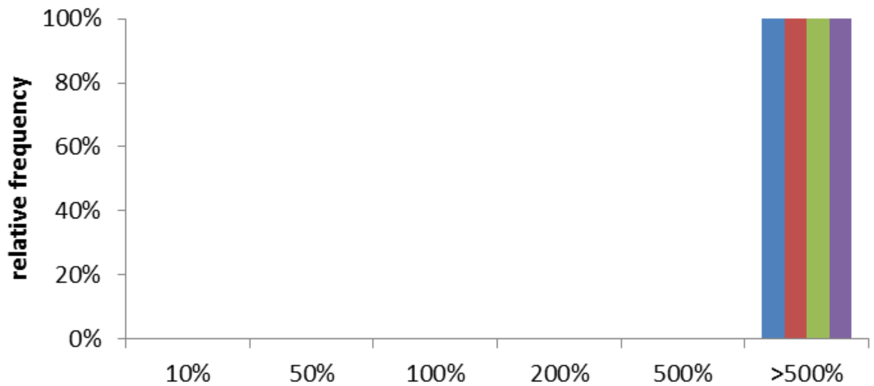


Figure Appendix 5.2.1.2.2c: Area (3.3) for L_∞ and buyer

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_x0.5) and seller (OTPS_x0.5)
L_2-norm constraint and value for buyer (OTPB_x1) and seller (OTPS_x1)
L_∞-norm constraint and value for buyer (OTPB_x2) and seller (OTPS_x2)

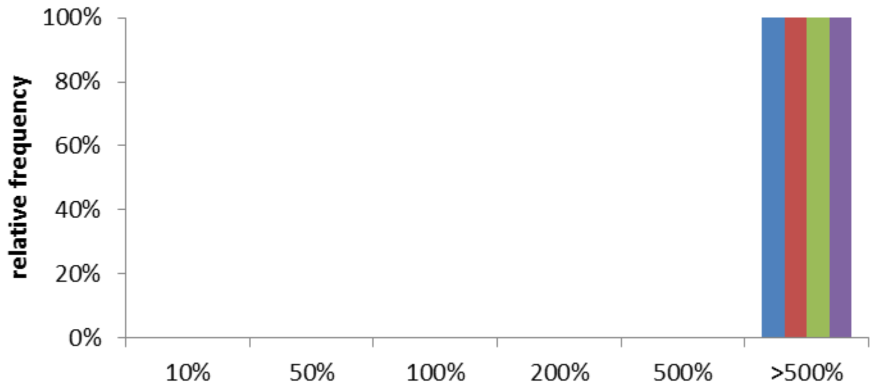


Figure Appendix 5.2.1.2.2d: Area (3.3) for L_1 and seller

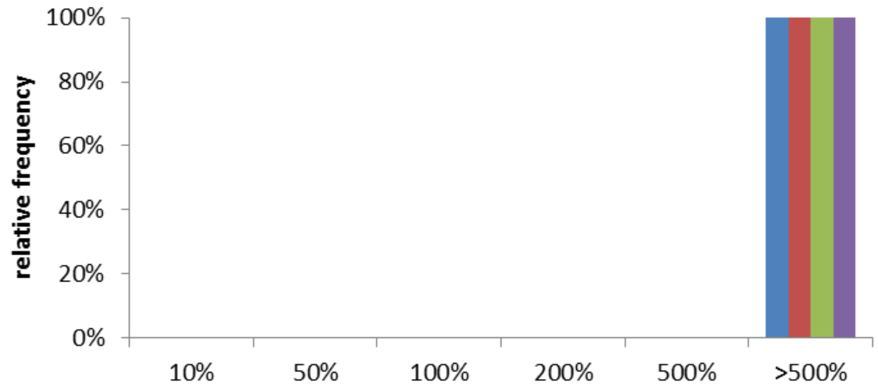


Figure Appendix 5.2.1.2.2e: Area (3.3) for L_2 and seller

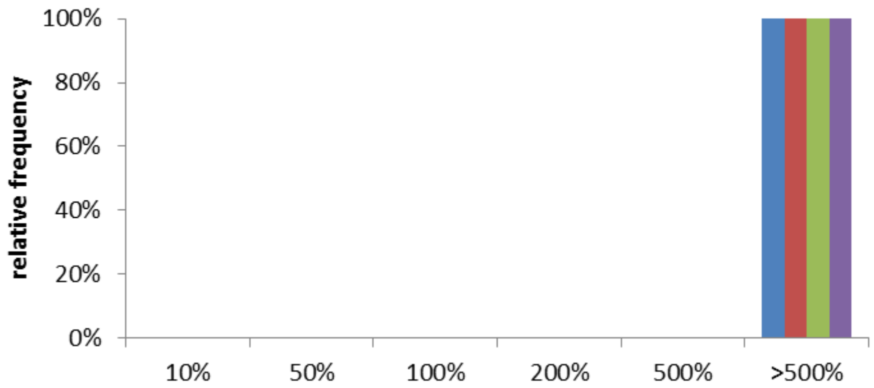


Figure Appendix 5.2.1.2.2f: Area (3.3) for L_∞ and seller

OTPS OTPS_x0.5 OTPS_x1 OTPS_x2

Non-negativity constraint and value for buyer (OTPB) and seller (OTPS)
L_1-norm constraint and value for buyer (OTPB_x0.5) and seller (OTPS_x0.5)
L_2-norm constraint and value for buyer (OTPB_x1) and seller (OTPS_x1)
L_∞-norm constraint and value for buyer (OTPB_x2) and seller (OTPS_x2)

– "Similarity" of price differences of different sizes of constraints on portfolio holdings broken down by factors

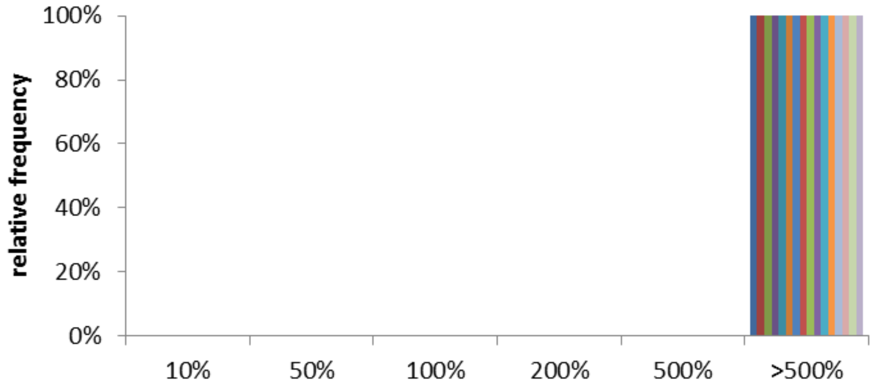


Figure Appendix 5.2.1.2.2g: Area (3.3) for L_1 and buyer broken down by factors

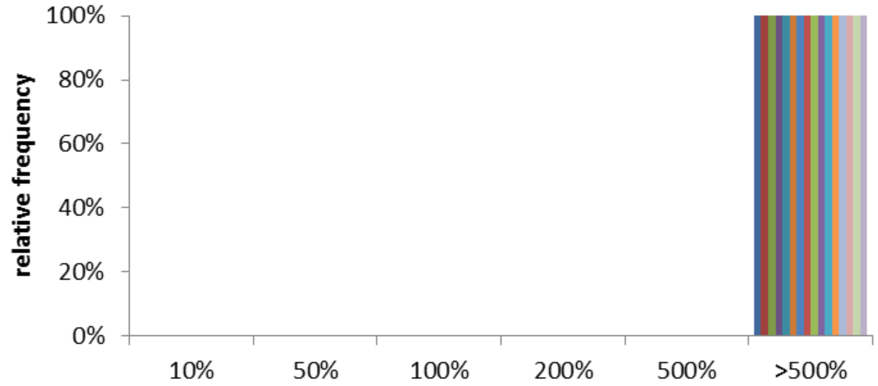


Figure Appendix 5.2.1.2.2h: Area (3.3) for L_2 and buyer broken down by factors

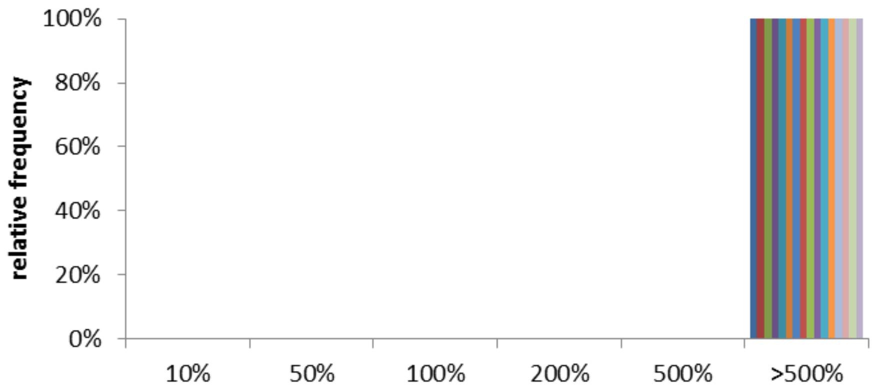


Figure Appendix 5.2.1.2.2i: Area (3.3) for L_∞ and buyer broken down by factors

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

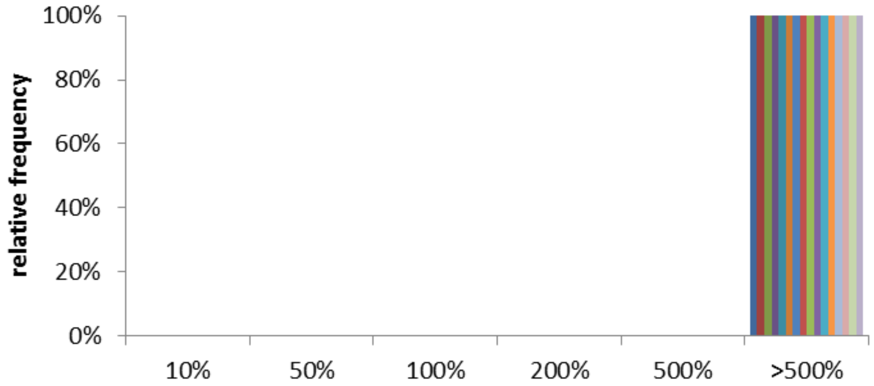


Figure Appendix 5.2.1.2.2j: Area (3.3) for L_1 and seller broken down by factors

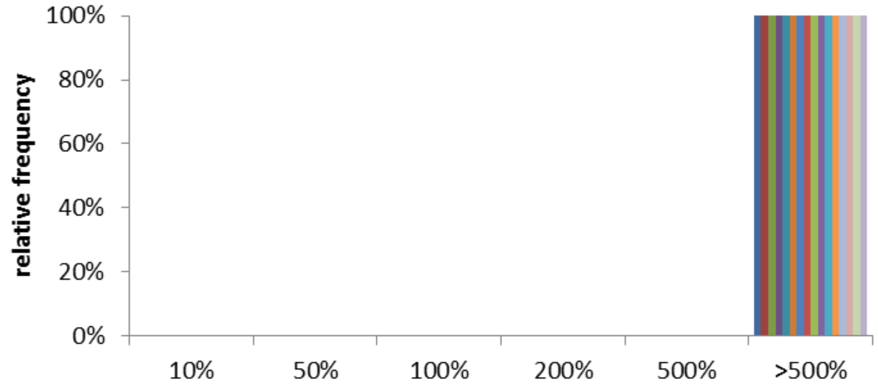


Figure Appendix 5.2.1.2.2k: Area (3.3) for L_2 and seller broken down by factors

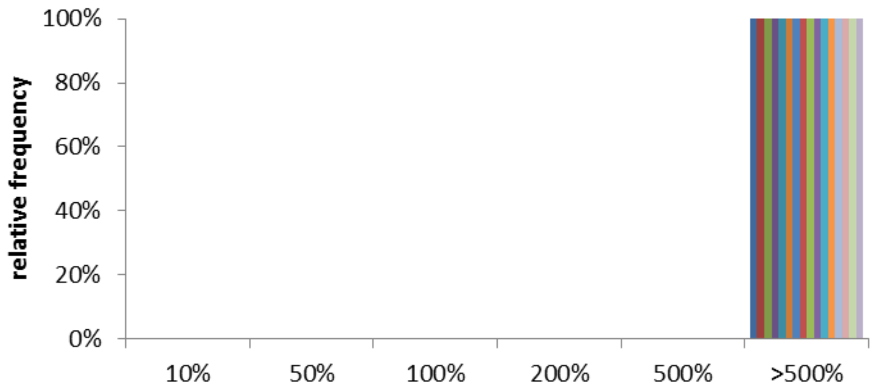


Figure Appendix 5.2.1.2.2l: Area (3.3) for L_∞ and seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

M1 to M11 are single-factor models (plus intercept)
M12 to M14 are two-factor models (plus intercept)
M15 is a three-factor model (plus intercept)
M16 is an eleven-factor model (plus intercept)

Appendix 5.2.2 Multi-period Version of the Optimize-the-Price-Approach (Section 3.3.3)

Appendix 5.2.2.1 Time Trend on Accounting Characteristics

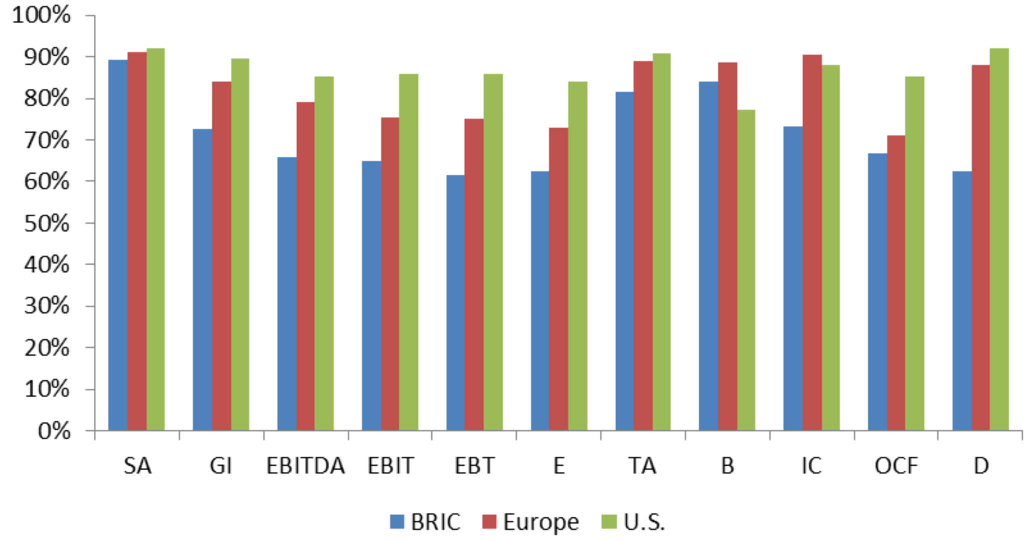


Figure Appendix 5.2.2.1: Time trend for selected accounting characteristics between 2010 and 2014 broken down by region with SA Net Sales or Revenues, GI Gross Income, EBITDA Earnings Before Interest, Taxes and Depreciation, EBIT Earnings Before Interest and Taxes, EBT Earnings Before Taxes, E Earnings, TA Total Assets, B Book Value of Common Equity, IC Invested Capital, OCF Operating Cash Flow, D Ordinary Cash Dividends

Appendix 5.2.2.2 “Magnitude”

– Buyer

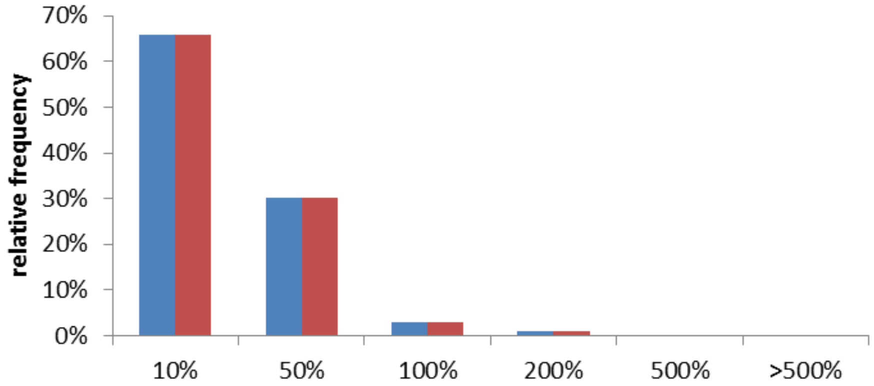


Figure Appendix 5.2.2.2a: |Ratio (3.1)| for the buyer

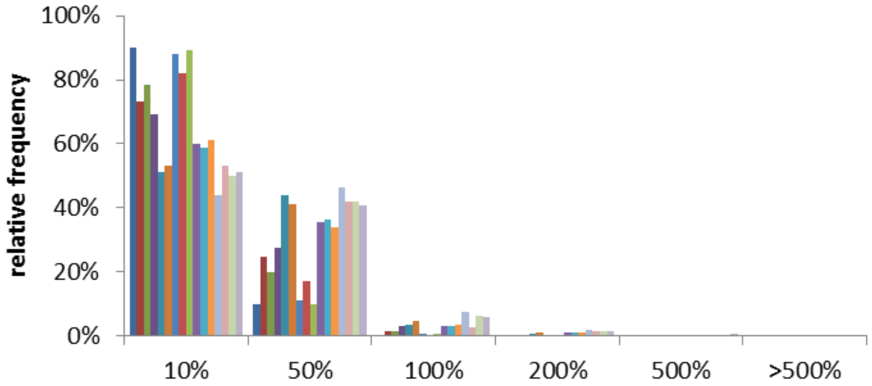


Figure Appendix 5.2.2.2b: |Ratio (3.1)| for the buyer broken down by factors

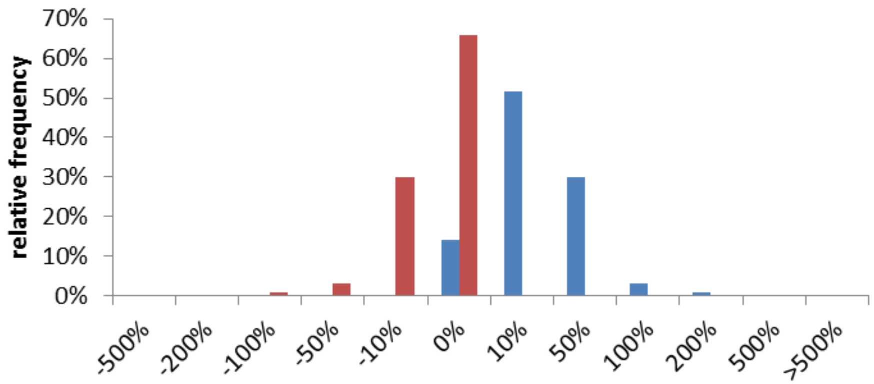


Figure Appendix 5.2.2.2c: Ratio (3.1) for the buyer

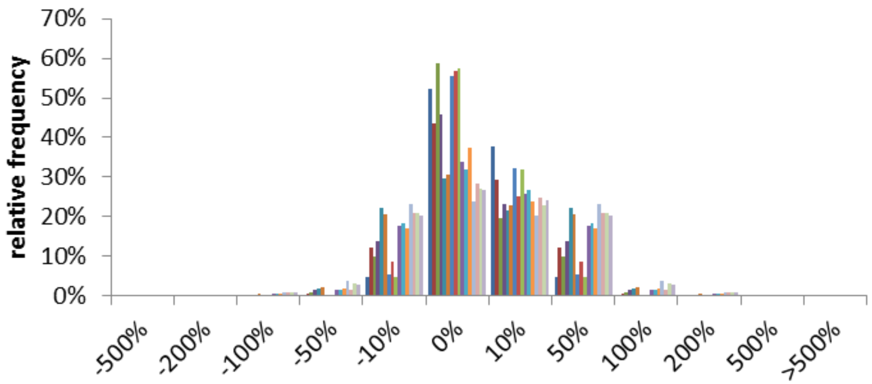


Figure Appendix 5.2.2.2d: Ratio (3.1) for the buyer broken down by factors

– **Seller**

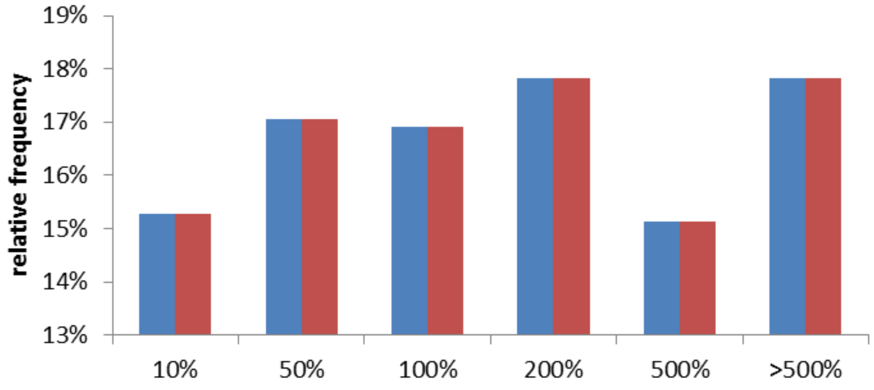


Figure Appendix 5.2.2.2e: |Ratio (3.1)| for the seller

OTPS OTPSmulti-period

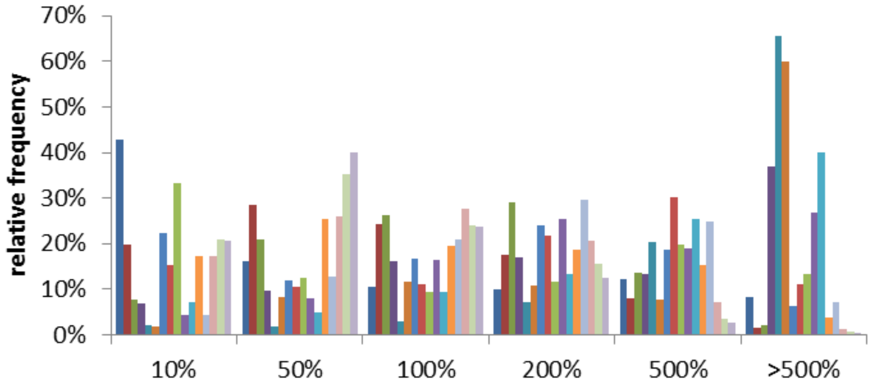


Figure Appendix 5.2.2.2f: |Ratio (3.1)| for the seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

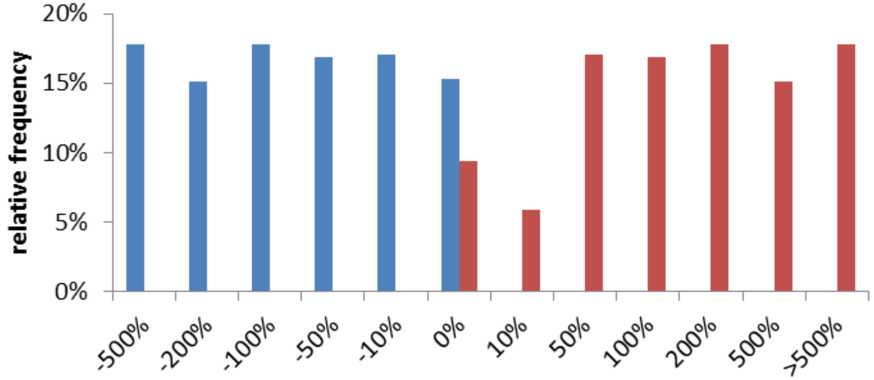


Figure Appendix 5.2.2.2g: Ratio (3.1) for the seller

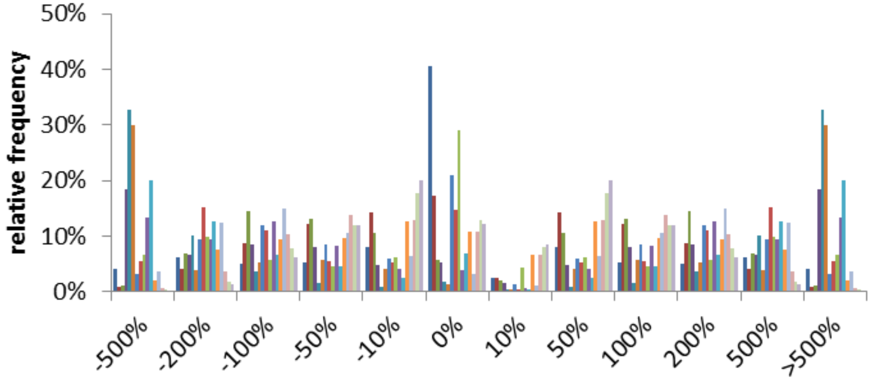


Figure Appendix 5.2.2.2h: Ratio (3.1) for the seller broken down by factors

Appendix 5.2.2.3 “Similarity”

– **Buyer**

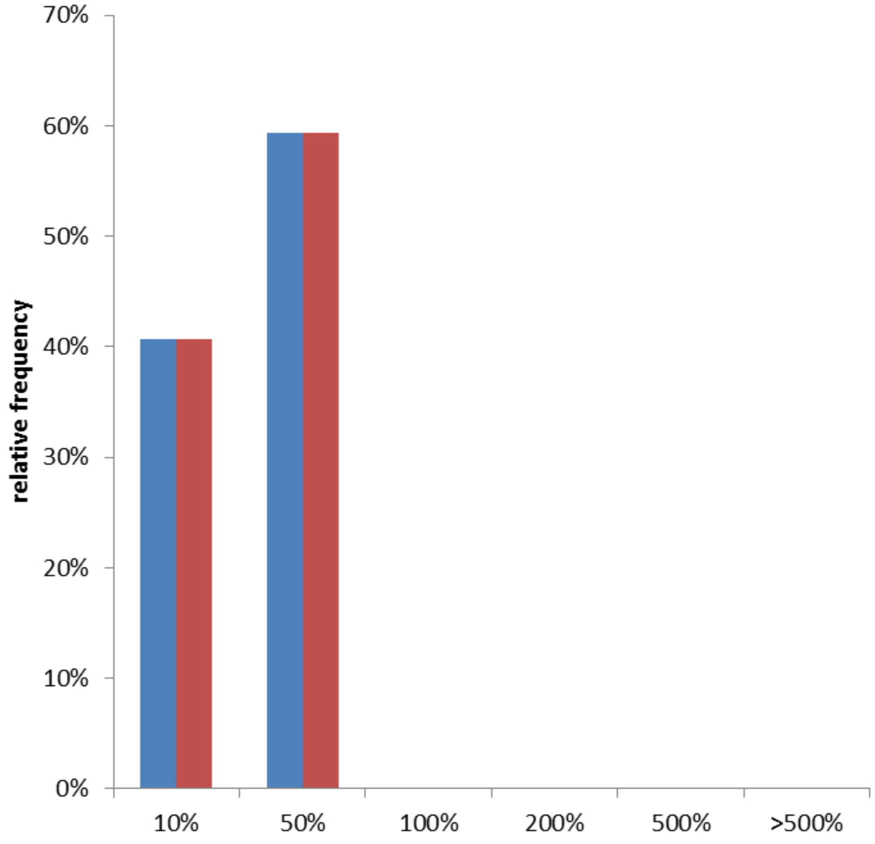


Figure Appendix 5.2.2.3a: Area (3.3) for the buyer

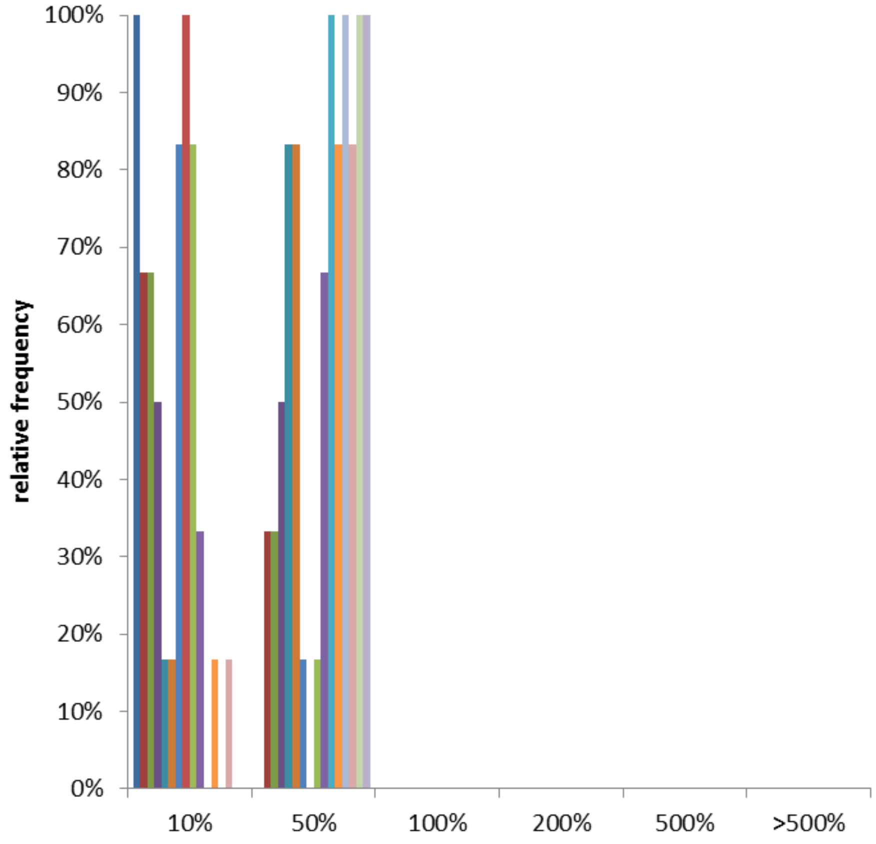


Figure Appendix 5.2.2.3b: Area (3.3) for the buyer broken down by factors

– Seller

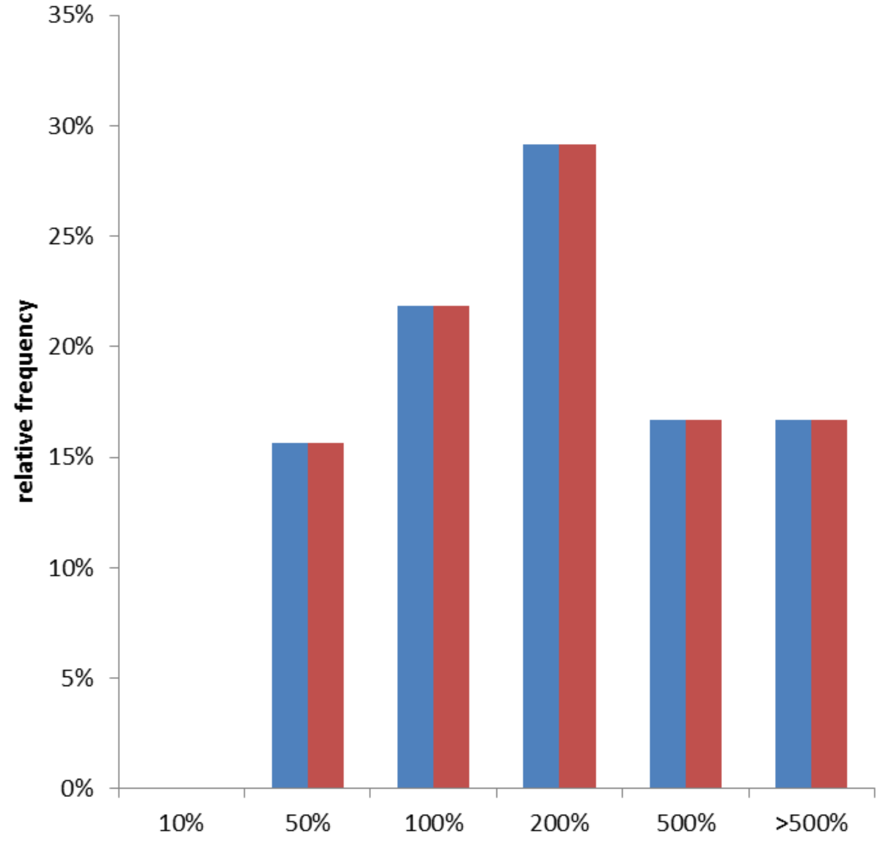


Figure Appendix 5.2.2.3c: Area (3.3) for the seller

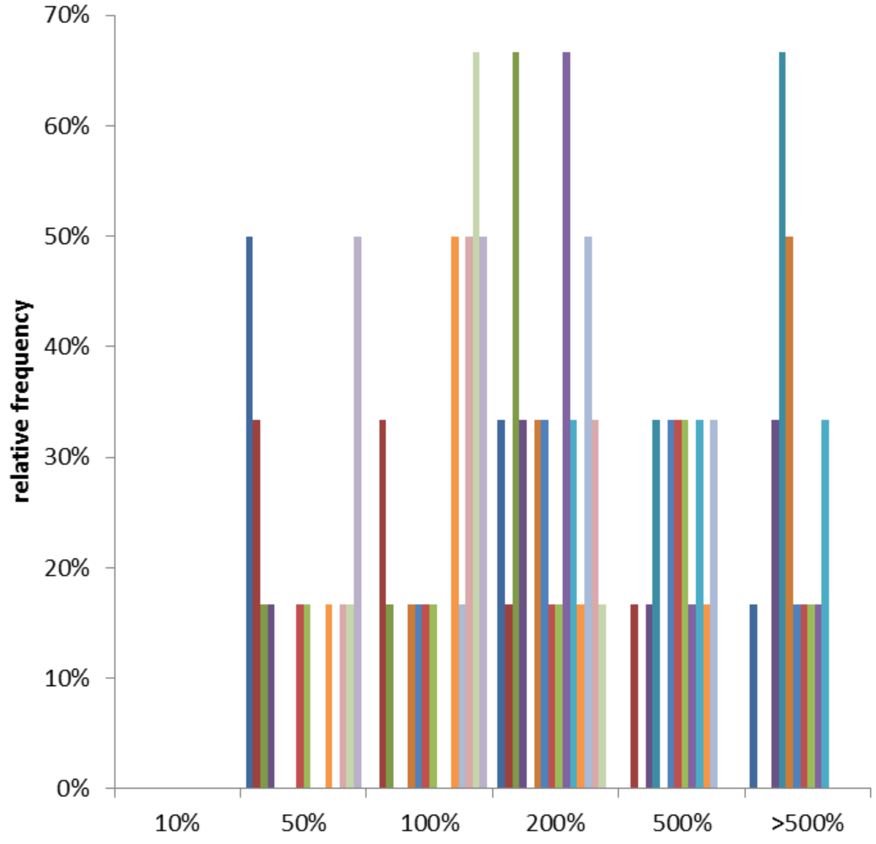


Figure Appendix 5.2.2.3d: Area (3.3) for the seller broken down by factors

Appendix 5.2.3 Optimize-the-Price-Approach versus Regressions (Section 3.3.4)

Appendix 5.2.3.1 Comparison of Regressions With and Without Constant β_0

Appendix 5.2.3.1.1 “Magnitude”

– OLS

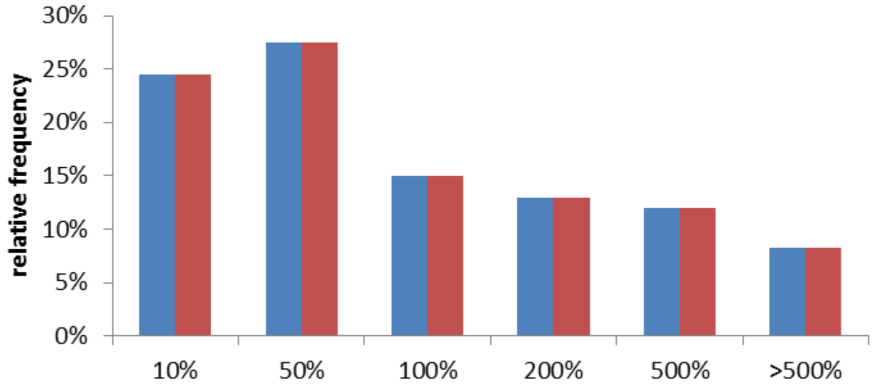


Figure Appendix 5.2.3.1.1a: |Ratio (3.1)|

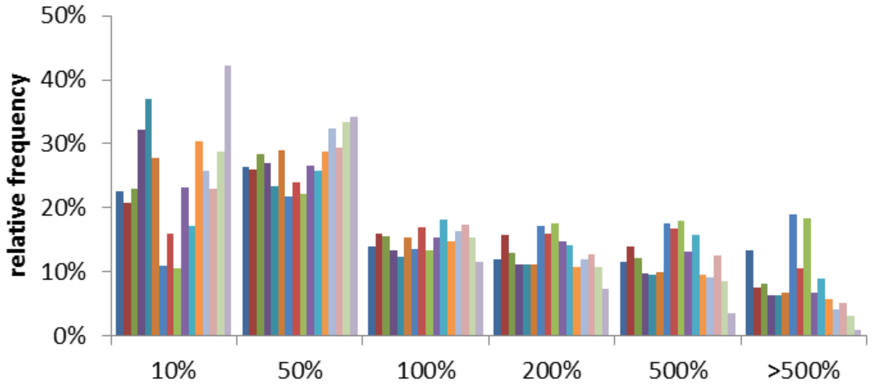


Figure Appendix 5.2.3.1.1b: |Ratio (3.1)| broken down by factors

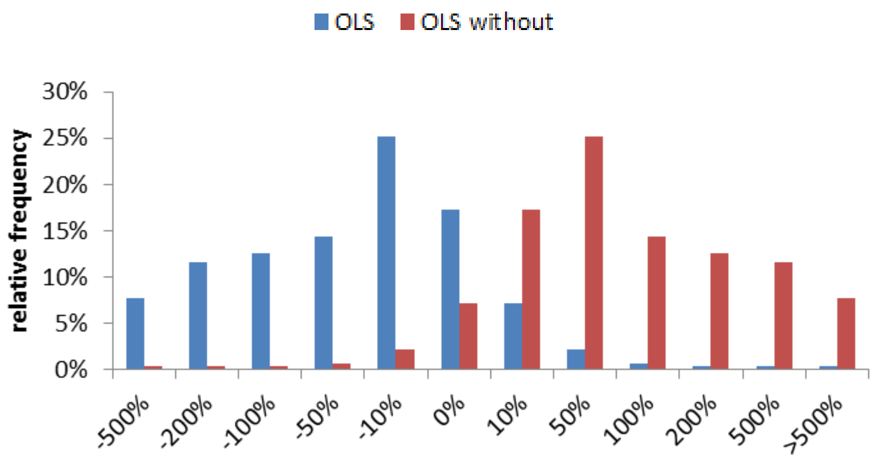


Figure Appendix 5.2.3.1.1c: Ratio (3.1)

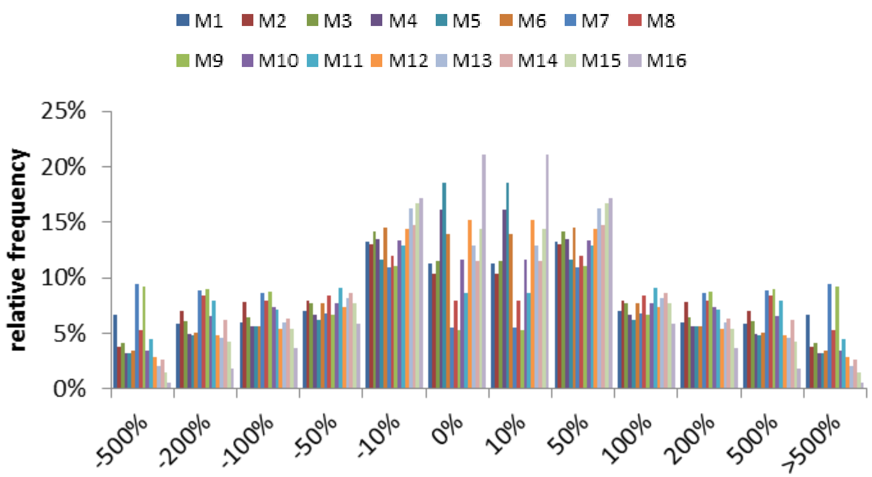


Figure Appendix 5.2.3.1.1d: Ratio (3.1) broken down by factors

– WLS

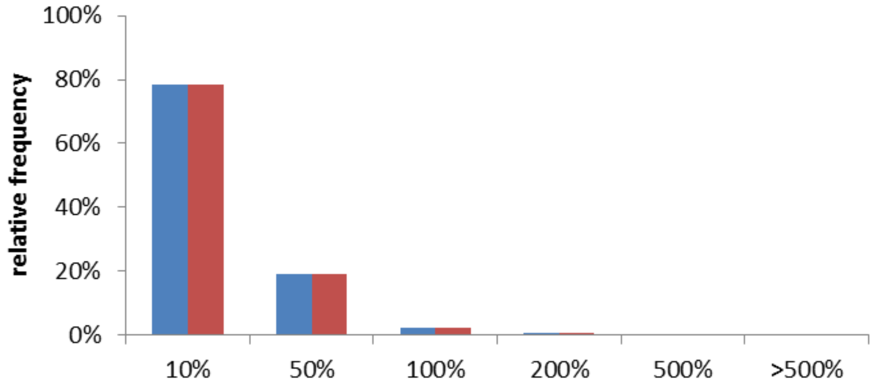


Figure Appendix 5.2.3.1.1e: |Ratio (3.1)|

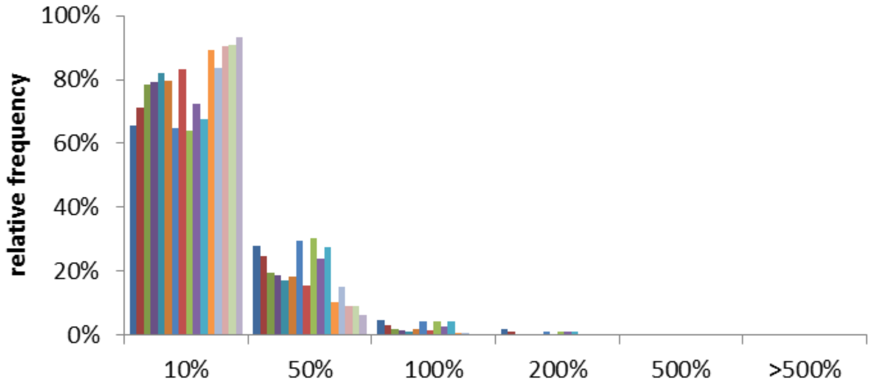


Figure Appendix 5.2.3.1.1f: |Ratio (3.1)| broken down by factors

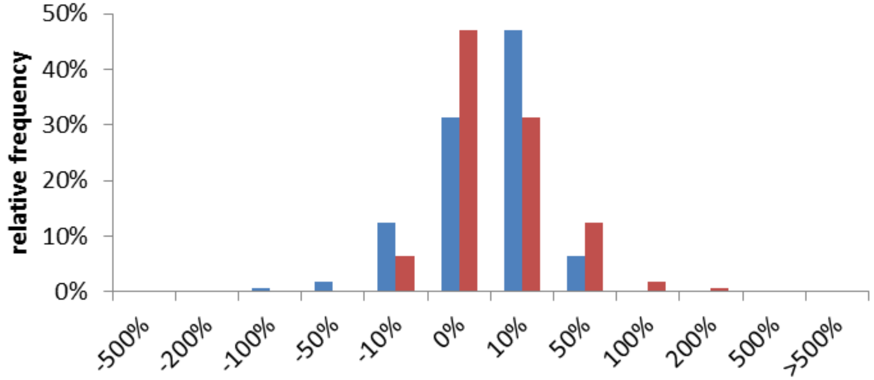


Figure Appendix 5.2.3.1.1g: Ratio (3.1)

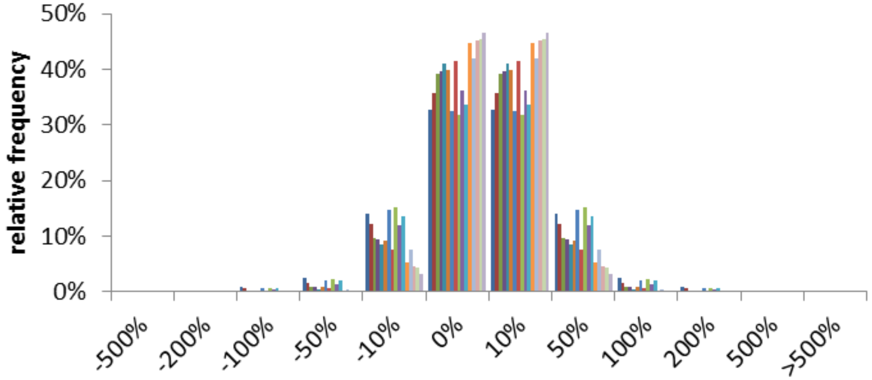


Figure Appendix 5.2.3.1.1h: Ratio (3.1) broken down by factors

– **Quantile (0.25)**

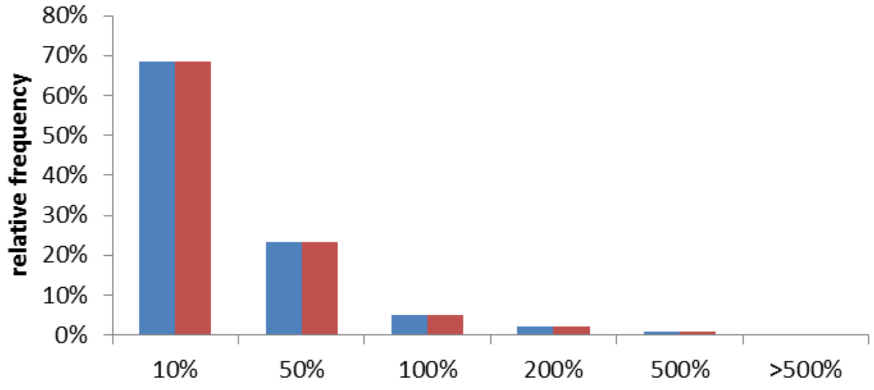


Figure Appendix 5.2.3.1.1i: |Ratio (3.1)|

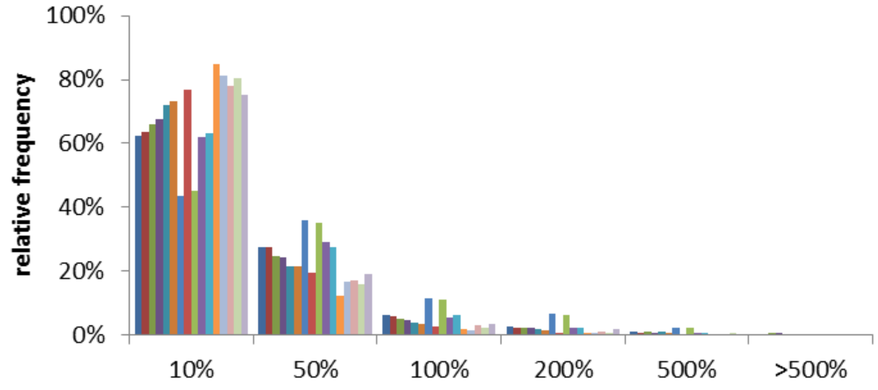


Figure Appendix 5.2.3.1.1j: |Ratio (3.1)| broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

Quantile(0.25) Quantile(0.25) without

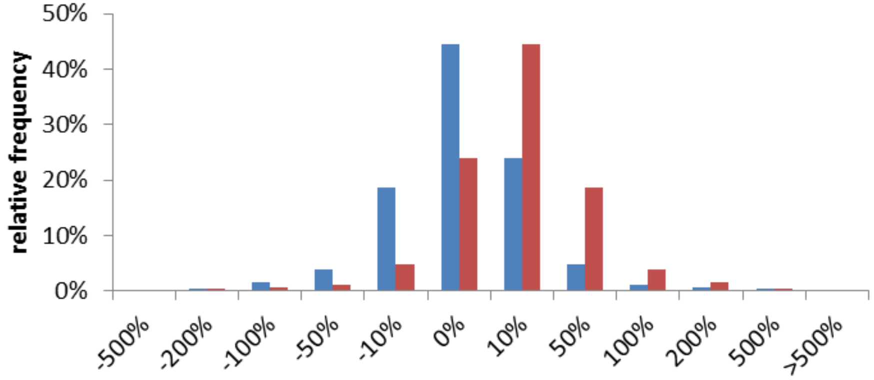


Figure Appendix 5.2.3.1.1k: Ratio (3.1)

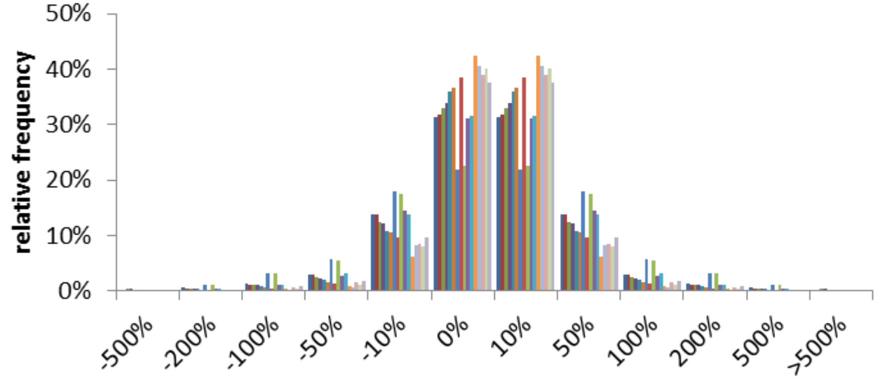


Figure Appendix 5.2.3.1.1l: Ratio (3.1) broken down by factors

– Quantile (0.50)

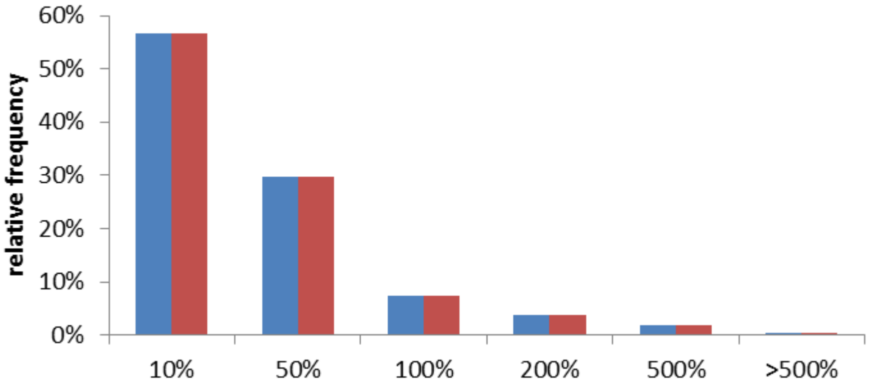


Figure Appendix 5.2.3.1.1m: |Ratio (3.1)|

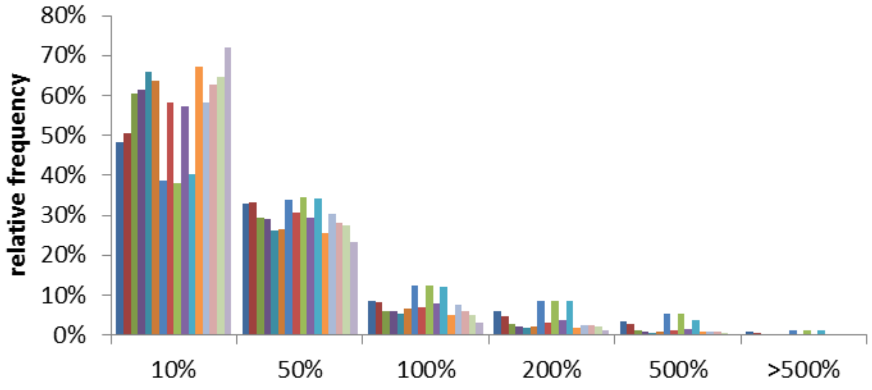


Figure Appendix 5.2.3.1.1n: |Ratio (3.1)| broken down by factors

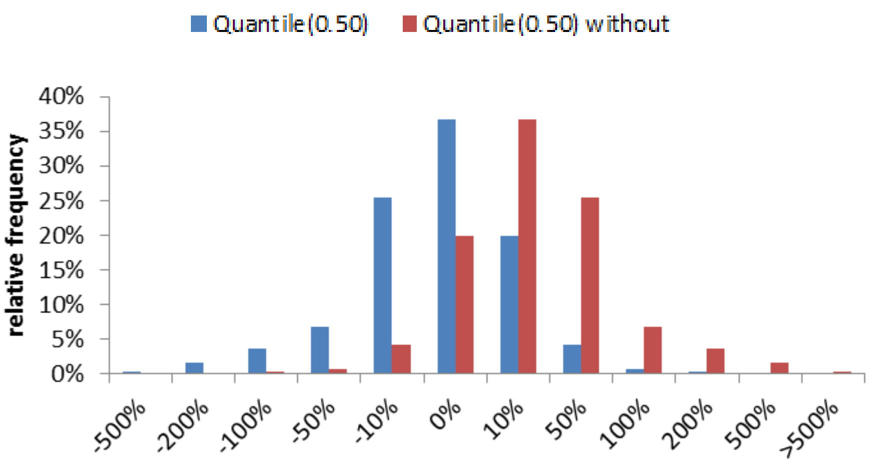


Figure Appendix 5.2.3.1.1o: Ratio (3.1)

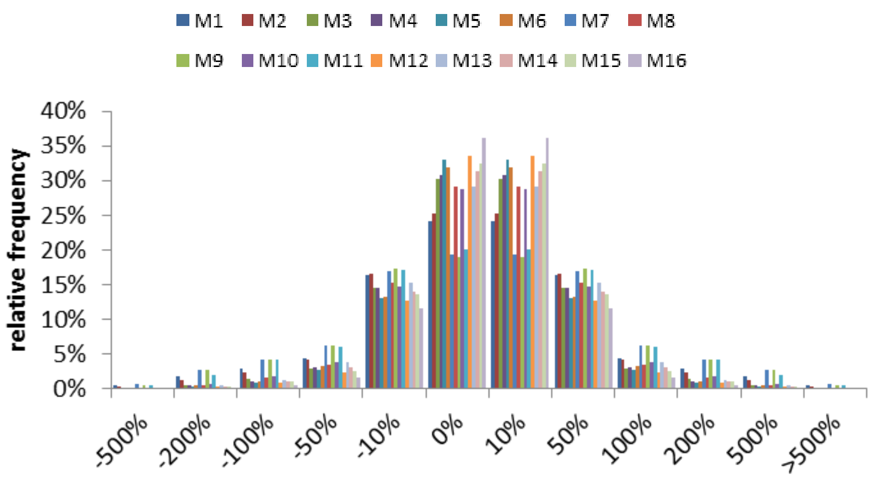


Figure Appendix 5.2.3.1.1p: Ratio (3.1) broken down by factors

– **Quantile (0.75)**

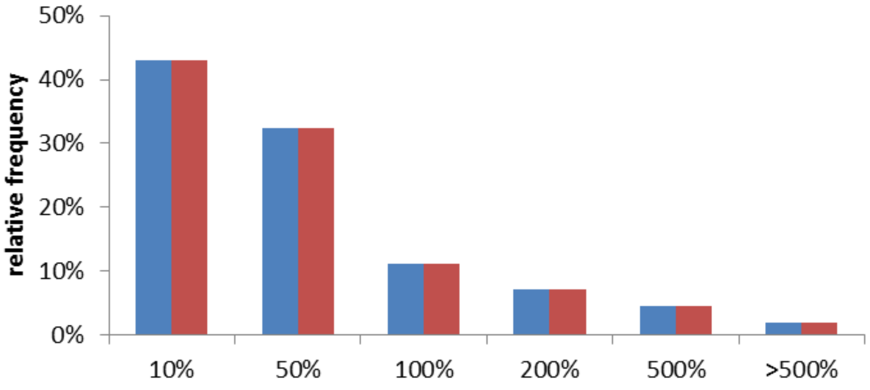


Figure Appendix 5.2.3.1.1q: |Ratio (3.1)|

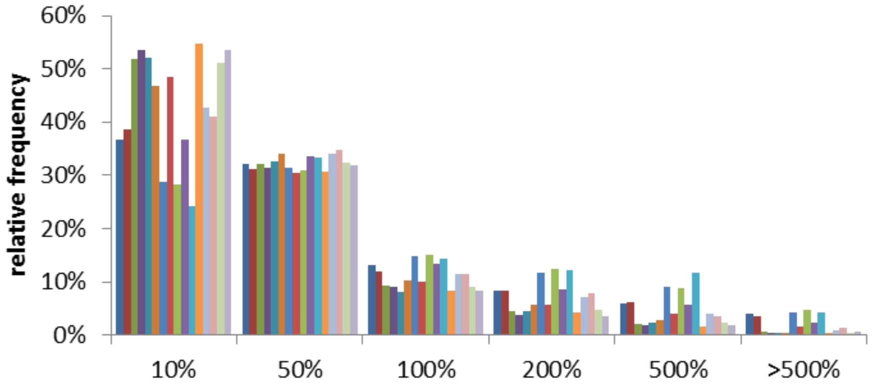


Figure Appendix 5.2.3.1.1r: |Ratio (3.1)| broken down by factors

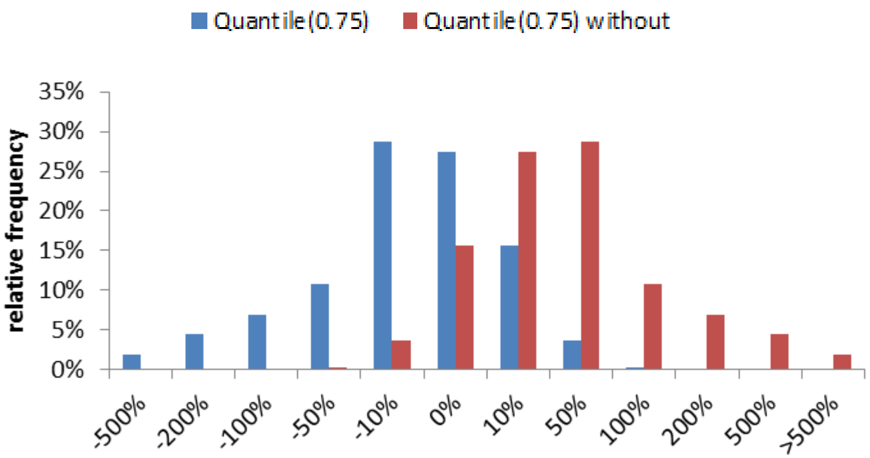


Figure Appendix 5.2.3.1.1s: Ratio (3.1)

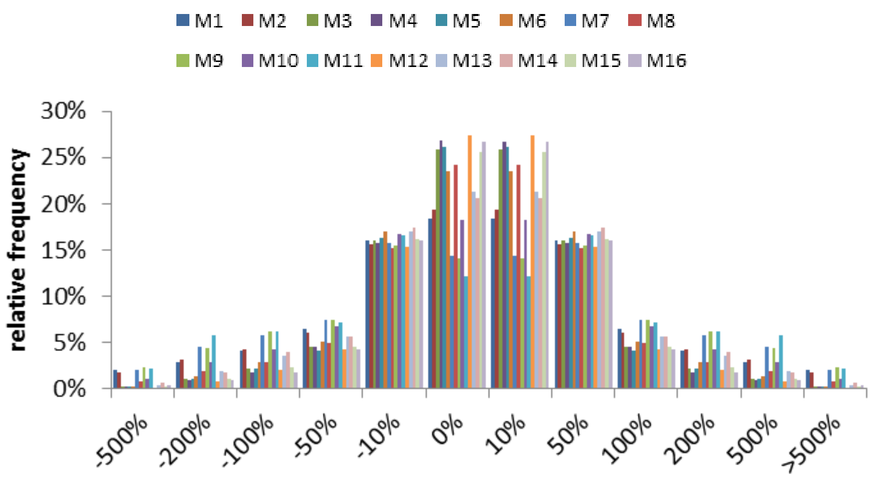


Figure Appendix 5.2.3.1.1t: Ratio (3.1) broken down by factors

Appendix 5.2.3.1.2 "Similarity"

– OLS

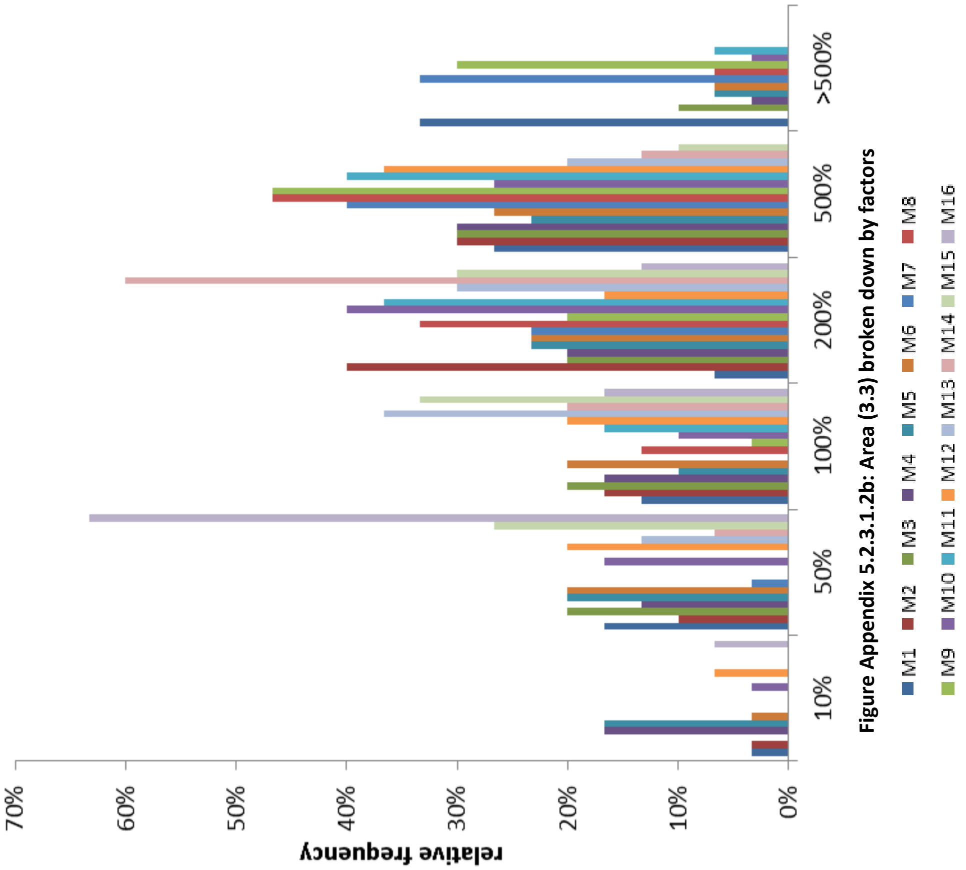


Figure Appendix 5.2.3.1.2b: Area (3.3) broken down by factors

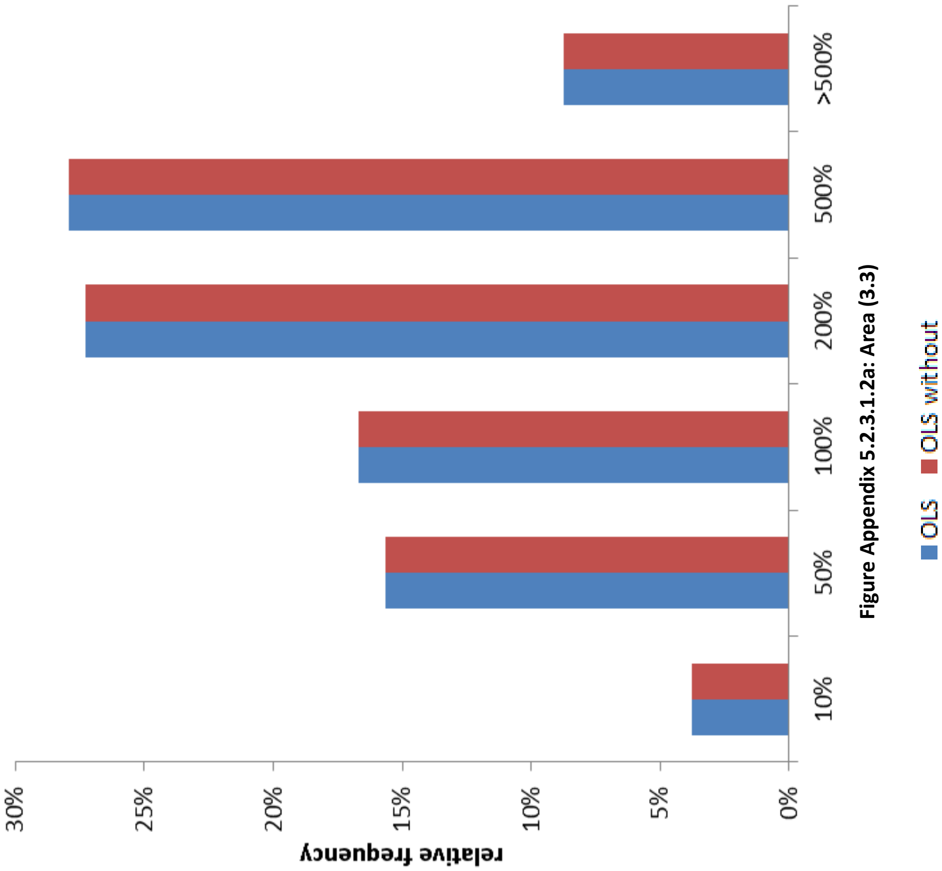


Figure Appendix 5.2.3.1.2a: Area (3.3)

– WLS

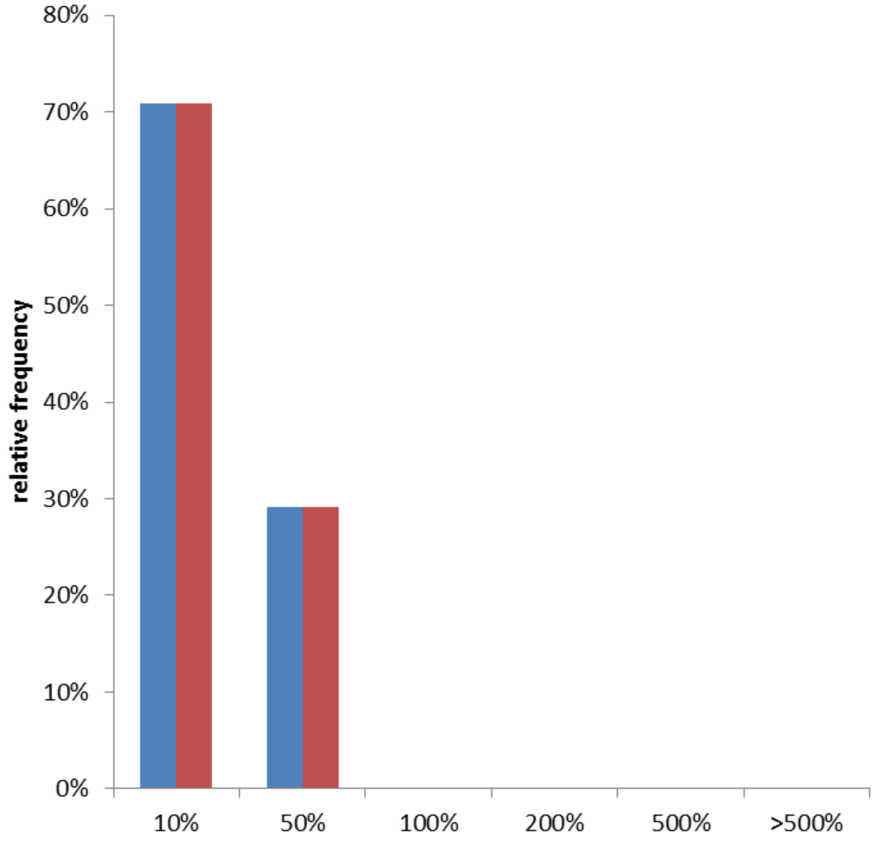


Figure Appendix 5.2.3.1.2c: Area (3.3)

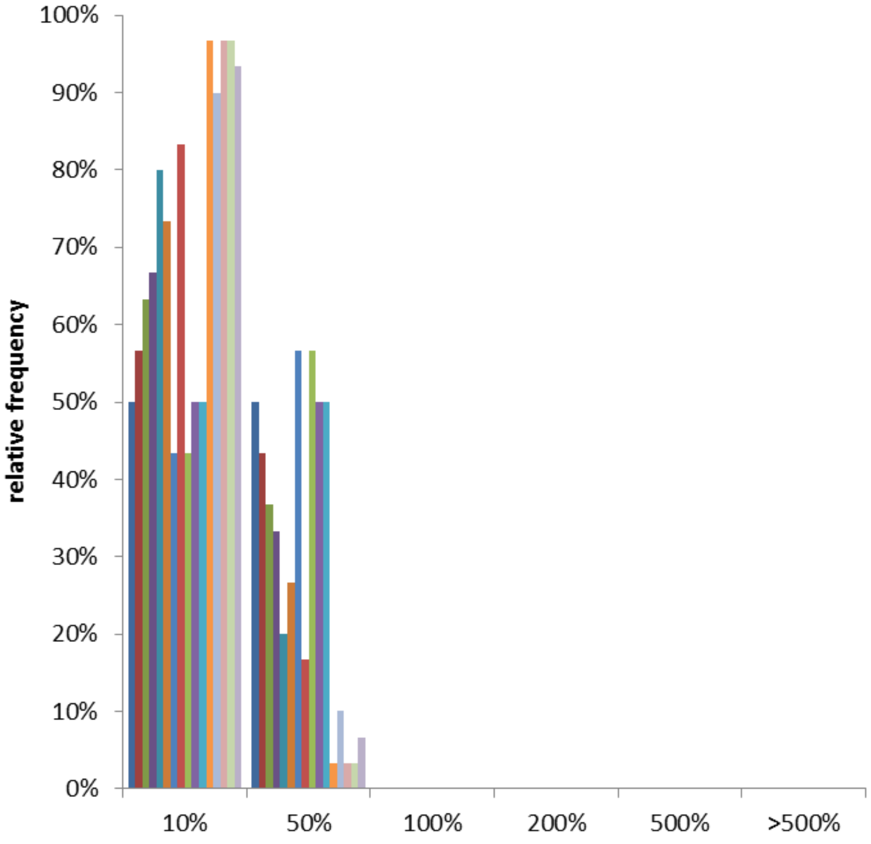


Figure Appendix 5.2.3.1.2d: Area (3.3) broken down by factors

– **Quantile (0.25)**

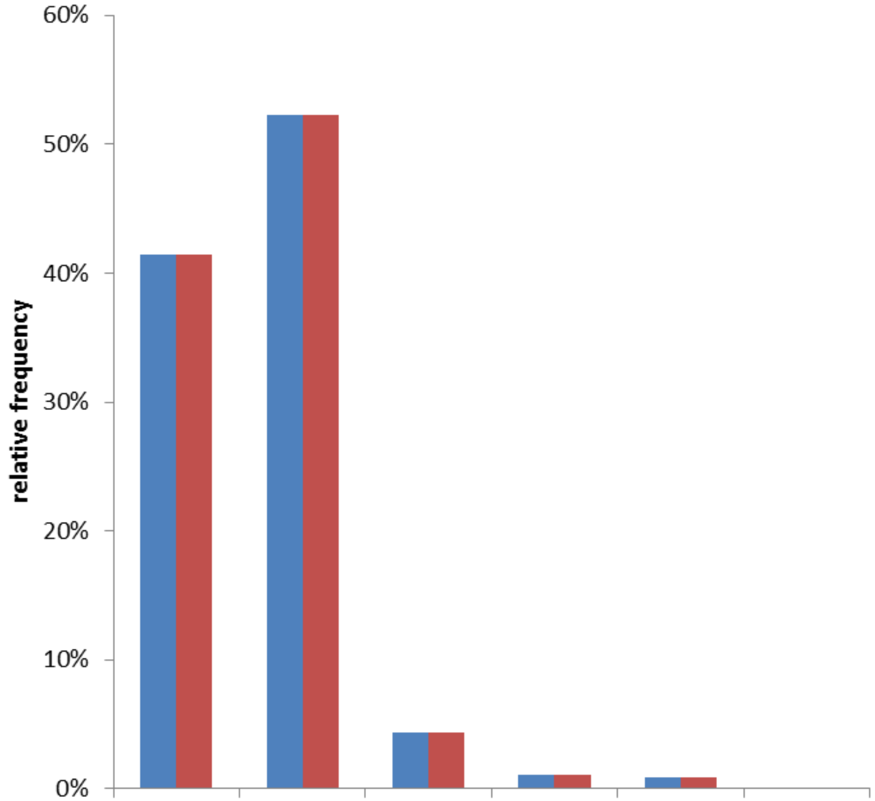


Figure Appendix 5.2.3.1.2e: Area (3.3)

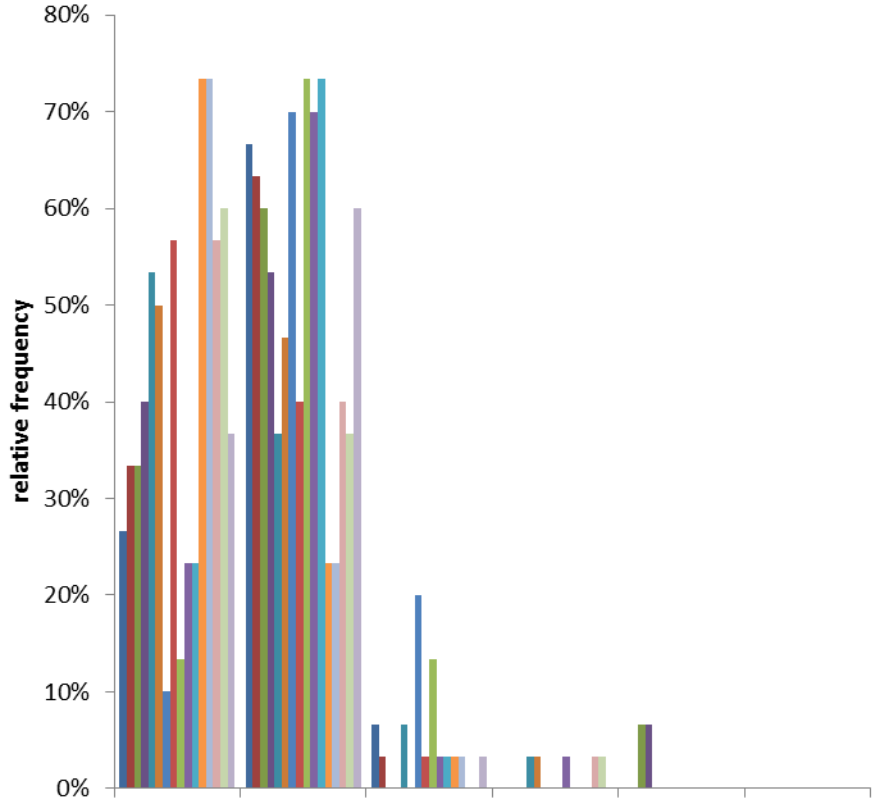


Figure Appendix 5.2.3.1.2f: Area (3.3) broken down by factors

– Quantile (0.50)

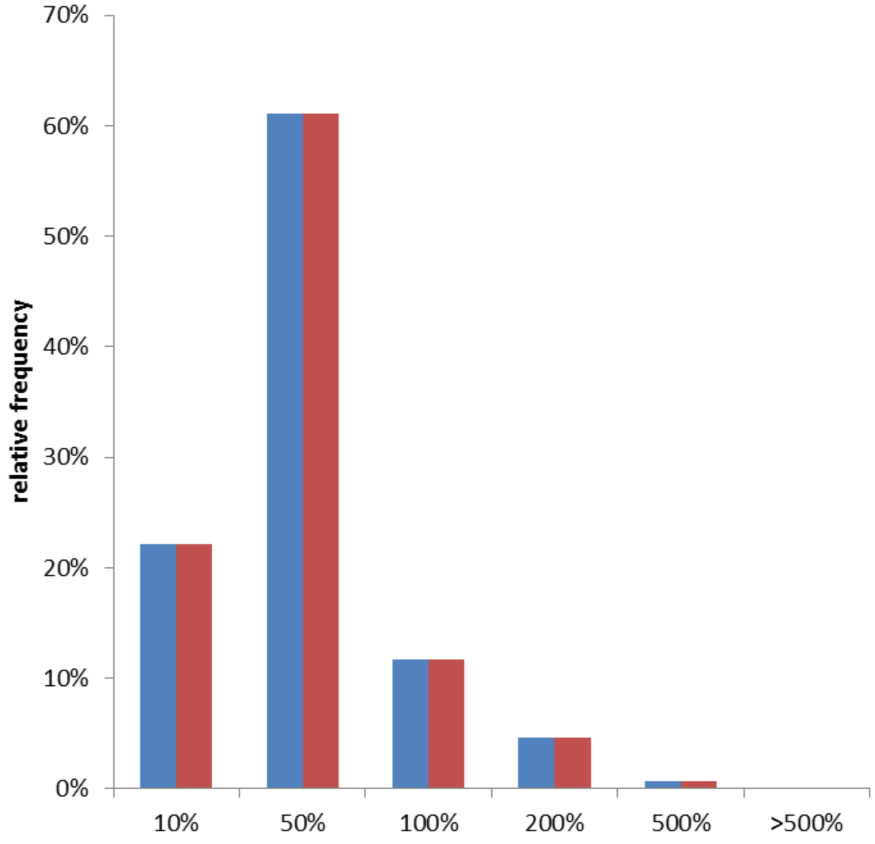


Figure Appendix 5.2.3.1.2g: Area (3.3)

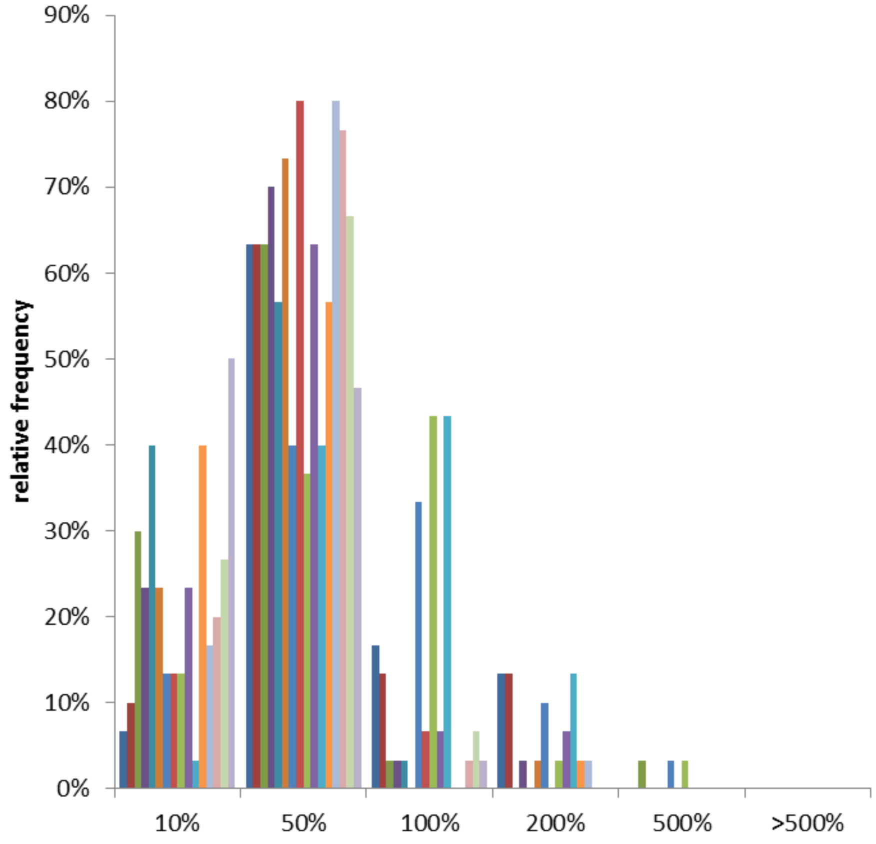


Figure Appendix 5.2.3.1.2h: Area (3.3) broken down by factors

– Quantile (0.75)

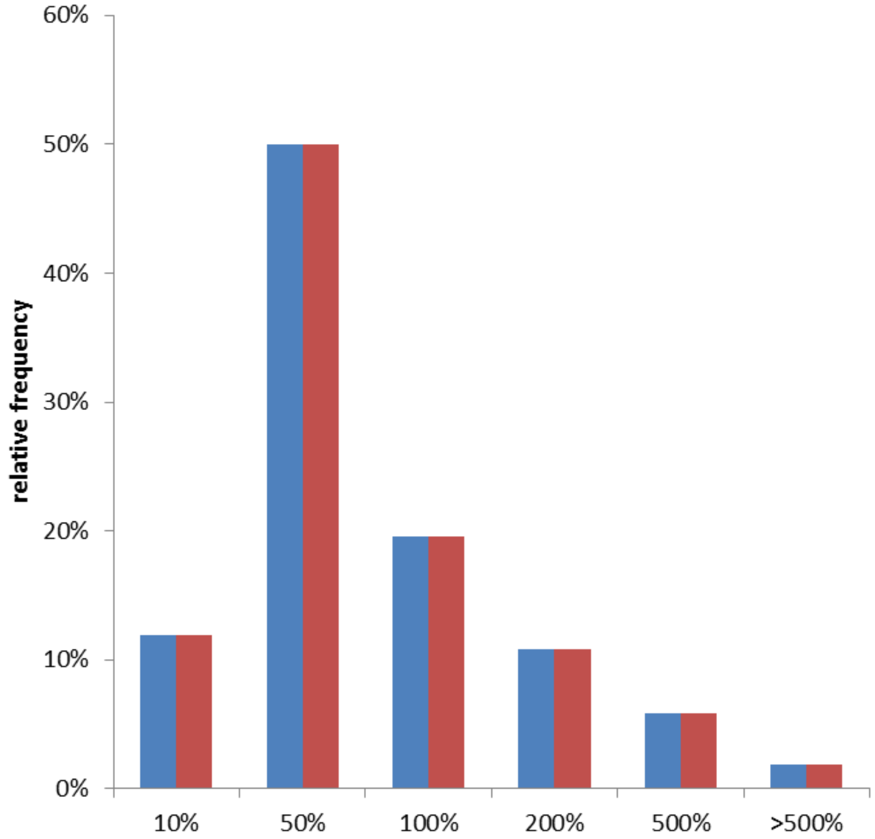


Figure Appendix 5.2.3.1.2i: Area (3.3)

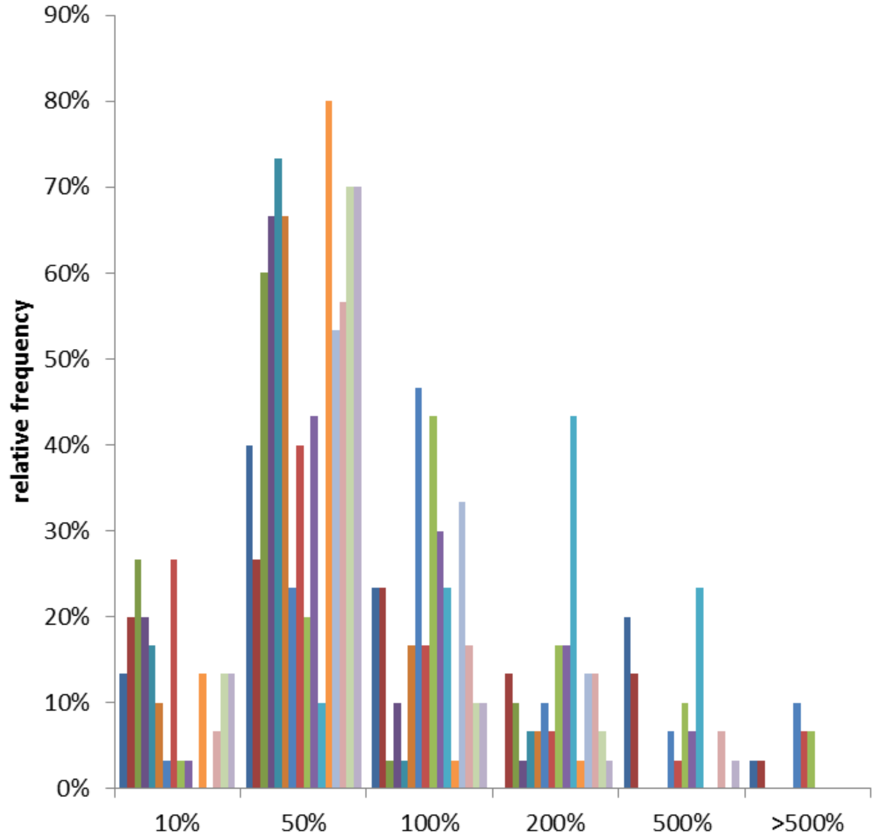


Figure Appendix 5.2.3.1.2j: Area (3.3) broken down by factors

Appendix 5.2.3.2 Optimize-the-Price-Approach versus Regressions With Constant $\boldsymbol{\beta_0}$

Appendix 5.2.3.2.1 “Magnitude”

– “Magnitude” of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Buyer

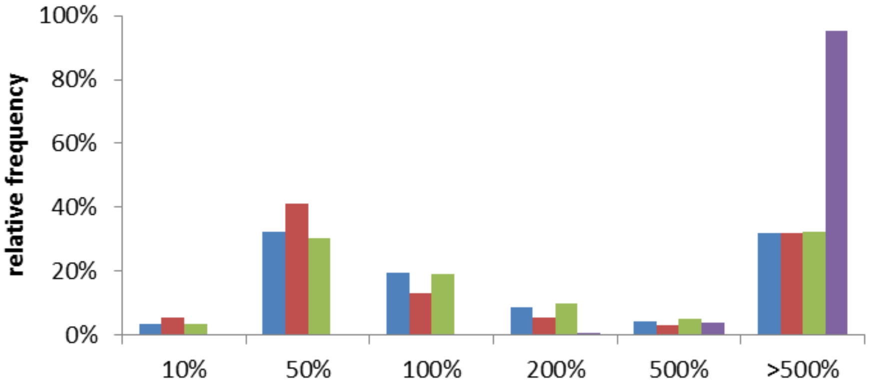


Figure Appendix 5.2.3.2.1a: |Ratio (3.1)| for the buyer

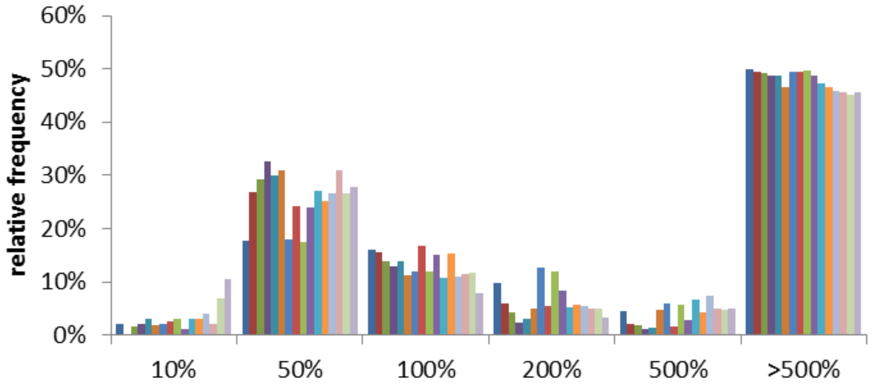


Figure Appendix 5.2.3.2.1b: |Ratio (3.1)| for the buyer broken down by factors

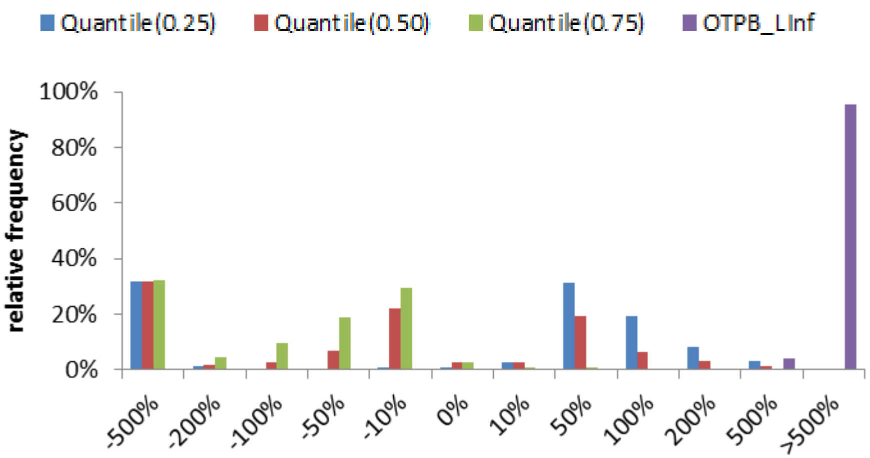


Figure Appendix 5.2.3.2.1c: Ratio (3.1) for the buyer

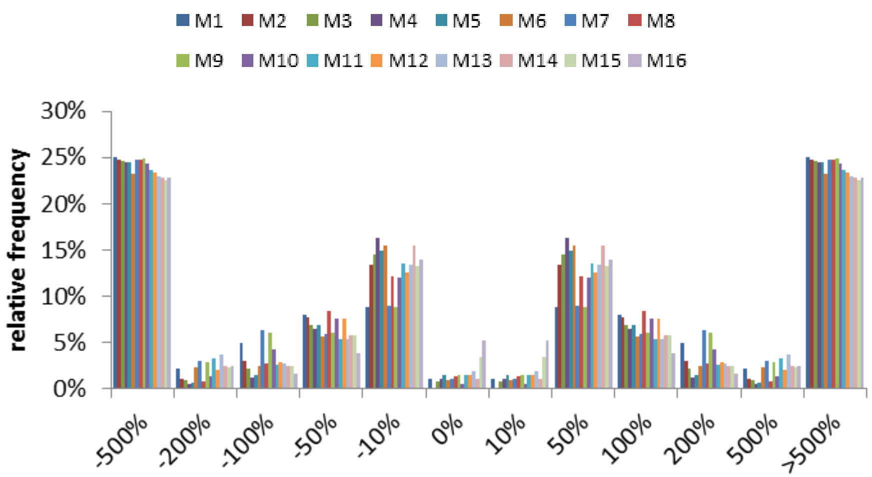


Figure Appendix 5.2.3.2.1d: Ratio (3.1) for the buyer broken down by factors

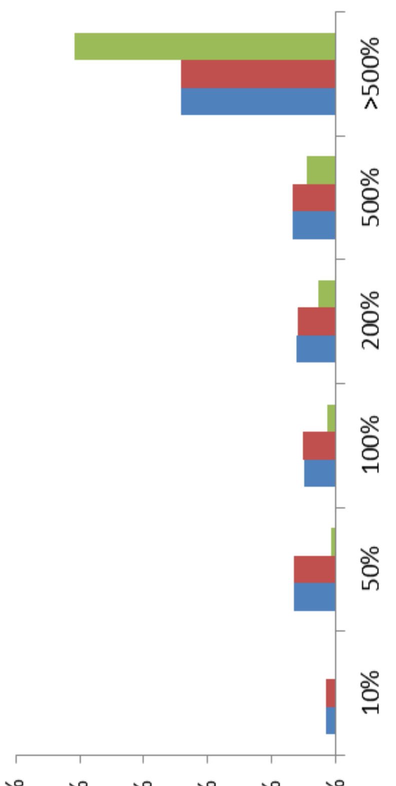


Figure Appendix 5.2.3.2.1e: |Ratio (3.1)| for the buyer

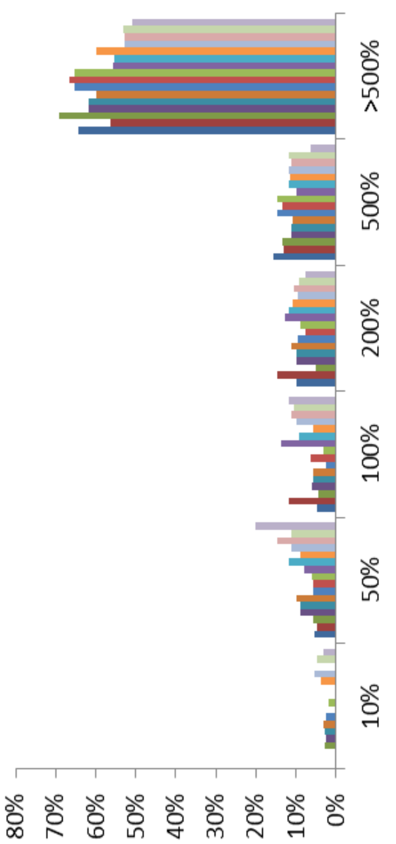


Figure Appendix 5.2.3.2.1f: |Ratio (3.1)| for the buyer broken down by factors

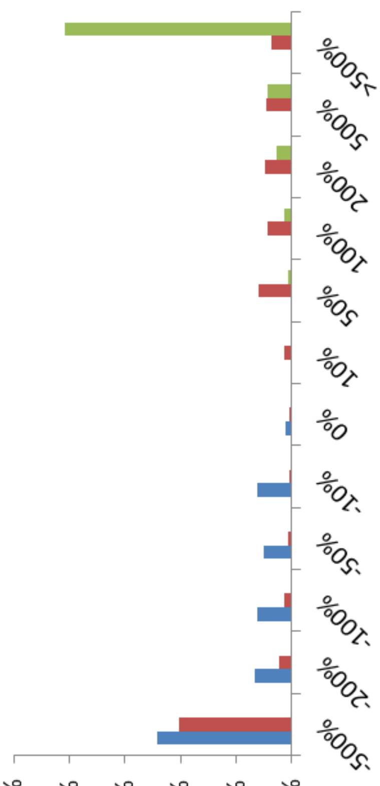


Figure Appendix 5.2.3.2.1g: Ratio (3.1) for the buyer

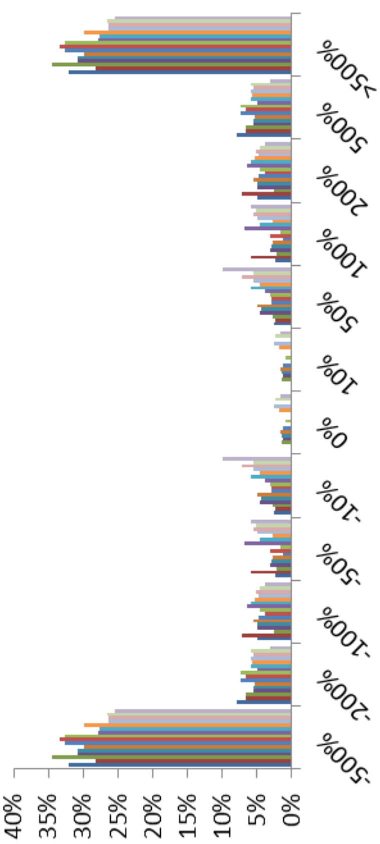


Figure Appendix 5.2.3.2.1h: Ratio (3.1) for the buyer broken down by factors

– “Magnitude” of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Seller

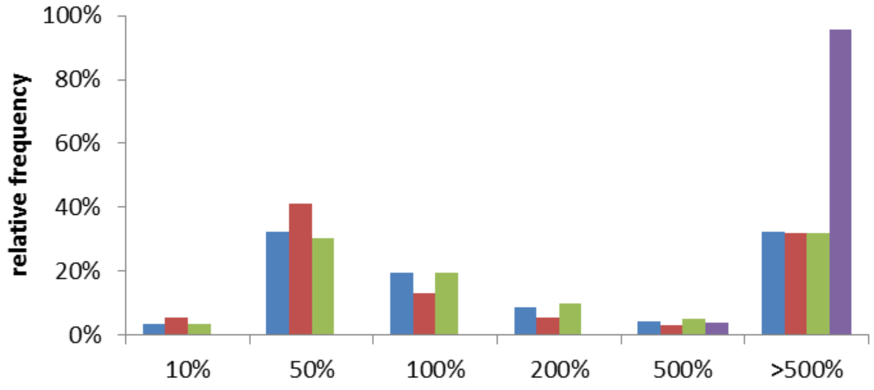


Figure Appendix 5.2.3.2.1i: |Ratio (3.1)| for the seller

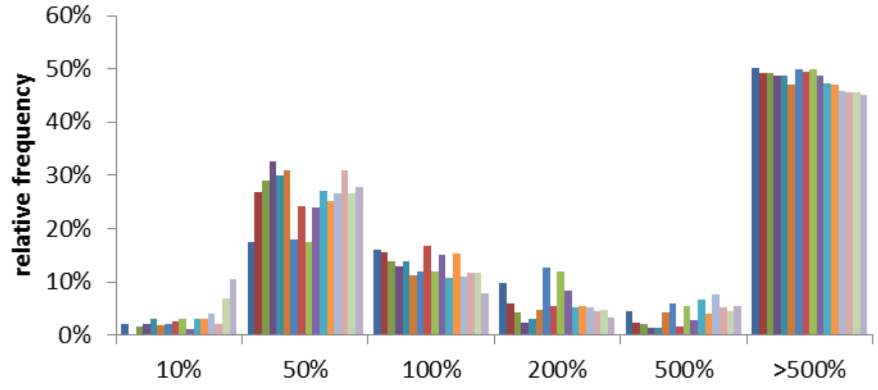


Figure Appendix 5.2.3.2.1j: |Ratio (3.1)| for the seller broken down by factors

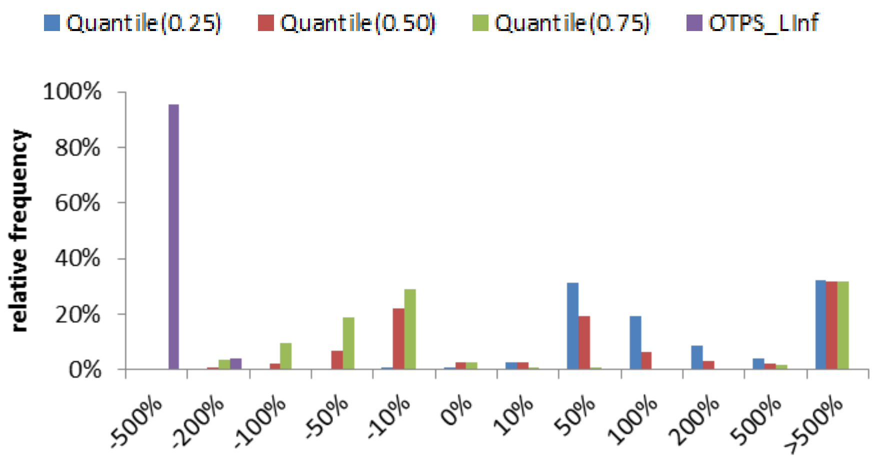


Figure Appendix 5.2.3.2.1k: Ratio (3.1) for the seller

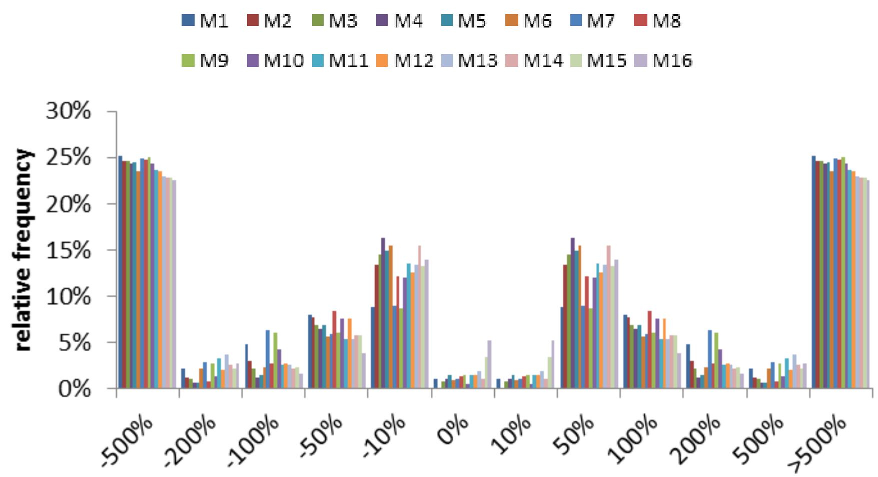


Figure Appendix 5.2.3.2.1l: Ratio (3.1) for the seller broken down by factors

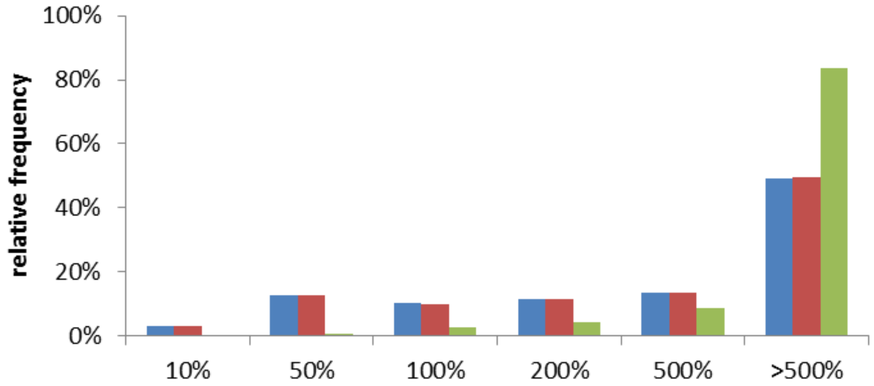


Figure Appendix 5.2.3.2.1m: |Ratio (3.1)| for the seller

OLS WLS OTPS_L2

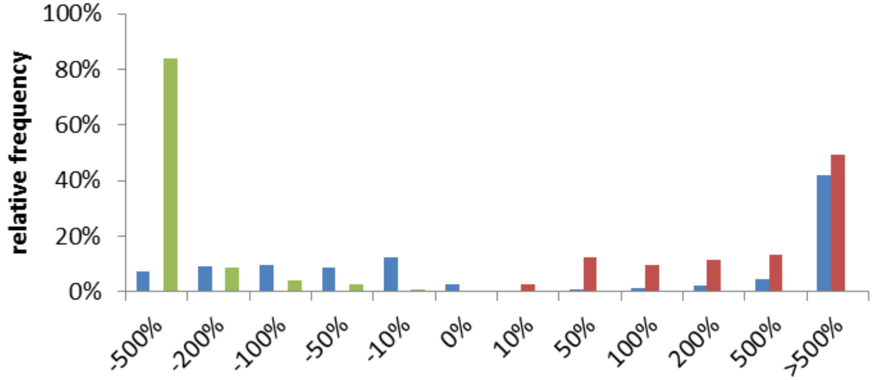


Figure Appendix 5.2.3.2.1o: Ratio (3.1) for the seller

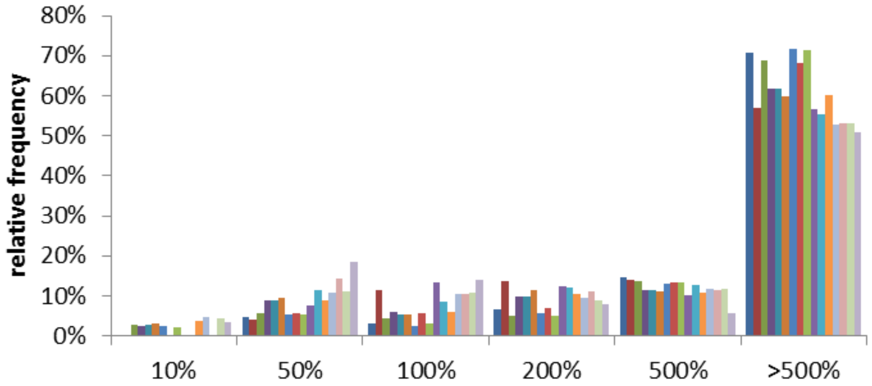


Figure Appendix 5.2.3.2.1n: |Ratio (3.1)| for the seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

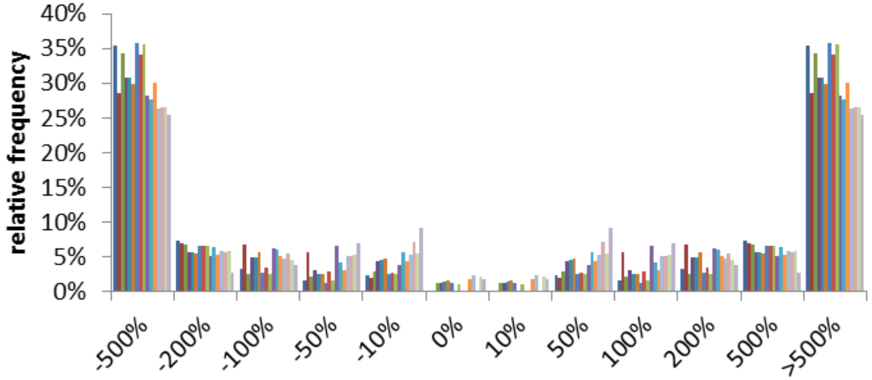


Figure Appendix 5.2.3.2.1p: Ratio (3.1) for the seller broken down by factors

Appendix 5.2.3.2.2 “Similarity”

– "Similarity" of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Buyer

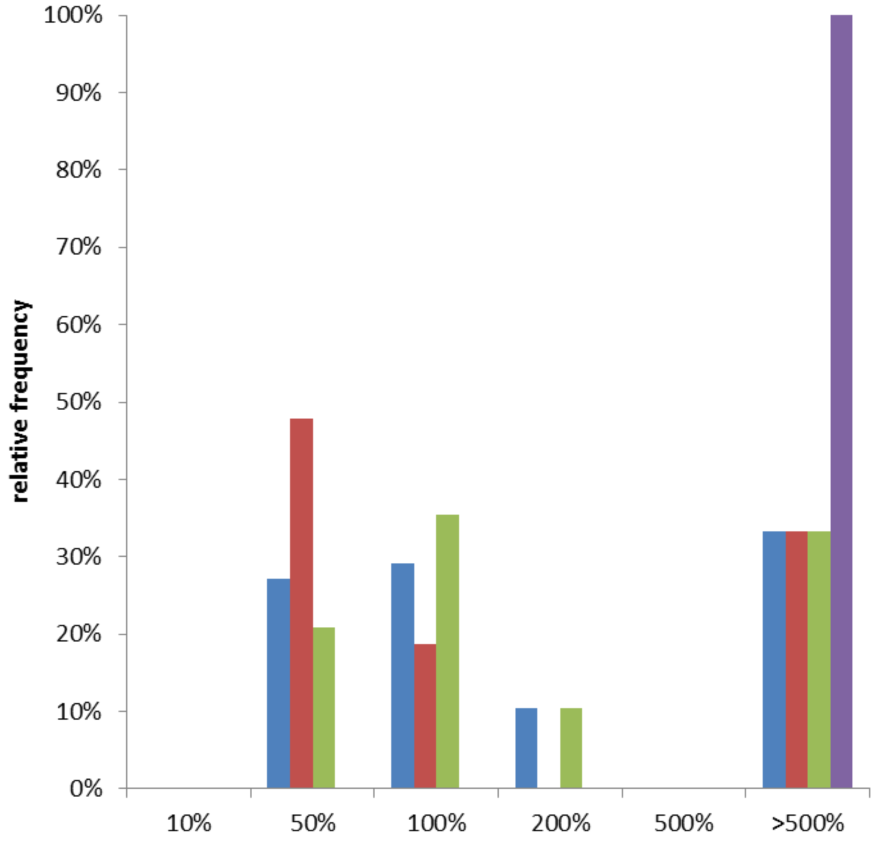


Figure Appendix 5.2.3.2.2a: Area (3.3) for the buyer

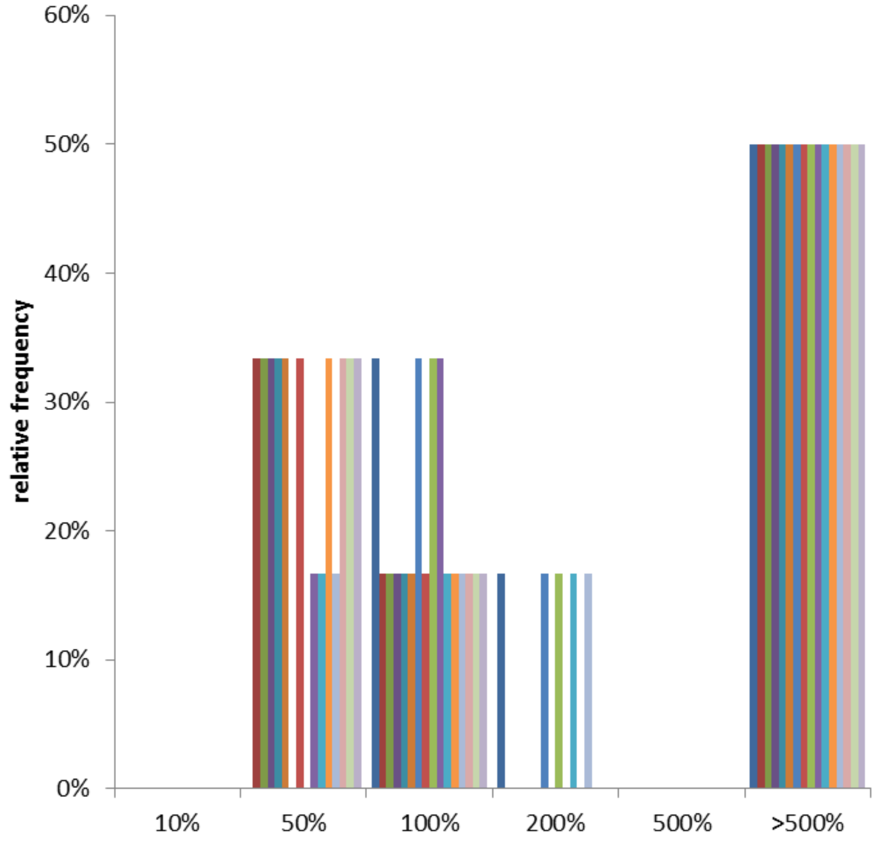


Figure Appendix 5.2.3.2.2b: Area (3.3) for the buyer broken down by factors

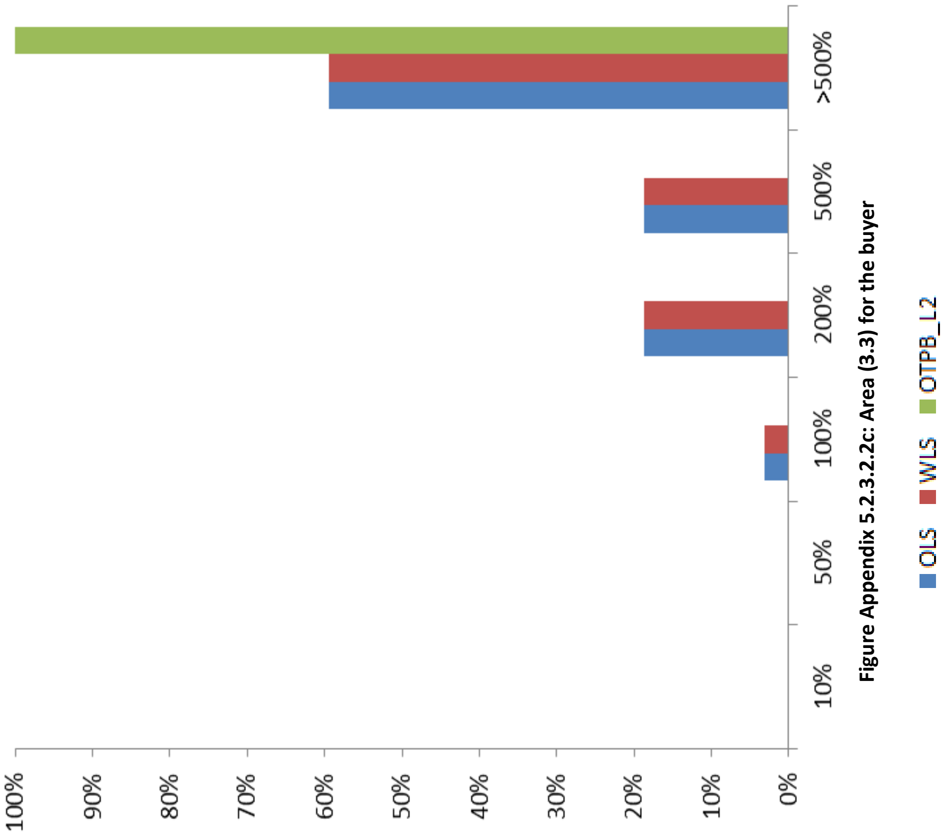


Figure Appendix 5.2.3.2.2c: Area (3.3) for the buyer

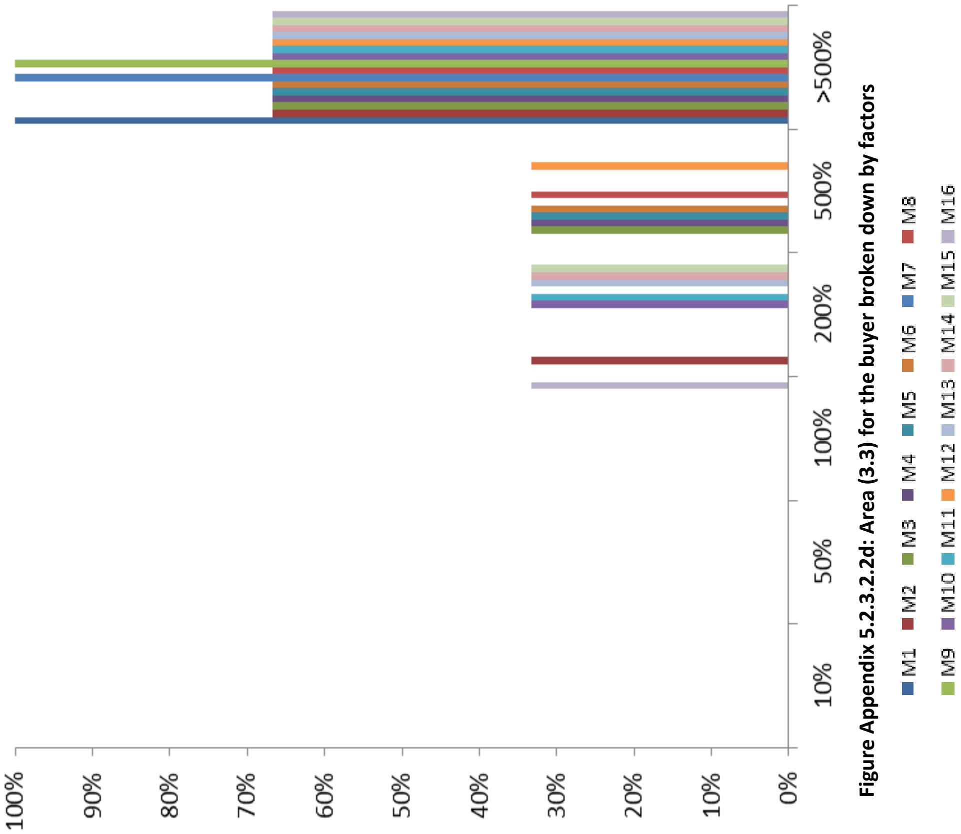


Figure Appendix 5.2.3.2.2d: Area (3.3) for the buyer broken down by factors

– "Similarity" of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Seller

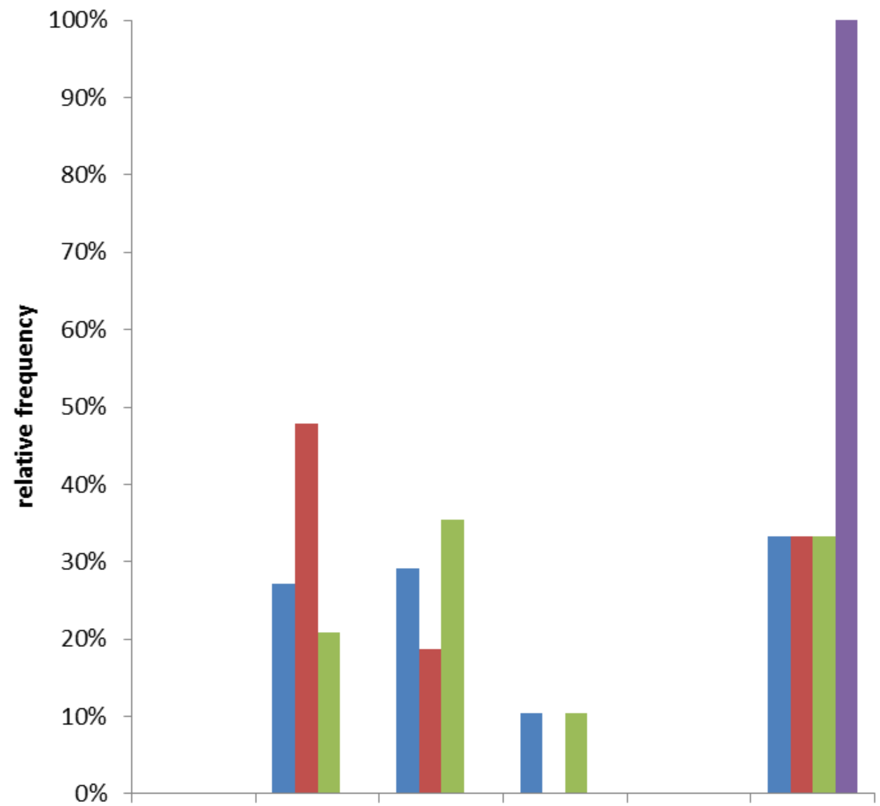


Figure Appendix 5.2.3.2.2e: Area (3.3) for the seller

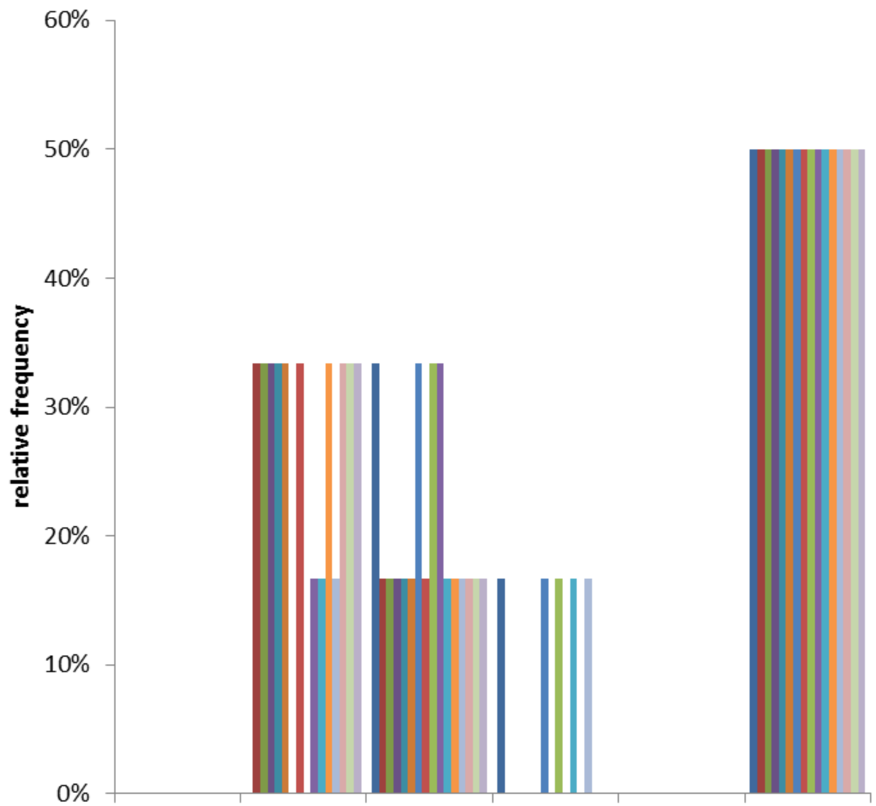


Figure Appendix 5.2.3.2.2f: Area (3.3) for the seller broken down by factors

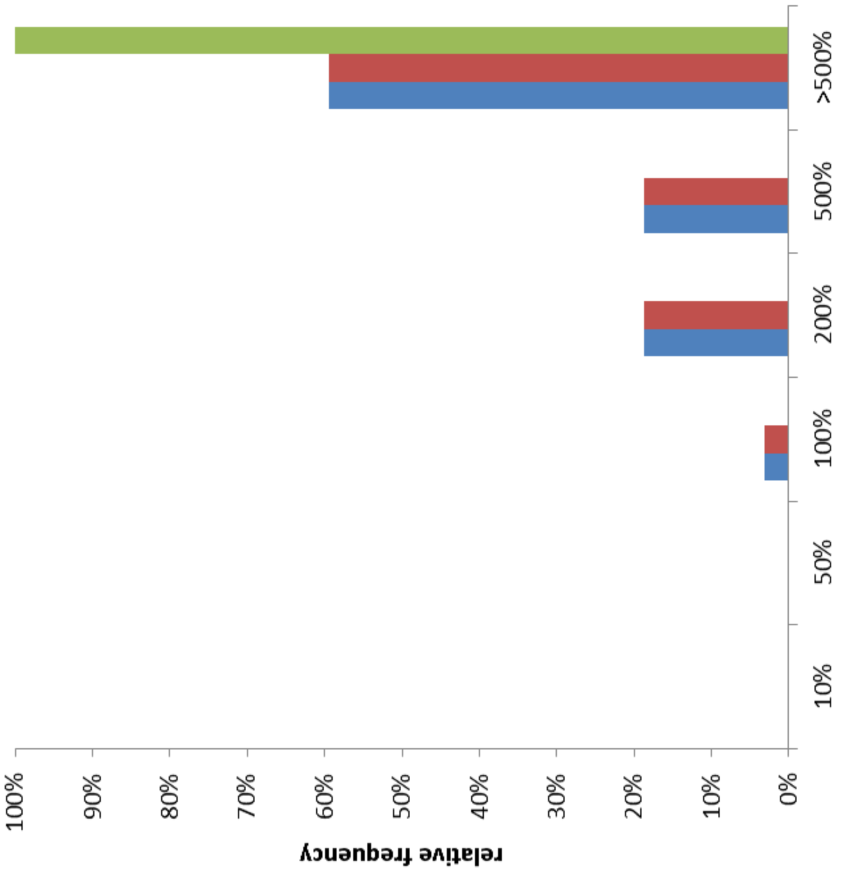


Figure Appendix 5.2.3.2.2g: Area (3.3) for the seller

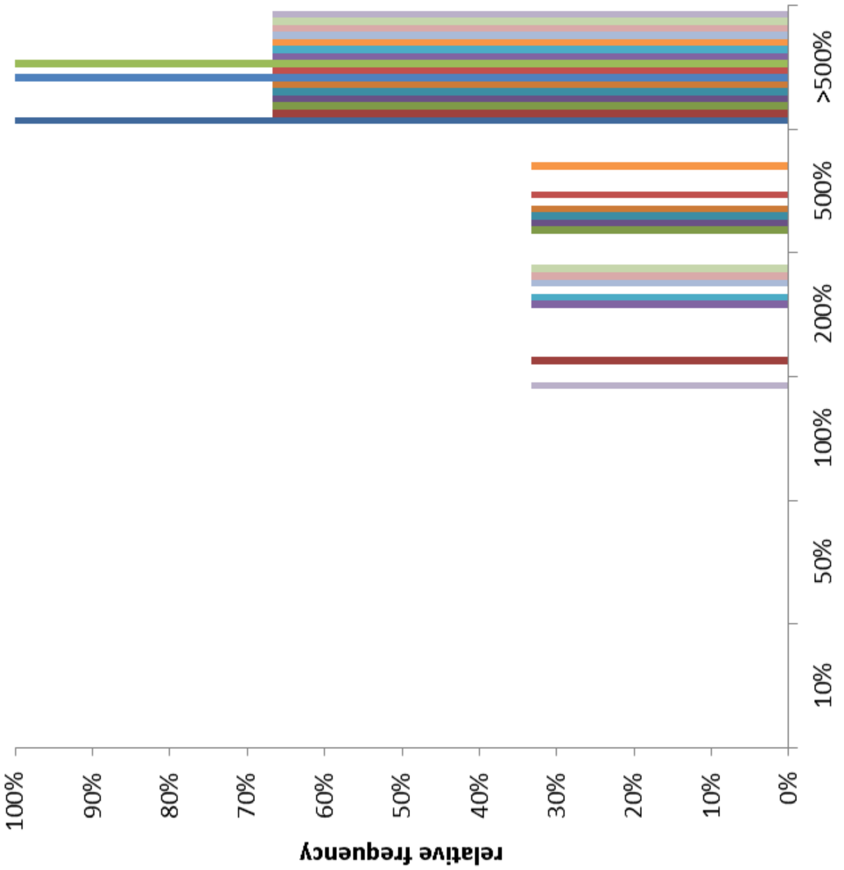


Figure Appendix 5.2.3.2.2h: Area (3.3) for the seller broken down by factors

Appendix 5.2.3.3 Optimize-the-Price-Approach versus Regressions Without Constant $\boldsymbol{\beta_0}$

Appendix 5.2.3.3.1 “Magnitude”

– “Magnitude” of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Buyer

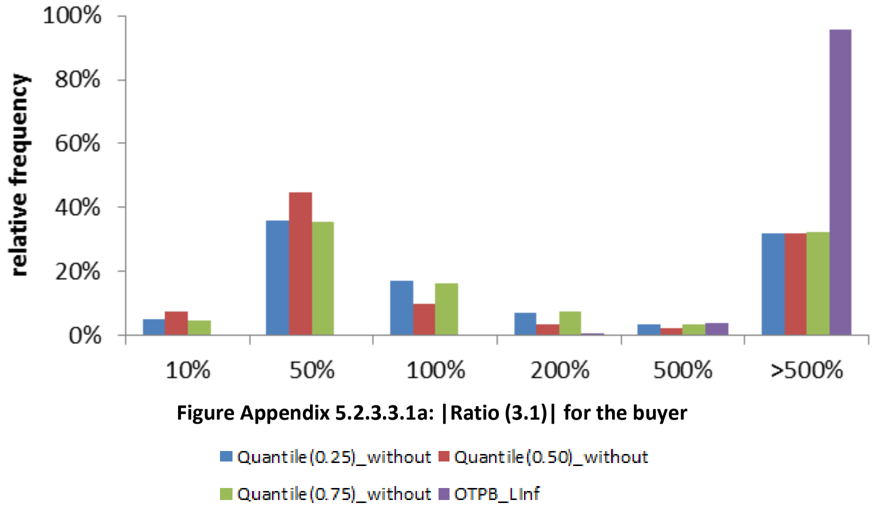


Figure Appendix 5.2.3.3.1a: |Ratio (3.1)| for the buyer

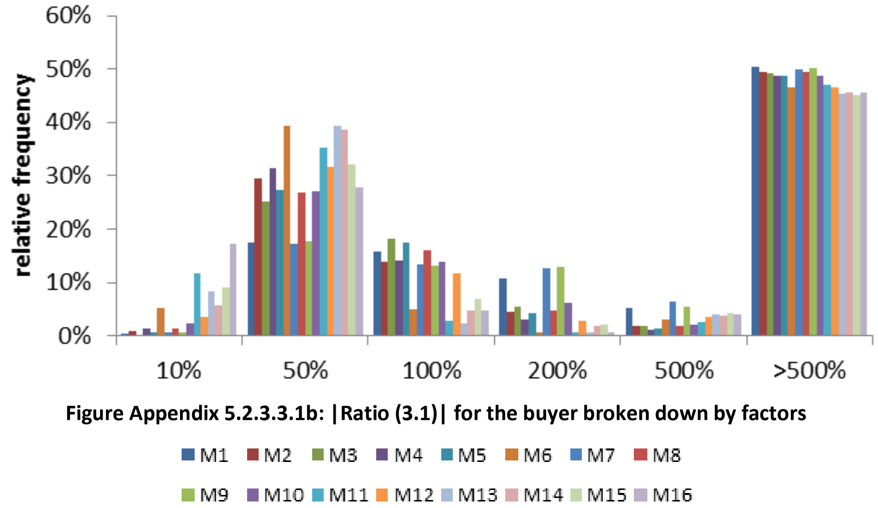


Figure Appendix 5.2.3.3.1b: |Ratio (3.1)| for the buyer broken down by factors

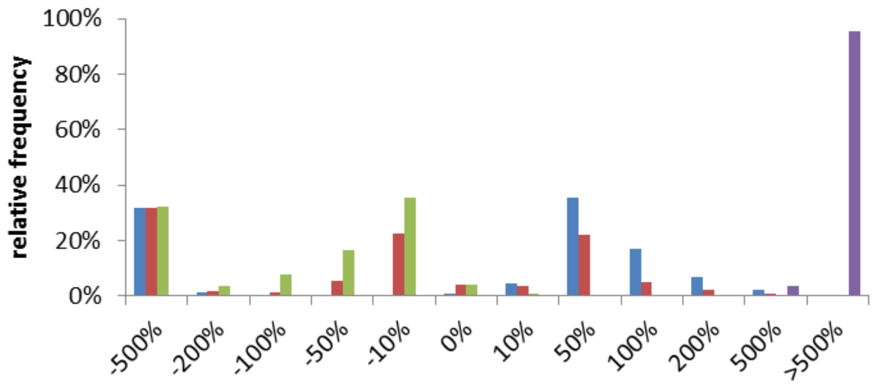


Figure Appendix 5.2.3.3.1c: Ratio (3.1) for the buyer

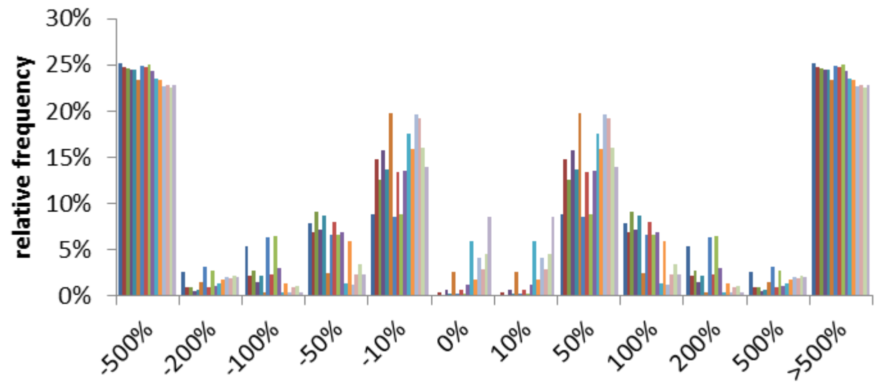


Figure Appendix 5.2.3.3.1d: Ratio (3.1) for the buyer broken down by factors

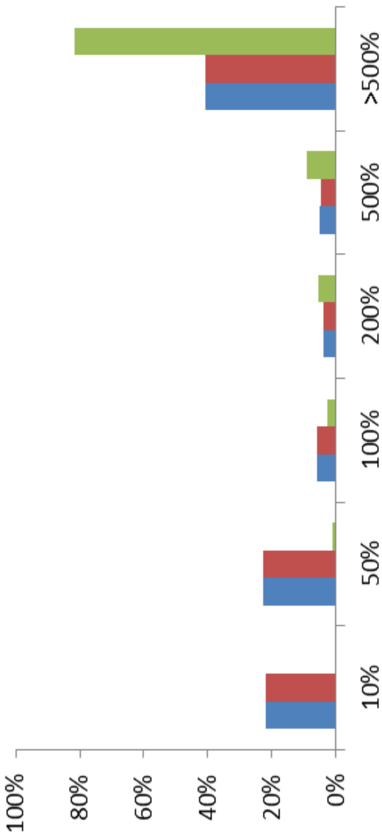


Figure Appendix 5.2.3.3.1e: |Ratio (3.1)| for the buyer

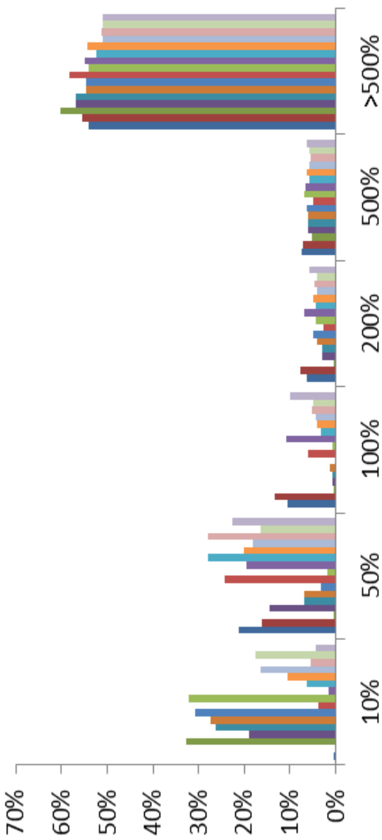


Figure Appendix 5.2.3.3.1f: |Ratio (3.1)| for the buyer broken down by factors

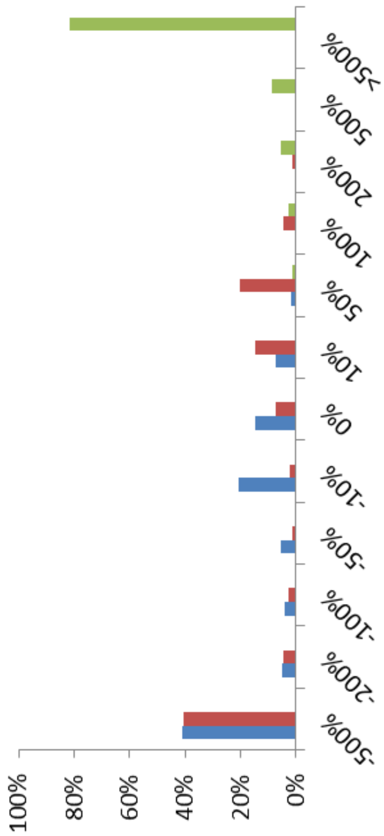


Figure Appendix 5.2.3.3.1g: Ratio (3.1) for the buyer

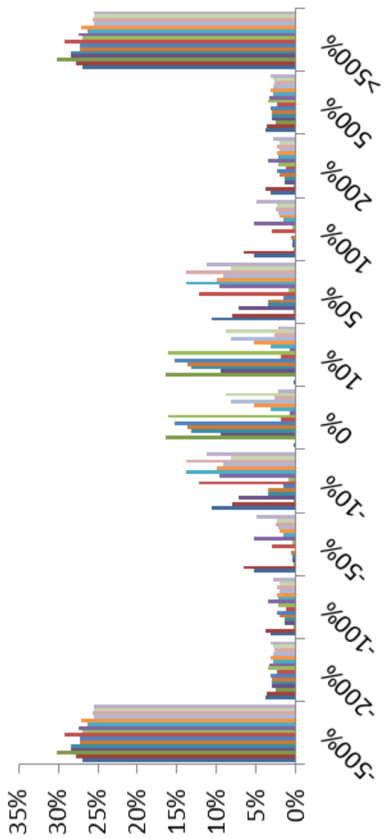


Figure Appendix 5.2.3.3.1h: Ratio (3.1) for the buyer broken down by factors

– "Magnitude" of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Seller

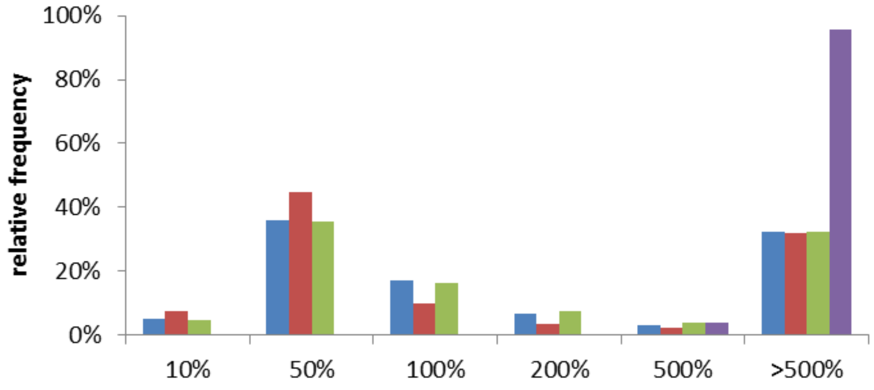


Figure Appendix 5.2.3.3.1i: |Ratio (3.1)| for the seller

Quantile(0.25)_without Quantile(0.50)_without
Quantile(0.75)_without OTPS_LInf

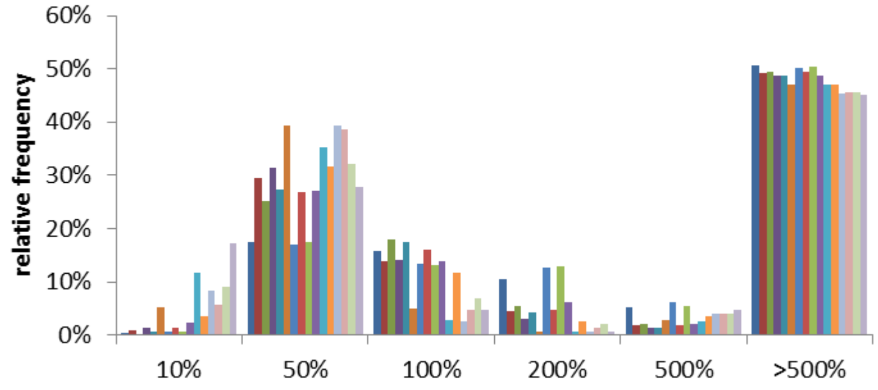


Figure Appendix 5.2.3.3.1j: |Ratio (3.1)| for the seller broken down by factors

M1 M2 M3 M4 M5 M6 M7 M8
M9 M10 M11 M12 M13 M14 M15 M16

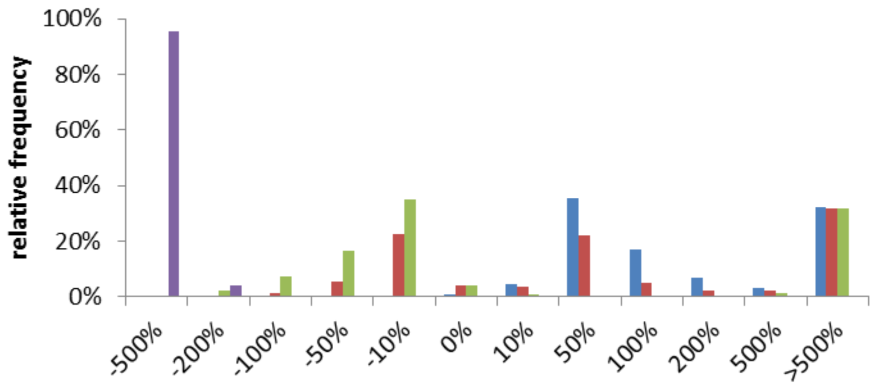


Figure Appendix 5.2.3.3.1k: Ratio (3.1) for the seller

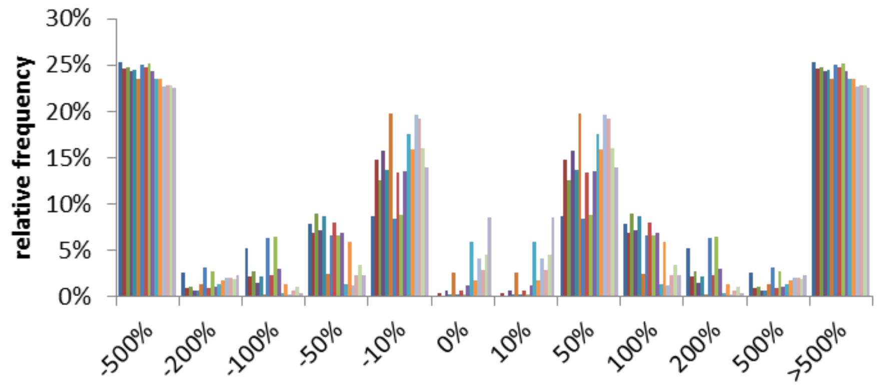


Figure Appendix 5.2.3.3.1l: Ratio (3.1) for the seller broken down by factors

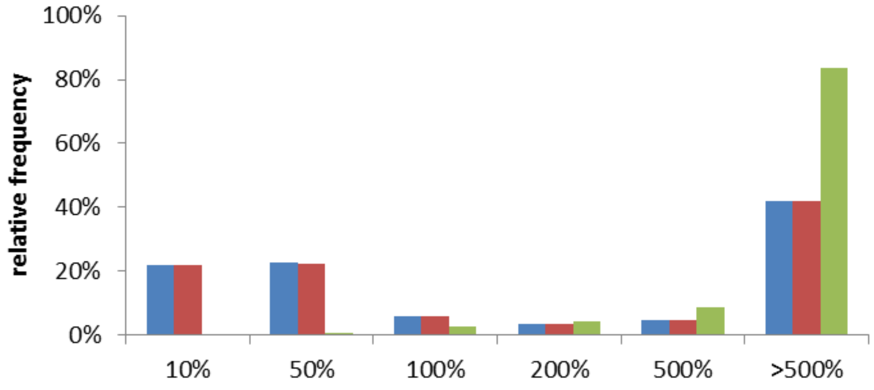


Figure Appendix 5.2.3.3.1m: |Ratio (3.1)| for the seller

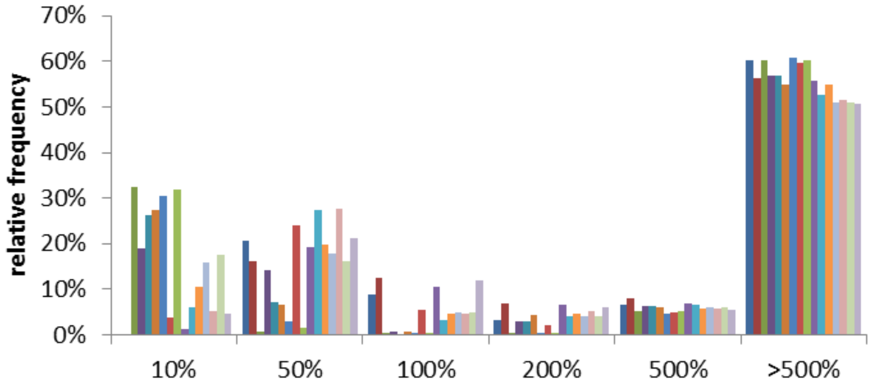


Figure Appendix 5.2.3.3.1n: |Ratio (3.1)| for the seller broken down by factors

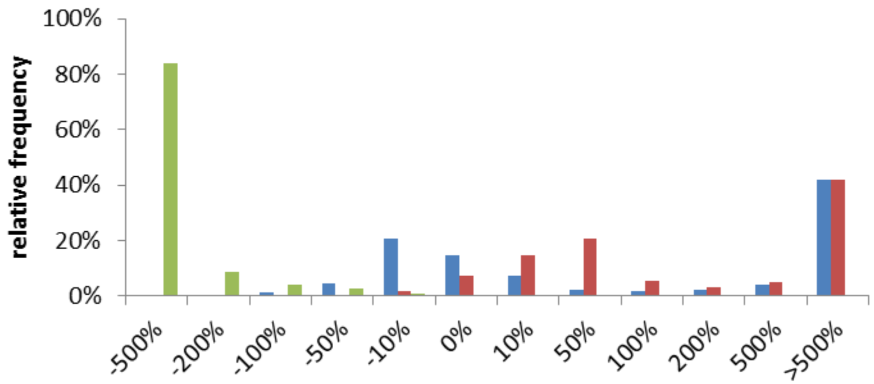


Figure Appendix 5.2.3.3.1o: Ratio (3.1) for the seller

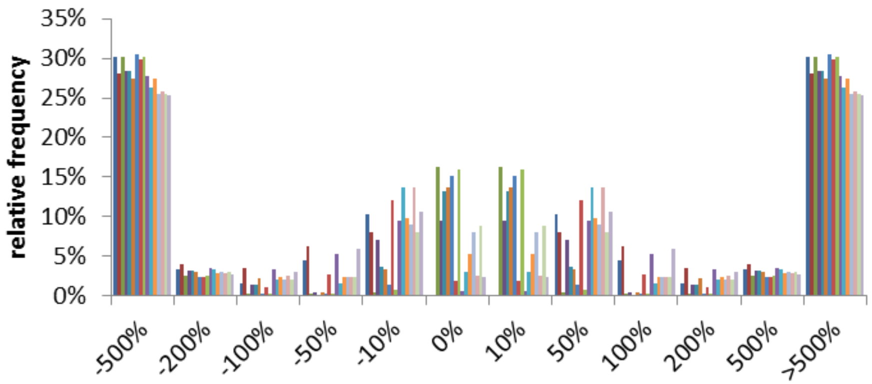


Figure Appendix 5.2.3.3.1p: Ratio (3.1) for the seller broken down by factors

Appendix 5.2.3.3.2 "Similarity"

– "Similarity" of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Buyer

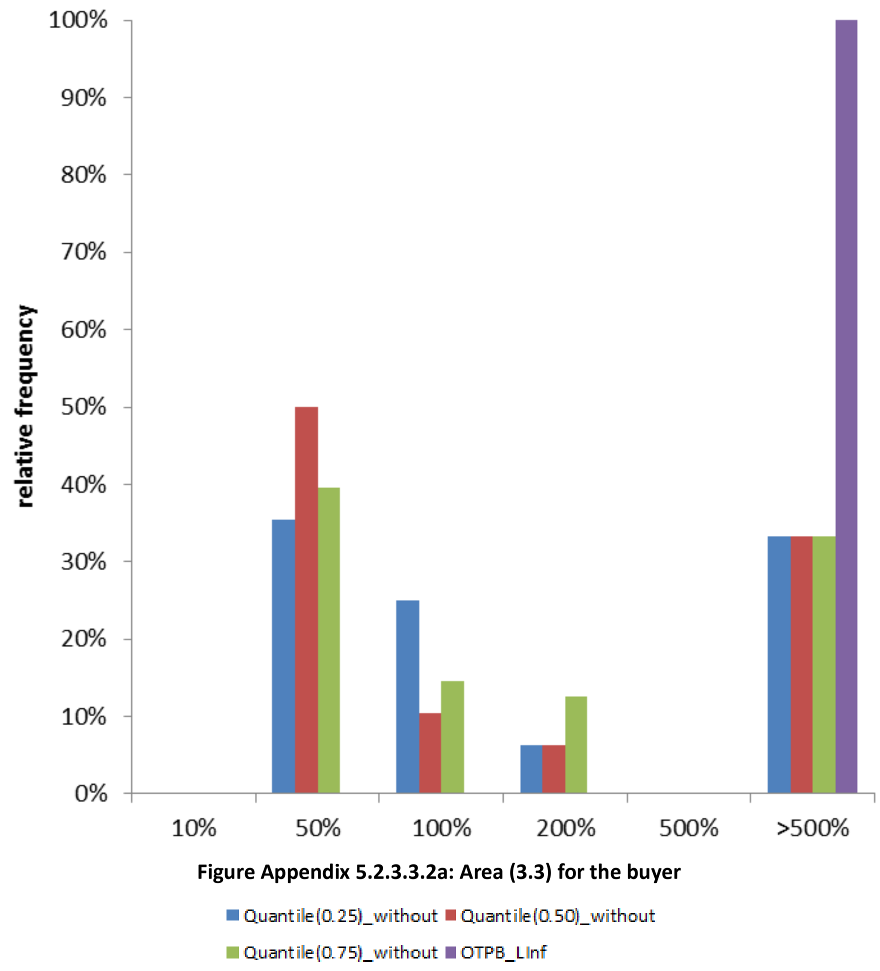


Figure Appendix 5.2.3.3.2a: Area (3.3) for the buyer

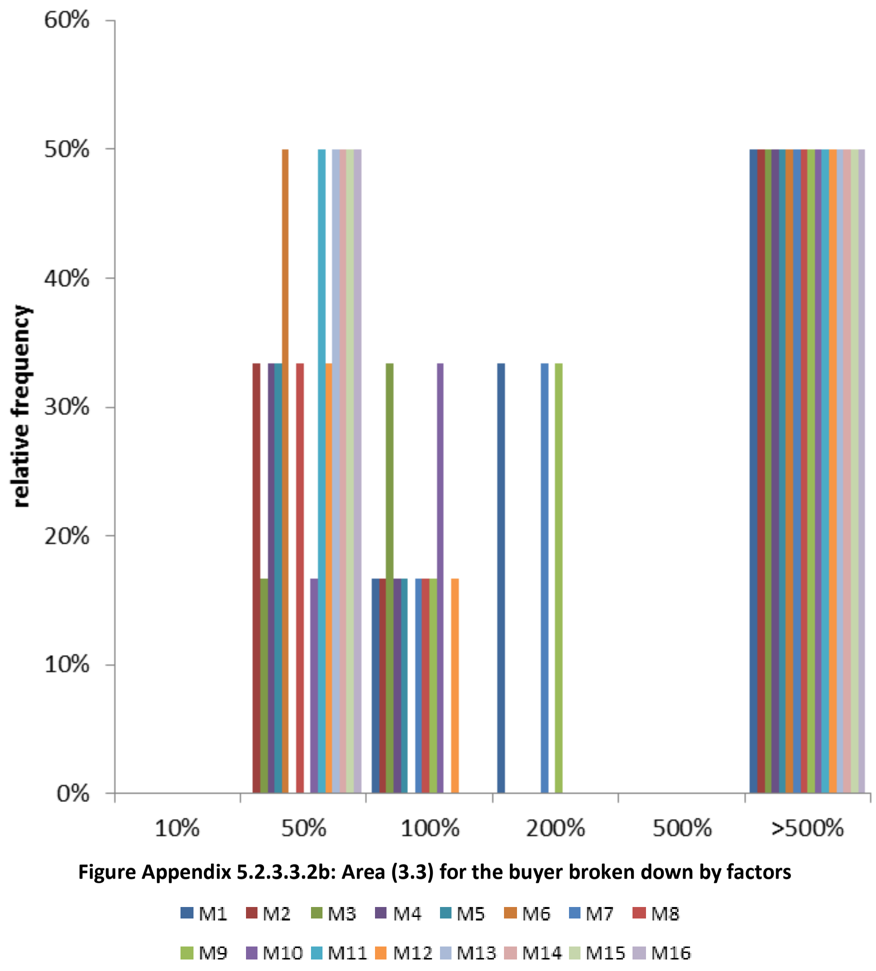


Figure Appendix 5.2.3.3.2b: Area (3.3) for the buyer broken down by factors

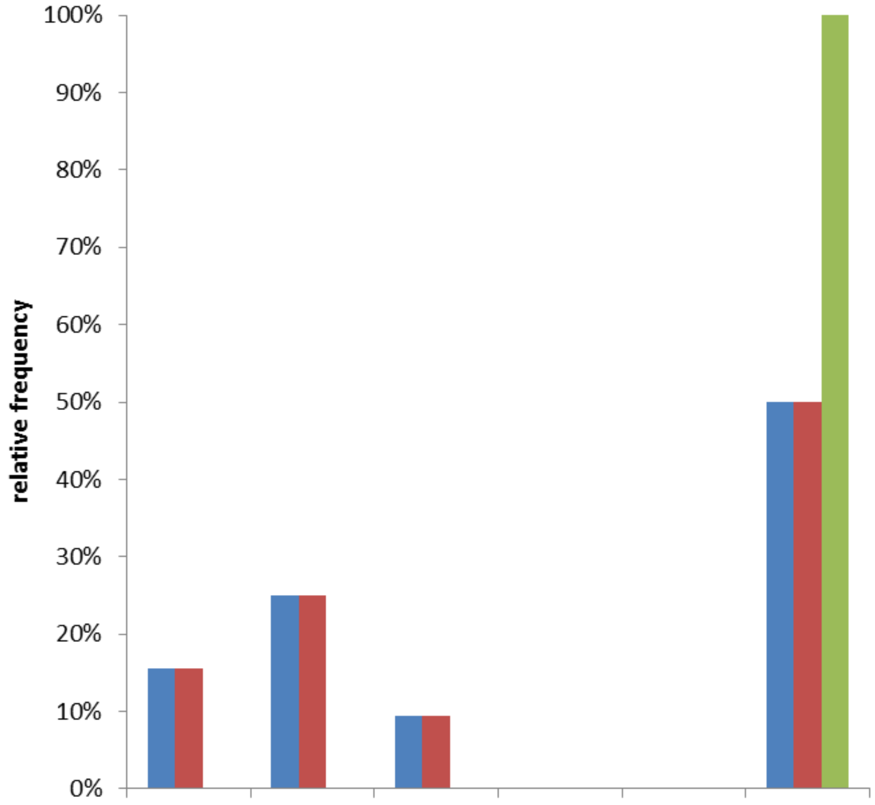


Figure Appendix 5.2.3.3.2c: Area (3.3) for the buyer

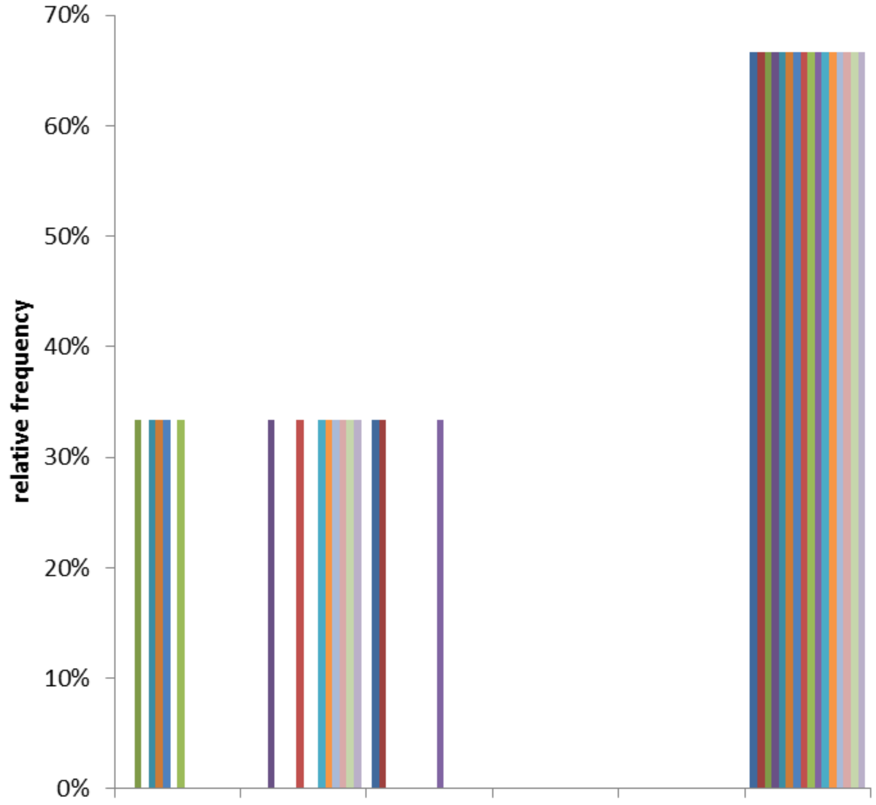


Figure Appendix 5.2.3.3.2d: Area (3.3) for the buyer broken down by factors

– “Similarity” of price differences between Integrated (Optimize-the-Price-Approach) versus Separated Approach (Regressions): Seller

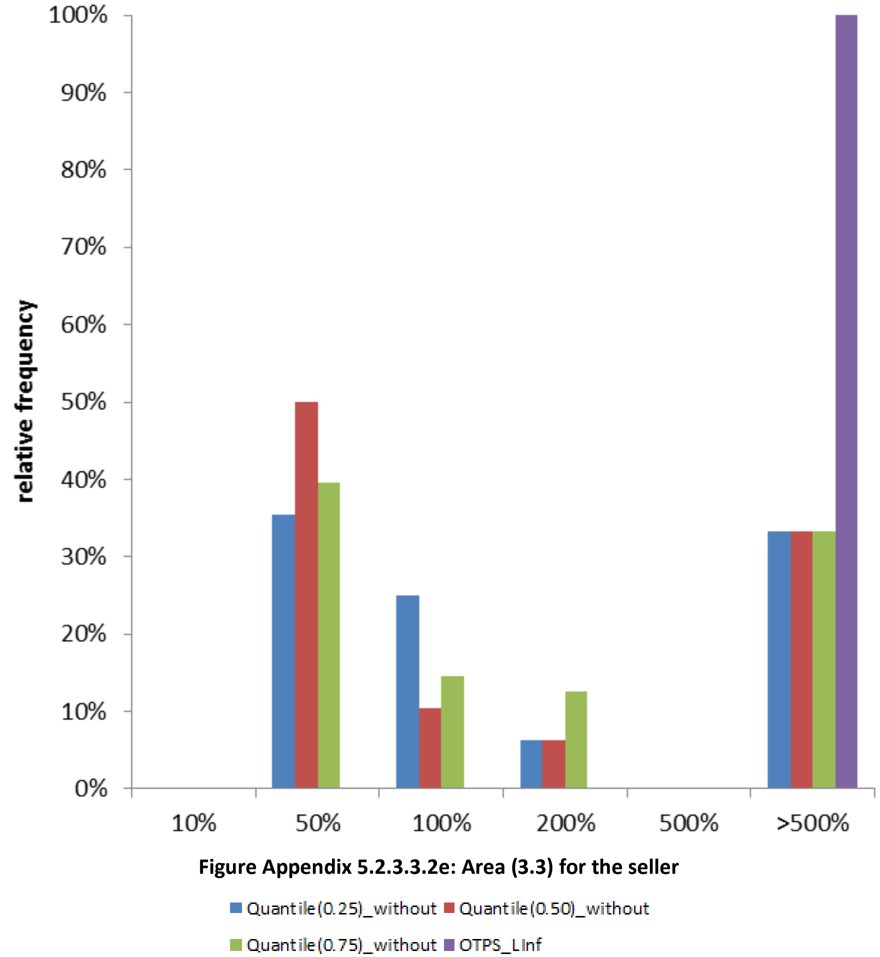


Figure Appendix 5.2.3.3.2e: Area (3.3) for the seller

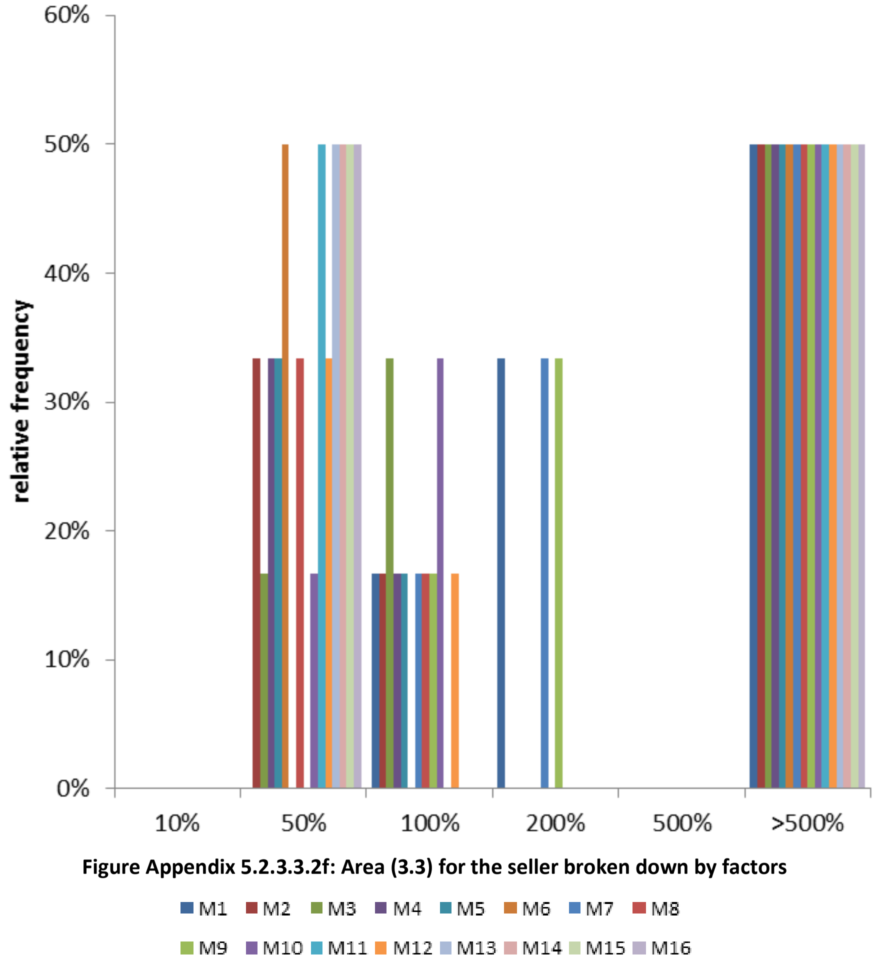


Figure Appendix 5.2.3.3.2f: Area (3.3) for the seller broken down by factors

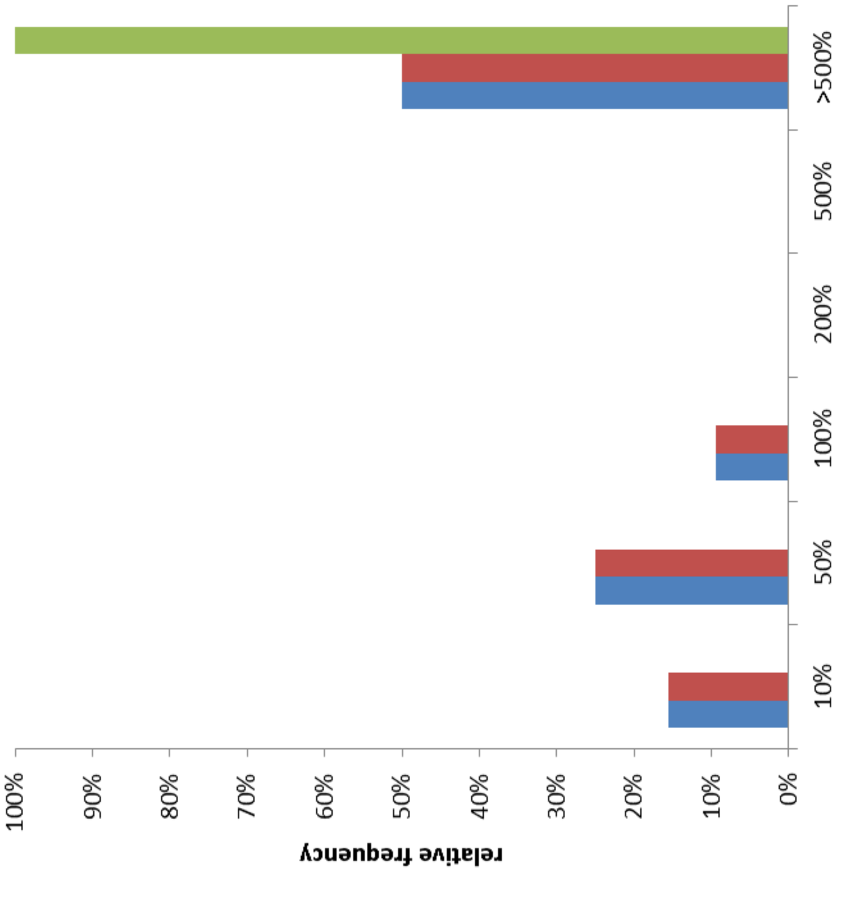


Figure Appendix 5.2.3.3.2g: Area (3.3) for the seller

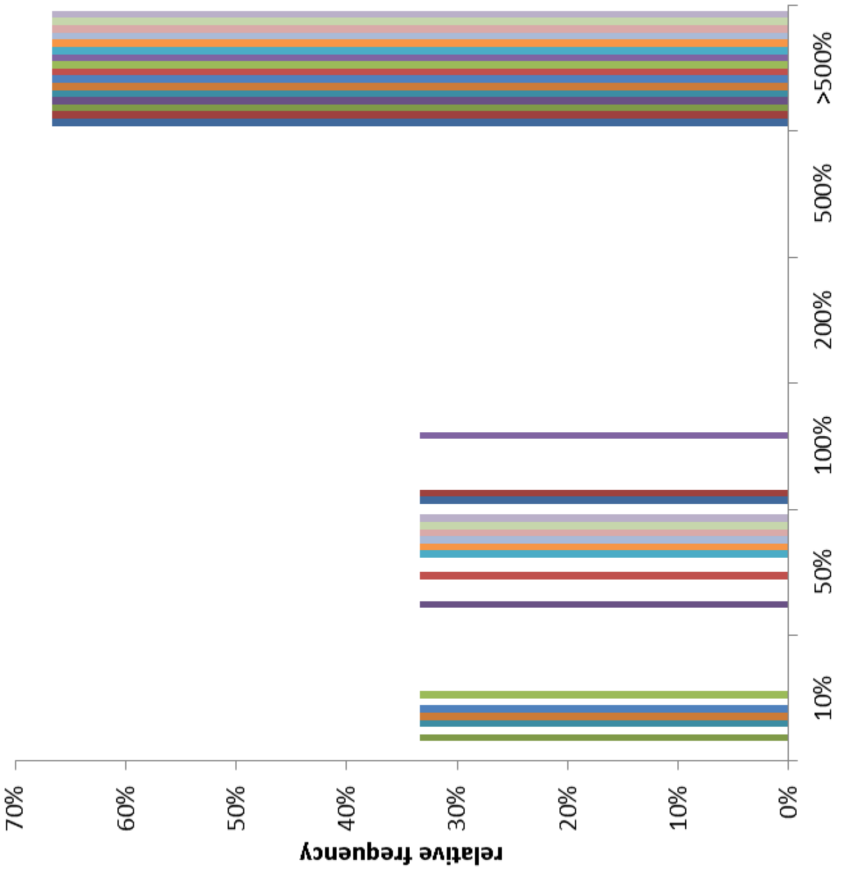


Figure Appendix 5.2.3.3.2h: Area (3.3) for the seller broken down by factors

Appendix 5.3 Cleaning the Results of the Numerical Optimization from Chapter V, Section 3.3.1

Data loss due to lack of convergence
Without cleaning there are 115 calculations per cell, 115x16 = 1,840 calculations per row sum (= sum over models), 115x20 = 2,300 calculations per column sum (=sum over methods) and 115x16x20 = 36,800 calculations in total.

	Model																
Method	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Σ
OTPB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB L1_x0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB L1_x1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	**2**
OTPB L1_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB L2_x0.5	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	**1**
OTPB L2_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB L2_x2	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**1**
OTPB LInf_x0.5	0	0	0	0	0	0	0	1	0	0	1	1	1	2	2	1	**9**
OTPB LInf_x1	1	0	1	0	0	0	1	0	1	0	0	0	1	0	0	1	**6**
OTPB LInf_x2	3	1	1	1	0	0	3	0	1	0	0	0	0	0	0	1	**11**
OTPS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L1_x0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L1_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L1_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L2_x0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L2_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L2_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS LInf_x0.5	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	**2**
OTPS LInf_x1	1	0	1	1	0	0	0	1	1	0	0	0	0	0	0	0	**5**
OTPS LInf_x2	0	0	0	0	0	1	1	1	4	0	2	0	0	1	0	0	**10**
Σ	**6**	**1**	**3**	**2**	**0**	**2**	**6**	**3**	**7**	**0**	**3**	**2**	**3**	**3**	**3**	**3**	**47**

Table Appendix 5.3a: Data loss due to lack of convergence

Data loss due to lack of compliance with constraints Without cleaning there are 115 calculations per cell, 115x16 = 1,840 calculations per row sum (= sum over models), 115x20 = 2,300 calculations per column sum (=sum over methods) and 115x16x20 = 36,800 calculations in total.																	
	Model																
Method	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	Σ
OTPB	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB L1_x0.5	9	12	9	9	15	23	4	17	7	18	20	23	18	23	24	20	**251**
OTPB L1_x1	2	7	3	28	2	10	1	4	8	10	8	7	10	14	19	20	**153**
OTPB L1_x2	4	6	3	2	1	4	1	5	2	8	4	3	2	14	13	11	**83**
OTPB L2_x0.5	3	18	2	7	8	1	4	3	3	9	2	2	2	4	3	5	**76**
OTPB L2_x1	1	2	0	0	0	0	5	4	3	0	0	0	0	0	0	1	**16**
OTPB L2_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB LInf_x0.5	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	**3**
OTPB LInf_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPB LInf_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L1_x0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L1_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L1_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L2_x0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L2_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS L2_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS LInf_x0.5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS LInf_x1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
OTPS LInf_x2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
Σ	**20**	**45**	**17**	**46**	**26**	**38**	**16**	**33**	**23**	**45**	**34**	**35**	**32**	**55**	**59**	**58**	**582**

Table Appendix 5.3b: Data loss due to lack of compliance with constraints

References

AbuGHAZALEH, Naser M. / AL-HARES, Osama M. / HADDAD, Ayman E. (2012): The Value Relevance of goodwill impairments: UK Evidence. International Journal of Economics and Finance 4 (2012), pp. 206-216.

AGGARWAL, Raj / MISHRA, Dev / WILSON, Craig (2018): Analyst recommendations and the implied cost of equity. Review of Quantitative Finance and Accounting 50 (2018), pp. 717-743.

AGRRAWAL, Pankaj / BORGMAN, Richard / CLARK, John M. / STRONG, Robert (2010): Using the Price-to-Earnings Harmonic Mean to Improve Firm Valuation Estimates. Journal of Financial Education 36 (2010), pp. 98-110.

ALEKSANYAN, Mark / KARIM, Khondkar (2013): Searching for value relevance of book value and earnings: a case of premium vs. discount firms. Review of Quantitative Finance and Accounting 41 (2013), pp. 489-511.

ALFORD, Andrew W. (1992): The Effect of the Set of Comparable Firms on the Accuracy of Price-Earnings Valuation Method. Journal of Accounting Research 30 (1992), pp. 94-108.

ALLEN, David E. / SINGH, Abhay K. / POWELL, Robert J. (2011): Asset Pricing, the Fama-French Factor Model and the Implications of Quantile-Regression Analysis, issued by Greg N. GREGORIOU / Razvan PASCALAU. 1. edition, London 2011, pp. 176-193.

ALMGREN, Robert / THUM, Chee / HAUPTMANN, Emmanuel / LI, Hong (2005): Direct Estimation of Equity Market Impact. Risk 18 (2005), pp. 58-62.

ASQUITH, Paul / MIKHAIL, Michael B. / AU, Andrea S. (2005): Information content of equity analyst reports. Journal of Financial Economics 75 (2005), pp. 245-282.

BABOUKARDOS, Diogenis (2018): The valuation relevance of environmental performance revisited: The moderating role of environmental provisions. The British Accounting Review 50 (2018), pp. 32-47.

BAKER, Malcolm / RUBACK, Richard S. (1999): Estimating Industry Multiples. June 1999.
http://www.people.hbs.edu/mbaker/cv/papers/EstimatingIndustry.pdf

BALACHANDRAN, Sudhakar / MOHANRAM, Partha (2011): Is the decline in the value relevance of accounting driven by increased conservatism?. Review of Accounting Studies 16 (2011), pp. 272-301.

BARILLAS, Francisco / SHANKEN, Jay (2017): Which Alpha?. The Review of Financial Studies 30 (2017), pp. 1316-1338.

BARILLAS, Francisco / SHANKEN, Jay (2018): Comparing Asset Pricing Models. The Journal of Finance 73 (2018), pp. 715-754.

BARTH, Mary E. / BEAVER, William H. / HAND, John R. M. / LANDSMAN, Wayne R. (2005): Accruals, Accounting-Based Valuation Models and the Prediction of Equity Values. Journal of Accounting, Auditing, and Finance 20 (2005), pp. 311-345.

BARTH, Mary E. / BEAVER, William H. / LANDSMAN, Wayne R. (2001): The relevance of the value relevance literature for financial accounting standard setting: another view. Journal of Accounting and Economics 31 (2001), pp. 77-104.

BARTH, Mary E. / LANDSMAN, Wayne R. / LANG, Mark H. / WILLIAMS, Christopher D. (2012): Are IFRS-based and US GAAP-based Accounting Amounts Comparable?. Journal of Accounting and Economics 54 (2012), pp. 68-93.

BARTH, Mary E. / LI, Ken / McCLURE, Charles (2018): Evolution in Value Relevance of Accounting Information. Stanford University Graduate School of Business Research Paper No. 17-24. September 2018.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2933197

BARTOV, Eli / GOLDBERG, Stephen R. / KIM, Myungsun (2005): Comparative Value Relevance Among German, U.S., and International Accounting Standards: A German Stock Market Perspective. Journal of Accounting, Auditing, and Finance 20 (2005), pp. 95-119.

BAUER, Rob / KOEDIJK, Kees / OTTEN, Roger (2005): International evidence on ethical mutual fund performance and investment style. Journal of Banking & Finance 29 (2005), pp. 1751-1767.

BEATTY, Randolph / RIFFE, Susan / THOMPSON, Rex (1999): The Method of Comparables and Tax Court Valuations of Private Firms: An Empirical Investigation. Accounting Horizons 13 (1999), pp. 177-199.

BECKMANN, Christoph / MEISTER, Jan M. / MEITNER, Matthias (2003): Das Multiplikatorverfahren in der kapitalmarktorientierten Unternehmensbewertungspraxis. Finanz Betrieb 5 (2003), pp. 103-105.

BERK, Jonathan / DeMARZO, Peter (2017): Corporate Finance: The Core. 4. edition, London 2017.

BLACK, Fischer / JENSEN, Michael C. / SCHOLES, Myron S. (1972): The Capital Asset Pricing Model: Some Empirical Tests, issued by Michael C. JENSEN. 1. edition, New York 1972.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=908569

BLACK, Fischer / SCHOLES, Myron S. (1973): The Pricing of Options and Corporate Liabilities. The Journal of Political Economy 81 (1973), pp. 637-654.

BOLLEN, Nicolas P. B. (2007): Mutual Fund Attributes and Investor Behavior. Journal of Financial and Quantitative Analysis 42 (2007), pp. 683-708.

BOYD, Stephen / VANDENBERGHE, Lieven (2009): Convex Optimization. 7. edition, Cambridge 2009.

BREALEY, Richard A. / MYERS, Stewart C. / ALLEN, Franklin (2016): Principles of Corporate Finance. 12. edition, New York 2016.

BROWN, Stephen / LO, Kin / LYS, Thomas (1999): Use of R^2 in accounting research: measuring changes in value relevance over the last four decades. Journal of Accounting and Economics 28 (1999), pp. 83-115.

CAMPBELL, John Y. / THOMPSON, Samuel B. (2008): Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?. The Review of Financial Studies 21 (2008), pp. 1509-1531.

CARHART, Mark M. (1997): On Persistence in Mutual Fund Performance. The Journal of Finance 52 (1997), pp. 57-82.

CASCINO, Stefano / CLATWORTHY, Mark / GARCÍA OSMA, Beatriz / GASSEN, Joachim / IMAM, Shahed / THOMAS, Jeanjean (2014): Who Uses Financial Reports and for What Purpose? Evidence from Capital Providers. Accounting in Europe 11 (2014), pp. 185-209.

CAZAVAN-JENY, Anne / JEANJEAN, Thomas (2006): The Negative Impact of R&D Capitalization: A Value Relevance Approach. European Accounting Review 15 (2006), pp. 37-61.

CHEN, Nai-Fu / ROLL, Richard / ROSS, Stephen A. (1986): Economic Forces and the Stock Market. The Journal of Business 59 (1986), pp. 383-403.

CHENG, C. S. Agnes / McNAMARA, Ray (2000): The Valuation Accuracy of the Price-Earnings and Price-Book Benchmark Valuation Methods. Review of Quantitative Finance and Accounting 15 (2000), pp. 349-370.

CLAUS, James / THOMAS, Jacob (2001): Equity Premia as Low as Three Percent? Evidence from Analysts' Earnings Forecasts for Domestic and International Stock Markets. The Journal of Finance 56 (2001), pp. 1629-1666.

COCHRANE, John H. (2005): Asset Pricing. Revised edition, Princeton/Oxford 2005.

COENENBERG, Adolf G. / SCHULTZE, Wolfgang (2002): Das Multiplikator-Verfahren in der Unternehmensbewertung: Konzeption und Kritik. Finanz Betrieb 4 (2002), pp. 697-703.

COLLATZ, Lothar (1964): Funktionalanalysis und numerische Mathematik. 1. edition, Berlin/Heidelberg 1964.

COLLINS, Daniel W. / PINCUS, Morton / XIE, Hong (1999): Equity Valuation and Negative Earnings: The Role of Book Value of Equity. The Accounting Review 74 (1999), pp. 29-61.

CORE, John E. / GUAY, Wayne R. / VAN BUSKIRK, Andrew (2003): Market valuations in the New Economy: an investigation of what has changed. Journal of Accounting and Economics 34 (2003), pp. 43-67.

COX, John C. / INGERSOLL, Jonathan E / ROSS, Stephen A. (1985): An Intertemporal General Equilibrium Model of Asset Prices. Econometrica 53 (1985), pp. 363-484.

DAMODARAN, Aswath (2006): Damodaran on valuation: security analysis for investment and corporate finance. 2. edition, Hoboken 2006.

DAMODARAN, Aswath (2012): Investment Valuation Tools and Techniques for Determining the Value of Any Asset. 3. edition, Hoboken 2012.

DECHOW, Patricia / GE, Weili / SCHRAND, Catherine (2010): Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics 50 (2010), pp. 344-401.

DEFOND, Mark L. / ZHANG, Jieying (2014): The Timeliness of the Bond Market Reaction to Bad Earnings News. Contemporary Accounting Research 31 (2014), pp. 911-936.

DEMIRAKOS, Efthimios G. / STRONG, Norman C. / WALKER, Martin (2004): What Valuation Models Do Analysts Use?. Accounting Horizons 18 (2004), pp. 221-240.

DITTMANN, Ingolf / MAUG, Ernst (2008): Biases and Error Measures: How to Compare Valuation Methods. August 2008. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=884337

EASTON, Peter D. (2004): PE Ratios, PEG Ratios, and Estimating the Implied Expected Rate of Return on Equity Capital. The Accounting Review 79 (2004), pp. 73-95.

EASTON, Peter D. / HARRIS, Trevor S. (1991): Earnings As an Explanatory Variable for Returns. Journal of Accounting Research 29 (1991), pp. 19-36.

EASTON, Peter D. / MONAHAN, Steven J. / VASVARI, Florin P. (2009): Initial Evidence on the Role of Accounting Earnings in the Bond Market. Journal of Accounting Research 47 (2009), pp. 721-766.

EASTON, Peter D. / SOMMERS, Gregory A. (2003): Scale and the Scale Effect in Market-based Accounting Research. Journal of Business Finance & Accounting 30 (2003), pp. 25-55.

FAMA, Eugene F. (1977): Risk-adjusted discount rates and capital budgeting under uncertainty. Journal of Financial Economics 5 (1977), pp. 3-24.

FAMA, Eugene F. (1991): Efficient Capital Markets: II. The Journal of Finance 46 (1991), pp. 1575-1617.

FAMA, Eugene F. (1996): Discounting Under Uncertainty. The Journal of Business 69 (1996), pp. 415-428.

FAMA, Eugene F. (1998): Market efficiency, long-term returns, and behavioral finance. Journal of Financial Economics 49 (1998), pp. 283-306.

FAMA, Eugene F. / FRENCH, Kenneth R. (1993): Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33 (1993), pp. 3-56.

FAMA, Eugene F. / FRENCH, Kenneth R. (1997): Industry Costs of Equity. Journal of Financial Economics 43 (1997), pp. 153-193.

FAMA, Eugene F. / FRENCH, Kenneth R. (2015): A five-factor asset pricing model. Journal of Financial Economics 116 (2015), pp. 1-22.

FAMA, Eugene F. / FRENCH, Kenneth R. (2018): Choosing Factors. Journal of Financial Economics 128 (2018), pp. 234-252.

FAMA, Eugene F. / MacBETH, James (1973): Risk, Return, and Equilibrium: Empirical Tests. Journal of Political Economy 38 (1973), pp. 607-636.

FELTHAM, Gerald A. / OHLSON, James A. (1995): Valuation and Clean Surplus Accounting for Operating and Financial Activities. Contemporary Accounting Research 11 (1995), pp. 689-731.

FRANCIS, Jennifer / LaFOND, Ryan / OLSSON, Per / SCHIPPER, Katherine (2005): The market pricing of accruals quality. Journal of Accounting and Economics 39 (2005), pp. 295-327.

FRANCIS, Jennifer / SCHIPPER, Katherine (1999): Have Financial Statements Lost Their Relevance?. Journal of Accounting Research 37 (1999), pp. 319-352.

GATHERAL, Jim (2010): No-Dynamic-Arbitrage and Market Impact. Quantitative Finance 10 (2010), pp. 749-759.

GEBHARDT, William R. / LEE, Charles M. C. / SWAMINATHAN, Bhaskaran (2001): Toward an Implied Cost of Capital. Journal of Accounting Research 39 (2001), pp. 135-176.

GIBBONS, Michael R. / ROSS, Stephen A. / SHANKEN, Jay (1989): A Test of the Efficiency of a Given Portfolio. Econometrica 57 (1989), pp. 1121-1152.

GIVOLY, Dan / HAYN, Carla / KATZ, Sharon (2017): The changing relevance of accounting information to debt holders over time. Review of Accounting Studies 22 (2017), pp. 64-108.

GRINOLD, Richard (2006): A Dynamic Model of Portfolio Management. Journal Of Investment Management 4 (2006), pp. 5-22.

GU, Shihao / KELLY, Bryan T. / XIU, Dacheng (2018): Empirical Asset Pricing via Machine Learning. Swiss Finance Institute Research Paper No. 18-71. July 2018.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3281018

HAJEK, Bruce (2015): Random Processes for Engineers. 1. edition, Cambridge 2015.

HAND, John R. M. / COYNE, Joshua G. / GREEN, Jeremiah R. / ZHANG, X. Frank (2017): The Use of Residual Income Valuation Methods by U.S. Sell-Side Equity Analysts. Journal of Financial Reporting 2 (2017), pp. 1-29.

HARRIS, Mary S. / MULLER, Karl A. (1999): The market valuation of IAS versus US-GAAP accounting measures using Form 20-F reconciliations. Journal of Accounting and Economics 26 (1999), pp. 285-312.

HARRISON, J. Michael / KREPS, David (1979): Martingales and arbitrage in multiperiod securities markets. Journal of Economic Theory 20 (1979), pp. 381-408.

HARVEY, Campbell R. (2017): Presidential Address: The Scientific Outlook in Financial Economics. The Journal of Finance 72 (2017), pp. 1399-1440.

HARVEY, Campbell R. / LIU, Yan (2019): A Census of the Factor Zoo. February 2019.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3341728

HARVEY, Campbell R. / LIU, Yan / ZHU, Heqing (2016): ... and the Cross-Section of Expected Returns. The Review of Financial Studies 29 (2016), pp. 5-68.

HENNING, Steven L. / LEWIS, Barry L. / SHAW, Wayne H. (2000): Valuation of the Components of Purchased Goodwill. Journal of Accounting Research 38 (2000), pp. 375-386.

HERATH, Hemantha S. B. / RICHARDSON, A. William / ROUBI, Raafat R. / TIPPETT, Mark (2015): Non-linear Equity Valuation: An Empirical Analysis. ABACUS 51 (2015), pp. 86-115.

HODGES, S. D. / SCHAEFER, S. M. (1977): A Model for Bond Portfolio Improvement. The Journal of Financial and Quantitative Analysis 12 (1977), pp. 243-260.

HOLTHAUSEN, Robert W. / WATTS, Ross L. (2001): The relevance of the value-relevance literature for financial accounting standard setting. Journal of Accounting and Economics 31 (2001), pp. 3-75.

HONG, Harrison / KACPERCZYK, Marcin (2009): The price of sin: The effects of social norms on markets. Journal of Financial Economics 93 (2009), pp. 15-36.

HORTON, Joanne / SERAFEIM, George (2010): Market reaction to and valuation of IFRS reconciliation adjustments: first evidence from the UK. Review of Accounting Studies 15 (2010), pp. 725-751.

HOU, Kewei / KAROLYI, G. Andrew / KHO, Bong-Chan (2011): What Factors Drive Global Stock Returns?. The Review of Financial Studies 24 (2011), pp. 2527-2574.

HOU, Kewei / XUE, Chen / ZHANG, Lu (2015): Digesting Anomalies: An Investment Approach. The Review of Financial Studies 28 (2015), pp. 350-705.

HUNG, Mingyi / SUBRAMANYAM, K. R. (2007): Financial statement effects of adopting international accounting standards: the case of Germany. Review of Accounting Studies 12 (2007), pp. 623-657.

IBIKUNLE, Gbenga / STEFFEN, Tom (2017): European Green Mutual Fund Performance: A Comparative Analysis with Their Conventional and Black Peers. Journal of Business Ethics 145 (2017), pp. 337-355.

IMAM, Shahed / BARKER, Richard / CLUBB, Colin (2008): The Use of Valuation Models by UK Investment Analysts. European Accounting Review 17 (2008), pp. 503-535.

INGERSOLL, Jonathan E. (1987): Theory of Financial Decision Making. 1. edition, Lanham 1987.

JIANG, Wei / STARK, Andrew W. (2013): Dividends, research and development expenditures, and the value relevance of book value for UK loss-making firms. The British Accounting Review 45 (2013), pp. 112-124.

KALLAPUR, Sanjay / KWAN, Sabrina Y. S. (2004): The Value Relevance and Reliability of Brand Assets Recognized by U.K. Firms. The Accounting Review 79 (2004), pp. 151-172.

KIM, Moonchul / RITTER, Jay R. (1999): Valuing IPOs. Journal of Financial Economics 53 (1999), pp. 409-437.

KOENKER, Roger (2005): Quantile Regression. 1. edition, Cambridge/New York/Melbourne/Madrid/Cape Town/Singapore/São Paulo 2005.

KOIJEN, Ralph S. J. / VAN NIEUWERBURGH, Stijn (2011): Predictability of Returns and Cash Flows. Annual Review of Financial Economics 3 (2011), pp. 467-491.

KOTHARI, S. P. (2001): Capital markets research in accounting. Journal of Accounting and Economics 31 (2001), pp. 105-231.

KOTHARI, S. P. / ZIMMERMAN, Jerold L. (1995): Price and return models. Journal of Accounting and Economics 20 (1995), pp. 155-192.

KOUSENIDIS, Dimitrios V. / LADAS, Anestis C. / NEGAKIS, Christos I. (2009): Value relevance of conservative and non-conservative accounting information. The International Journal of Accounting 44 (2009), pp. 219-238.

KYLE, Albert S. (1985): Continuous Auctions and Insider Trading. Econometrica 53 (1985), pp. 1315-1336.

LEV, Baruch / ZAROWIN, Paul (1999): The Boundaries of Financial Reporting and How to Extend Them. Journal of Accounting Research 37, pp. 353-385.

LEWELLEN, Jonathan (2015): The Cross-section of Expected Stock Returns. Critical Finance Review 4 (2015), pp. 1-44.

LEWELLEN, Jonathan / NAGEL, Stefan / SHANKEN, Jay (2010): A skeptical appraisal of asset pricing tests. Journal of Financial Economics 96 (2010), pp. 175-194.

LIU, Jing / NISSIM, Doron / THOMAS, Jacob (2002): Equity Valuation Using Multiples. Journal of Accounting Research 40 (2002), pp. 135-172.

LONDON STOCK EXCHANGE GROUP PLC (2016): Industry Classification Benchmark (Equity) – Ground Rules (v2.2), October 2016.

MacKINLAY, A. Craig (1985): Multifactor models do not explain deviations from the CAPM. Journal of Financial Economics 38 (1995), pp. 3-28.

MÖLLS, Sascha H. / STRAUß, Michael (2007): Bewertungsrelevanz der Rechnungslegung - Stand und Implikationen der empirischen Forschung für Aktionäre und Regulierer. Zeitschrift für Betriebswirtschaft 9 (2007), pp. 955-995.

MUNN, Glenn G. (1983): Encyclopedia of Banking and Finance. 8. edition, Boston 1983.

NIETERT, Bernhard / OTTO, Thomas (2018): Vergleichsverfahren der Unternehmensbewertung - eine empirische Untersuchung, issued by Nils CRASSELT / Elmar LUKAS / Sascha MÖLLS / Christian TIMMRECK. 1. edition, Stuttgart 2018, pp. 162-182.

NOFSINGER, John / VARMA, Abhishek (2014): Socially responsible funds and market crises. Journal of Banking & Finance 48 (2014), pp. 180-193.

NOVY-MARX, Robert (2013): The other side of value: The gross profitability premium. Journal of Financial Economics 108 (2013), pp. 1-28.

OHLSON, James A. (1995): Earnings, Book Values, and Dividends in Equity Valuation. Contemporary Accounting Research 11 (1995), pp. 661-687.

OHLSON, James A. (2005): On Accounting-Based Valuation Formulae. Review of Accounting Studies 10 (2005), pp. 323-347.

OHLSON, James A. / JUETTNER-NAUROTH, Beate E. (2005): Expected EPS and EPS Growth as Determinants of Value. Review of Accounting Studies 10 (2005), pp. 349-365.

OUTA, Erick Rading / OZILI, Peterson K. / EISENBERG, Paul (2017): IFRS Convergence and Revisions: Value Relevance of Accounting Information from East Africa. Journal of Accounting in Emerging Economies 7 (2017), pp. 352-368.

PÁSTOR, Ľuboš / STAMBAUGH, Robert F. (2003): Liquidity Risk and Expected Stock Returns. Journal of Political Economy 111 (2003), pp. 642-685.

PEASNELL, K. / YIN, Y. (2014): How analysts make stock recommendations and take account of risk in doing so. American Accounting Association Annual Meeting conference paper, Atlanta 2014.

PEEMÖLLER, Volker H. / MEISTER, Jan M. / BECKMANN, Christoph (2002): Der Multiplikatoransatz als eigenständiges Verfahren in der Unternehmensbewertung. Finanz Betrieb 4 (2002), pp. 197-209.

PENMAN, Stephen H. / REGGIANI, Francesco / RICHARDS, Scott A. / TUNA, A. Irem (2018): A framework for identifying accounting characteristics for asset pricing models, with an evaluation of book-to-price. European Financial Management 24 (2018), pp. 488-520.

R CORE TEAM (2019): R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna 2019. http://www.R-project.org/

RAO, C. Radhakrishna / TOUTENBURG, Helge / SHALABH / HEUMANN, Christian (2008): Linear Models and Generalizations – Least Squares and Alternatives. 3. edition, Berlin/Heidelberg 2008.

RAO, Chaitanya (2012): The mean of a random variable in terms of its cdf. July 2012.
https://ckrao.wordpress.com/2012/07/18/the-mean-of-a-random-variable-in-terms-of-its-cdf

RAPACH, David / ZHOU, Guofu (2013): Forecasting Stock Returns, issued by Graham ELLIOTT / Allan TIMMERMANN. 2A. edition, Oxford 2013, pp. 328-383.

REGULATION (EU) No 236/2012 of the European Parliament and of the Council of 14 March 2012 on short selling and certain aspects of credit default swaps.

RENNEBOOG, Luc / TER HORST, Jenke / ZHANG, Chendi (2008): The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. Journal of Corporate Finance 14 (2008), pp. 302-322.

ROSS, Stephen A. / WESTERFIELD, Randolph W. / JAFFE, Jeffrey F. / ROBERTS, Gordon S. (2015): Corporate Finance. 7. edition, Toronto 2015.

ROSS, Stephen A. / WESTERFIELD, Randolph W. / JORDAN, Bradford D. (2015): Essentials of Corporate Finance. 9. edition, New York 2015.

SAMI, Heibatollah / ZHOU, Haiyan (2004): A comparison of value relevance of accounting information in different segments of the Chinese stock market. The International Journal of Accounting 3 (2004), pp. 403-427.

SCHREINER, Andreas (2007): Equity Valuation Using Multiples: An Empirical Investigation. 1. edition, Diss. Wiesbaden 2007.

SIMKO, Paul J. (1999): Financial Instrument Fair Values and Nonfinancial Firms. Journal of Accounting, Auditing, and Finance 14 (1999), pp. 247-274.

TAN, Hongping / YU, Changju (2018): DCF Please! Evidence from a Content Analysis of Analyst Reports. July 2018. https://scholarspace.manoa.hawaii.edu/bitstream/10125/59350/HARC_2019_paper_192.pdf

THE AMERICAN STATISTICAL ASSOCIATION (2016): Statement on Statistical Significance and P-Values. March 2016. https://www.amstat.org/asa/files/pdfs/P-ValueStatement.pdf

THOMPSON REUTERS (2015): Worldscope Database – Data Definitions Guide (Issue 14.3), 17th Aug 2015.

THOMPSON REUTERS (2016): Thompson Reuters Global Equity Indices – Index Methodology, January 2016.

TSALAVOUTAS, Ioannis / ANDRÉ, Paul / EVANS, Lisa (2012): The transition to IFRS and the value relevance of financial statements in Greece. The British Accounting Review 44 (2012), pp. 262-277.

WELCH, Ivo / GOYAL, Amit (2008): A Comprehensive Look at the Empirical Performance of Equity Premium Prediction. The Review of Financial Studies 21 (2008), pp. 1455-1508.

WILHELM, Jochen / BRÜNING, Lars (1992): Die Fristigkeitsstruktur der Zinssätze: Theoretisches Konstrukt und empirische Evaluierung. Kredit und Kapital, issued by Werner EHRLICHER / Hans-Hermann FRANCKE / Hans-Jacob KRÜMMEL. 1. edition, Berlin 1992, pp. 259-294.

WOOLDRIDGE, Jeffrey M. (2012): Introductory Econometrics: A Modern Approach. 5. edition, Boston 2012.

YAN, Su (2018a): Does the Convergence of IFRS Enhance Value Relevance? Evidence from A- and B-Share Markets in China. April 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3152714

YAN, Su (2018b): IFRS Voluntary Adoption Effect on the Value Relevance of Accounting Information: Evidence from Japanese Listed Firms. August 2018. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3213041

Deutschsprachige Zusammenfassung

Gemeinsame Grundlage aller empirischen rechnungslegungsbasierten Asset-Pricing-Modelle ist der Versuch, die heutigen Assetpreise oder -renditen mit heute beobachtbaren Rechnungslegungsgrößen zu erklären. Technisch gesehen ist das empirische rechnungslegungsbasierte Asset Pricing in der Literatur durch eine Vielzahl statistischer Methoden implementiert: Regressionsansätze, Multiplikatorverfahren und Fehlermaße, was zu mehreren Problemen führt.

Erstes Problem

Angesichts der Tatsache, dass sich Regressionsansätze, Multiplikatorverfahren und Fehlermaße mit der empirischen Preisermittlung von Vermögenswerten befassen, ist die Vielzahl der konzeptionell unterschiedlichen und nicht verbundenen Ansätze verblüffend und wirft zwei Fragen auf:

(i) Wenn Regressionsansätze, Multiplikatorverfahren und Fehlermaße empirisch angewendet werden, können sie zu sehr unterschiedlichen Bewertungsergebnissen führen. Wäre es daher nicht sinnvoll, konzeptionelle Ähnlichkeiten und Unterschiede zwischen diesen statistischen Methoden herauszuarbeiten und sogar eine übergeordnete Kategorie zu finden?

(ii) In Bezug auf Regressionsansätze verwendet die vorhandene Literatur nur eine kleine Teilmenge möglicher statistischer Methoden für die empirische Preisermittlung von Vermögenswerten, d.h. OLS-, WLS- oder Quantilsregressionen. Wäre es nicht vernünftig, diese Untermenge von Regressionsansätzen durch Verwendung anderer Funktionen der Residuen zu vergrößern, z.B. Verwendung einer höheren (und nicht ersten oder zweiten) Ordnung der Absolutwerte der Residuen oder des maximalen Fehlers?
Wäre es in Bezug auf das Multiplikatorverfahren nicht sinnvoll, eine Bewertungsformel zu haben, die verschiedene Methoden der Mittelwertberechnung sowie die Verwendung mehrerer Rechnungslegungsgrößen integrieren kann?

> Wäre es in Bezug auf Fehlermaße nicht sinnvoll, einen Bewertungsrahmen (= Zielfunktion) zu haben, der mit dem Fehlermaß (= Qualitätsbewertung) übereinstimmt?

In Anbetracht dieser Fragen besteht das erste Ziel dieser Arbeit in Kapitel II darin zu analysieren, welche der vorhandenen empirischen Ansätze zur Preisermittlung von Vermögenswerten konzeptionell ähnlich sind, d.h. zu einer übergeordneten Kategorie zusammengefasst werden können und statistische Methoden vorzustellen, die als quasi-natürliche Erweiterungen zu bestehenden empirischen Asset-Pricing-Modellen betrachtet werden können.

<u>Zweites Problem</u>

Basierend auf dieser Übersicht über empirische Asset-Pricing-Modelle und der Literatur kann stark davon ausgegangen werden, dass die gewählten Faktoren (Anzahl und spezifische Auswahl von erklärenden Variablen) sowie die spezifisch verwendete statistische Methode (z.B. OLS-Regression, Quantilsregression) einen wichtigen Einfluss auf die Erklärungskraft einer empirischen Analyse haben. Da die Mehrheit der vorhandenen Arbeiten nur die zuvor erwähnte Erklärungskraft betrifft, kann davon ausgegangen werden, dass sie sich mit der statistischen Signifikanz von Faktoren/spezifischen statistischen Methoden befassen, während die ökonomische Relevanz weitaus weniger analysiert wird.

Da in der Praxis der entscheidende Aspekt von Bewertungsmodellen Preisunterschiede sind und keine statistische Signifikanz, ist eine Analyse ihrer ökonomischen Signifikanz unabdingbar und unumgänglich. Niemand wird einen höheren Preis für ein Unternehmen zahlen, nur weil eine bestimmte Bewertungsmethode einen hohen out-of-sample R^2-Wert ergibt. Darüber hinaus sollten Geschäftsentscheidungen nicht nur darauf beruhen, ob ein p-Wert eine bestimmte Schwelle überschreitet, da die statistische Signifikanz (p-Wert) die Größe eines Effekts oder die Wichtigkeit eines Ergebnisses nicht messen kann.

Daher ist es das zweite Ziel dieser Arbeit in Kapitel III, die ökonomische Bedeutung verschiedener Faktoren/spezifischer statistischer Methoden zu analysieren.

Drittes Problem

Wenn jedoch unterschiedliche Faktoren/spezifische statistische Methoden zu ökonomisch signifikanten Preisunterschieden führen, ist ein Modellauswahlkriterium erforderlich, das auf ökonomischen statt statistischen Kriterien basiert. Während die Arbitrage-Theorie eine allgemeine Richtlinie für die ökonomische Modellbewertung für theoretische Asset-Pricing-Modelle darstellt (d.h. die Preise müssen eine lineare Funktion ihrer zukünftigen Cashflows sein), stützen sich empirische Asset-Pricing-Modelle nicht auf Barwerte der Cashflows, sondern auf angenommene Beziehungen zwischen Rechnungslegungsgrößen/Faktorrenditen und Unternehmenspreisen/-renditen. Aus diesem Grund existieren keine theoretischen Richtlinien in Bezug auf die Komponenten des Modells. Insbesondere gibt es weder Hinweise auf die Anzahl und Art der erklärenden Variablen noch auf den spezifischen statistischen Ansatz.

Angesichts des hohen Bedarfs an einem ökonomischen Bewertungskriterium besteht das dritte Ziel dieser Arbeit in Kapitel IV darin, ein ökonomisches Bewertungskriterium zu entwickeln und eine ökonomische Einordnung verschiedener empirischer Modelle zu erstellen.

Viertes Problem

Aus Sicht der Asset-Pricing-Theorie ist ein solches Modellbewertungskriterium überflüssig, da das richtige Unternehmensbewertungsmodell klar ist: der Barwert zukünftiger Cashflows. Prognosen für die Zukunft sind in der Praxis schwierig und insbesondere die Ermittlung von Abzinsungsfaktoren problematisch. Daher ist es möglicherweise besser, ein theoretisch weniger überzeugendes aber leichter anwendbares Modell zu verwenden—z.B. die Verwendung von Rechnungslegungsgrößen—als ein theoretisch überlegenes aber nicht angemessen implementierbares Modell—Barwert. Die überlegene Praktikabilität bestehender rechnungslegungsbasierter Bewertungen hat jedoch seinen Preis: einer relativ schwache Begründung in der Asset-Pricing-Theorie:

(i) Multiplikatoren

Multiplikatoren argumentieren im Wesentlichen, dass ähnliche Rechnungslegungsgrößen zu ähnlichen Preisen führen sollten.

Probleme aus Sicht der Asset-Pricing-Theorie: Eine solche Bewertung ist zwar intuitiv, wird jedoch nicht durch die Asset-Pricing-/Arbitrage-Theorie gestützt, die besagt: Identische Cashflow-Ströme müssen identische Preise aufweisen. Mit anderen Worten, es gibt drei Unterschiede zwischen Multiplikatoren und Arbitrage-Theorie. Erstens werden Rechnungslegungsgrößen anstelle von Cashflow-Strömen berücksichtigt. Zweitens werden ähnliche statt identische Positionen untersucht. Drittens wird eine Rechnungslegungsgröße als ausreichend angesehen um ein Unternehmen vollständig zu charakterisieren.

(ii) Implementierung von Discounted-Cashflow-Modellen mit Hilfe von Rechnungslegungsgrößen

In der Literatur gibt es Discounted-Cashflow-Modelle, die (Funktionen von) Rechnungslegungsgrößen verwenden um Cashflows, den Restwert und/oder den Abzinsungsfaktor auszudrücken.

Probleme aus Sicht der Asset-Pricing-Theorie: Unabhängig von der konkreten Einbeziehung der Rechnungslegungsgrößen in die Discounted-Cashflow-Modelle können sie nur als Annäherung dienen, d.h. die Modelle enthalten Annahmen, die in der Realität im Allgemeinen nicht zutreffen.

(iii) Empirische rechnungslegungsbasierte Ansätze

Empirische rechnungslegungsbasierte Ansätze erklären Aktienkurse anhand von Rechnungslegungsgrößen.

Probleme aus Sicht der Asset-Pricing-Theorie: Diese empirischen rechnungslegungsbasierten Ansätze gehören zum Bereich der Value-Relevance-Studien und interessieren sich daher nur für die statistische Signifikanz von Rechnungslegungsgrößen, nicht aber für die ökonomische Signifikanz, d.h. sie leiten keine Preisaussagen ab. Grundsätzlich können die Regressionskoeffizienten von Value-Relevance-Studien auch zur Ermittlung von Unternehmenswerten herangezogen werden. Die Preisunterschiede zwischen den verschiedenen Regressionsansätzen sind jedoch groß und diese Modelle haben im Gegensatz zum ökonomischen Prinzip einen schwachen ökonomischen Rückhalt.

Alle diese Probleme unterstreichen den Kompromiss zwischen einer strengen Preisermittlung von Vermögenswerten und der praktischen Umsetzbarkeit von Modellen: Barwertmodelle sind theoretisch überlegen aber ihre praktische Umsetzung in Form von konstanten Abzinsungsfaktoren und Restwertmodellen ist ökonomisch gesehen alles andere als überzeugend. Rechnungslegungsbasierte Modelle zeichnen sich durch eine weniger strenge Anlehnung an die Asset-Pricing-Theorie aus, können jedoch ohne viel von ihrer theoretischen Grundlage zu verlieren implementiert werden. Bessere Asset-Pricing-Modelle zu erhalten bedeutet daher entweder die Implementierung von Barwertmodellen oder die theoretischen Grundlagen von rechnungslegungsbasierten Modellen zu verbessern. Zwei Gründe sprechen für die Verbesserung der theoretischen Grundlagen von rechnungslegungsbasierten Modellen. Einerseits hat die Rechnungslegungsliteratur das Asset-Pricing-Potenzial rechnungslegungsbasierter Modelle bisher nicht voll ausgeschöpft: Sie kann ohne Einbußen bei der Praktikabilität sichtbar gesteigert werden. Andererseits führen rein empirische Modelle immer zu einem Rechtfertigungsproblem: Wer würde für ein Unternehmen einen höheren Preis zahlen, weil Umsatzmultiplikatoren zu höheren Preisen führen als Gewinnmultiplikatoren? Wer würde einen höheren Preis für ein Unternehmen zahlen, weil ein niedrigerer Abzinsungsfaktor für die Gewinne verwendet wird? Wer würde einen höheren Preis für ein Unternehmen zahlen, weil ein empirisches Schätzverfahren mit einem höheren R^2 einen höheren Preis empfiehlt als andere empirische Schätzverfahren?

Daher ist es das vierte Ziel dieser Arbeit in Kapitel V, die Praktikabilität rechnungslegungsbasierter Bewertungsmodelle mit der theoretischen Strenge der Asset-Pricing-Theorie zu verknüpfen.