# Fundamental Indexation for Corporate Bond Markets
## A practical view and a quantitative analysis

*Author*
RONALD SMITS*
Tilburg University
rhsmits@me.com

*University supervisor*
PAVEL CIZEK
Tilburg University

*Company supervisor*
WILLIAM DE VRIES
Kempen Capital Management

**Abstract**

This thesis extends the examination of fundamental indexation outperformance to the corporate bond markets. We follow the methodology as described by Shepherd (2014) with slight adaptions in the regarded universe to analyse robustness. The fundamental indexation methodology takes on additional decisions in the switch of constituent weighing from market prices (outstanding debt) to company fundamentals; those decisions are disentangled one by one. We find that the weighing mechanism is insensitive to most adjustments, only the rebalancing frequency affects performance substantially but outperformance persists in all cases. This raises the idea that market mis-pricings occur in corporate bond markets, both structurally as a consequence of the 'bums problem', and randomly due to market sentiments as found in equity markets. We find that these results are observed most clearly by regarding broader investment universes, indicating that the fundamental approach is particularly good in diversifying risks. Subsequently, we verify whether the fundamental portfolio satisfies the index requirements. We do so by investigating whether the portfolio can be replicated and used for benchmarking purposes, and find that this can be done by a mixed integer linear optimisation approach. Finally, we address the critique that the fundamental index does not provide an equilibrium portfolio and suggest a "hybrid" methodology that uses both price and fundamentals. The approach corrects prices on the basis of six factors that are found by a lasso-approach on the pricing error. The pricing error is modelled prior to the lasso-approach by using the Fama and MacBeth (1973)-approach. Our methodology adds value as it is more likely to attain all index requirements and produces better returns and lower volatility, while also no clear additional risk exposures are found. Hence, direct modelling of the pricing error allows incorporation of more information while also bypassing (a part of) the pricing errors. The thesis concludes by stating the need for a consensus on the notion "index", in order to formalise the index comparisons further.

*The author thanks Kempen Capital Management for the opportunity and help to write this thesis. Moreover, the author received valuable feedback from Research Affiliates.

## Glossary

This list contains the main definitions and notions that are used as general wisdom but that might need proper specification.

**Bums problem:** the presumed phenomenon that bond prices do not fully incorporate for the total amount of outstanding debt of a company at the time of a new issuance. There exists confidence among investors and researchers that bond returns are therefore too low initially (priced too expensive) and afterwards become higher.

**Composite rating:** the average rating of Moody's, Standard&Poor, Fitch, and DBRS. First, each rating is ranked among the distinguished ratings per rating company then the average is rounded to an integer value. More information is in the bond index almanac of Bank of America.

**Debt-weighted index:** the index that weighs constituents based on the outstanding debt, that is the issue size times its price plus accrued interest. This is regarded as the market-capitalisation counterpart for bond markets, also see the documentation of the G0BC-index of Bank of America.

**Equal-weighted index:** the index methodology that assigns equal weights to all bond constituents, *e.g.* if there are 100 bonds, the equal-weighted index assigns all a weight of 1%.

**Fundamental-weighted index:** a specific subset of valuation-indifferent indexes that use company fundamentals in constructing an index.

**Fundamental universe:** the bonds that publish fundamental data such that fundamental weights can be calculated and assigned. It is therefore a subset of all available bonds.

**Index:** a financial index assigning weights among its constituents; for a list of its properties see Section 2.III. It is hence a subset of all possible portfolios, namely the ones that satisfy these properties.

**Market portfolio:** the (theoretically) optimal portfolio, as implied by the Capital Asset Pricing Model, that provides the best trade-off between risk and return attainable for investors.

**Option adjusted spread (OAS):** a bond risk measure found by regarding how much the discount rate curve (the government spot curve) should be shifted such that the price of the bond equals its discounted cash flows. The measure is adjusted for bonds with embedded options in a specific way. Generally, the worst-case scenario is taken in the calculation. For the exact rules see the bond index almanac of Bank of America Merrill Lynch.

**Portfolio:** the notion 'portfolio' is used throughout to describe a financial portfolio, that is, a combination of financial assets that are held by an investor. Any allocation of investable money to assets specifies a particular portfolio.

**Research Affiliates (RAFI):** American research company specialising in fundamental indexes for different asset classes. This paper largely follows the approach of their Shepherd (2014) White paper.

**Valuation-indifferent index:** any index methodology that bypasses (bond) prices in the weight determination.

**Yield:** an alternative pricing measure for bonds. The yield of a bond is the discount rate such that the sum of all its projected cash flows are equal to the price.

## Acknowledgments

In the process of writing this thesis the author received valuable feedback from the Fixed Income team of Kempen Capital Management. Talks with William de Vries, Hans Kamminga, Kim Lubbers, and Emma Weeder provided the necessary practical knowledge of corporate bond markets.

Kempen Capital Management moreover enabled a partnership with Research Affiliates, a world leader in fundamental index research. The author thanks Jason Hsu in sharing his vision on key opportunities for fundamental indexation in corporate bond markets. Ashish Garg helped considerably in the comparison of the first results leading to interesting discussions.

# Contents

**1 Introduction** **6**
1.I Research questions . . . 6
1.II Structure of the paper . . . 6
1.III About the company . . . 6

**2 Literature background** **8**
2.I The Capital Asset Pricing Model and the Market Portfolio . . . 8
2.II Are market-capitalization weighted indexes optimal? . . . 8
2.III What is an index? . . . 9
2.IV The fundamental indexation discussion for equities . . . 10
2.V Fundamental indexation for corporate bonds . . . 12
2.VI The approach of Shepherd (2014) . . . 14

**3 Construction of the bond dataset** **17**
3.I RAFI provided data . . . 17
3.II Data matching . . . 17
3.III Data description . . . 18

**4 Formalisation of the indexation methods** **22**

**5 Analysis of the RAFI findings** **27**
5.I Index performance . . . 27
5.II Index robustness . . . 34

**6 Replication of the RAFI index** **38**
6.I The model . . . 38
6.II Mathematical representation . . . 40
6.III Sampling portfolio performance . . . 44

**7 Towards enhancing the RAFI method** **49**
7.I Modelling the pricing error . . . 49
7.II Explaining the pricing error . . . 53
7.III Our “fundamental indexation” approach . . . 58

**8 Conclusions and recommendations** **62**

**References** **65**

**Appendices** **68**
A Index performance detail . . . 68
B Example of variable replacement . . . 68
C Proof of proposition 6.1 . . . 69
D Matlab implementation of stratified sampling MILP . . . 71
E Overview of fundamentals . . . 73
F Lasso results . . . 75
G Detail of sector exposures . . . 76
H Detail of the under- and over-weights of our method . . . 77

# 1. Introduction

Whereas fundamental indexation for equities has attracted a lot of attention, both academically and in practice, less research exists on other asset classes. While the methodology is yet to be widely accepted, it yields a promising and elegant result; if one wants full exposure to a certain market and growth along with it, traditional indexes based on market prices are not the best available proxy. With many critiques to beat, the fundamental indexation stronghold is its empirical outperformance without additional risk exposures. To give additional evidence to (or to break down) this perception, this thesis examines the fundamental indexation performance in the alternative asset class of corporate bonds.

## 1.I Research questions

The thesis will greatly follow the methodology as introduced by Research Affiliates (RAFI) in the Shepherd (2014) white paper. The research questions of this thesis are:

1. can we find the same results as RAFI on fundamental indexation outperformance for corporate bonds,
2. and, if so, can we get additional insights in what drives this outperformance.

Secondly, we will concentrate on whether the RAFI methodology conforms the properties of an index. In Section 2.III the requirements for an index are introduced and we will examine whether:

3. the RAFI portfolio can be replicated by smaller, investable portfolios,
4. and, finally, whether the method can be improved.

*Hypothesis:* My hypothesis is that the RAFI methodology adds value as it bypasses price in the weight determination and thereby is not affected by market mis-pricing. However, I am anxious that the RAFI method incorporates for too few information and therefore expect undesirable concentration shifts or radical changes in certain exposures. For this reason I do not expect it to be possible to replicate the RAFI portfolio returns, and think that improvements can be found.

## 1.II Structure of the paper

Section 2 gives an overview of the literature on indexation of equities; why are market-capitalisation weighted indexes so broadly seen as the only possible indexes? Subsequently the discussion on fundamental indexation for equities is presented, finally the transition to bond markets is made along with a review of existing research for fundamental indexation in this market. Section 3 describes in detail the steps taken in the data construction process. The indexation methodologies are formulated precisely in Section 4. The performance of the RAFI fundamental indexation is then documented in Section 5 along with an examination of its possible drivers, and robustness checks. In addition, the index properties are further examined in Sections 6 and 7. Finally, in Section 8 we conclude with an overview of the results, answers to the research questions, and provide recommendations.

## 1.III About the company

This thesis was assigned by Kempen Capital Management (KCM), a unique, specialist asset management company with a focus on niche markets; high-yield stocks, fixed income, and funds of hedge funds. During my research I joined the Fixed Income Team of the KCM Department that concentrates on passive strategies in bond markets, both sovereign and corporate. KCM has previously launched a fundamental index product for sovereign bonds; its success is one of the main reasons for exploration of additional opportunities. This shows that for this area the distinction between academics and practice is a thin one; theoretical concepts are to be translated in solid investment products and advices.

## 2. Literature background

The main purpose of this section is to give a better understanding of the academic arrival on fundamental indexation for corporate bonds. First, the origins and main developments of financial investing are described. Next, we present the discussion on market-capitalisation weighted index optimality and introduce the concept of fundamental indexation as alternative. Finally, the transition from equities to corporate bonds is discussed and the additional difficulties and possibilities that arise in this particular asset class.

### 2.I The Capital Asset Pricing Model and the Market Portfolio

For decades the problem of optimally allocating financial assets has been troubling academics and practitioners. Markowitz (1952) suggests regarding the expected (mean) return of a portfolio together with its variance. This mean-variance framework was further developed by Sharpe (1964) and Treynor (1961)[1], to be clarified later by Lintner (1965) resulting in the famous capital asset pricing model (CAPM).

Jensen et al. (1972) summarize the underlying assumptions as all investors (1) choose portfolios based on mean and variance; (2) maximize single period risk-averse utility of final wealth; (3) have the same views on joint probability distributions of asset returns; (4) and have limitless access to a risk-free rate.

The main implication of CAPM is the existence of a “market portfolio” that is mean-variance optimal. Under the above conditions, every asset yields, on average, an excess return (over the risk-free on the regarded period) that is proportional to its co-variance with the market. Deviations are possible due to idiosyncratic (firm-specific) risk and the market portfolio is the one that diversifies these optimally. Hence, at this point the search of finding an optimal trading strategy is reduced to identifying the market portfolio.

While this seems very promising, the last decades show that *ex ante* determination is a daunting task. Approaches like Markowitz (1952) require forecasting future returns and their covariances of many assets. This approach is very sensitive to the underlying data (see *e.g.* Best and Grauer (1991)). More robust methods, like the model proposed by Black and Litterman (1992), overcome this issue but can lead to overly risky portfolios as shown by Da Silva et al. (2009). The existence of the market portfolio is promising, but its construction/composition unclear.

### 2.II Are market-capitalization weighted indexes optimal?

A new discussion came to life when Fama (1965) asserted point (3) from above to be correct beliefs that are moreover based on all available information.[2] This is called the efficient market hypothesis (EMH) and implies that investors are right about prices on average. In fact, any new inflow of information is completely unpredictable leading to a “random walk” of prices.

Markowitz (1959) and Sharpe (1965) show that, under CAPM and EMH, it is optimal to hold market-capitalization weighted (cap-weighted) indexes as they are *nearly* mean-variance

[1] The contribution of Treynor (1961) was unfortunately never published, but Sharpe (1964) states the importance of his work.

[2] In fact, three sub-characterizations are given. One suggests that all information, even insider information, is priced in (strong form). The semi-strong form states that prices are based on all publicly available information, the weak form only considers historic prices. Mathematically the proposition can be stated as: $P_t = E(m_t D_t \mid \Omega_t)$, where $\Omega_t$ denotes all available information, $P_t$ the price, and $D_t$ discounted future cash flows; all regarded at time $t$. Finally, $m_t$ is the (presumably CAPM) pricing kernel that establishes the relationship.

optimal.[3] They argue that one cannot do better than this portfolio. While academics are all but in agreement on this, practice has since used cap-weighted indexes pervasively and turned it into a trillion dollar market.

First questions arose when Jensen et al. (1972) empirically test EMH and conclude to reject its validity, implying that assets are incorrectly priced at least at some times. This attack on EMH need not be problematic says Shleifer (2000). He suggests a weaker version of EMH stating that prices can deviate from their true values, but are random and over time average to it (by the presence of "rational arbitrageurs" as Shleifer (2000) dubs them). As deviations are random, they cannot be exploited and one still cannot outperform cap-weighted indexes (since they are still based on all available information). Fama (1998) shows that empirically there is an even split between supposed over- and under-reactions, which is in line with the proposition of Shleifer (2000).

The new formulation makes tests very hard, if not impossible, because of the "joint hypothesis problem". Any comparison test on observed market prices requires a model for "correct" prices. This leads to ambiguity as it is unclear whether a misfit indicates inefficiency of the market or a bad pricing model. Other approaches are elegant as shown by Malkiel (2003) and Shleifer and Vishny (1997) but lead to opposing conclusions. Malkiel (2003) directly regards the outperformance of 'financial experts' and concludes markets seem efficient as no consistent outperformance is found. While Shleifer and Vishny (1997) show inefficient market behaviour in their famous, specific Royal Dutch/Shell case.

Markowitz (2005) shows what happens if assumptions leading to the CAPM are adjusted. He shows that if assumption (4) from above is replaced by a more realistic one of limited borrowing/lending the market portfolio itself need not even be mean-variance optimal, let alone its supposed proxy portfolio. Summarising, cap-weighted optimality depends directly on CAPM and EMH validity and while both are hard to falsify, we conclude two things: (1) it seems particularly interesting to regard alternative indexes as cap-weighted optimality is doubtful, and (2) outperformance of other indexes provide an alternative route of tests on CAPM and EMH.

### 2.III What is an index?

In order to define an alternative to cap-weighing, one needs a proper definition of what is an index in the first place. As we will see later, much of the discussion on alternative indexes originates from a disagreement on the notion of "index". In fact, as Siegel (2003) shows, indexes have been used over time for different purposes and we introduce his definition as follows.

**Definition 2.1.** A portfolio is moreover an index if and only if it can be used for the following properties, dubbing it a 'triple duty':

(i) as an index fund, meaning that any fund manager must be able to match the holdings of an index and reap the same returns, minus expenses (that tend to be low for indexes),

(ii) as a benchmark for actively managed funds,

(iii) and as a proxy for asset classes in asset allocation. △

Mainly item (iii) distinguishes various investment strategies from indexes whereas for the other points this is harder to conclude. For example, an equally-weighted portfolio would mean

[3] Ultimately the theoretically optimal "market portfolio" should include all assets, including *i.a.* real estate and human capital. With the term "nearly" is meant that cap-weighted indexes are, at least, assumed to be the closest and a sufficient proxy of the market portfolio.

that, if any investor would hold it, all companies are valued the same, regardless size (or cash flows etc.), which is obviously unrealistic. Similarly, for style portfolios, this would mean that large parts of the investment universe are excluded (often decile portfolios are considered meaning 90% is then cutoff) and therefore are not indexes. It does remain, however, hard to say how an index *should* look like. At least, cap-weighing is not the only possibility according to Siegel (2003). He states that any index is a trade-off between:

1. completeness,
2. investability (liquidity),
3. clear, published rules and open governance structure,
4. accurate and complete data,
5. low turnover and related transaction costs.

Siegel (2003) also states acceptability from investors and the presence of derivatives and tradable products as properties. Not only are these conditions harder to define but they are also more or less the consequence of fulfilling the others. Moreover, it seems that any of the latter conditions are directly satisfied if an alternative index is held by all investors (meaning that the alternative index becomes the cap-weighted index). While this seems a somewhat technical reasoning, it does provide insight in what a fundamental index is trying to achieve; in order to become the cap-weighted portfolio it *must* be that intrinsic values of the regarded assets are better reflected. More precisely, this is achieved indirectly as the alternative set of weights leads to new prices. If those prices are a better reflection of the true fundamental values, the index will lead to returns that are more in line with its underlying industry/asset class leading to a higher mean-variance efficiency. This is the motivation of fundamental indexation as introduced by Arnott et al. (2005) and the reason why the reported outperformance (corrected for risk factors) is so promising.

To conclude, we will refer to any index solely based on market prices as the traditional indexing approach. For equities that means constituents are weighted by market-capitalisation (cap-weighing) and in bond markets by the value of the outstanding debt (debt-weighing). Furthermore, we use the term valuation-indifferent weighing for any index that bypasses price as factor, and fundamental indexes as the subset that use company fundamentals specifically (the methodology of Chen et al. (2007) regarding smoothed cap-weights is not seen as fundamental indexation as they partly use prices).

## 2.IV The fundamental indexation discussion for equities

Hsu (2004), Treynor (2005), and Siegel (2006) are the first to analyse cap-weighted indexes mathematically. As motivation they argue the following, if the gap between fair value and price (error) is uncorrelated with the fair value, then it must be positively correlated with price.[4] This implies that higher prices are more overvalued and the cap-weighing suffers from a performance drag. In Arnott et al. (2005) an alternative to cap-weighted indexes is suggested by using fundamentals describing size instead. They report robust superior results and argue this is due

[4]Assume the following (simple, illustrating) model: $P = V + \varepsilon$, with $P$ the price, $V$ the fair value, and $\varepsilon$ the pricing error. Then $\text{cov}(P, \varepsilon) = \text{cov}(V, \varepsilon) + \text{var}(\varepsilon)$. The term $\text{cov}(V, \varepsilon)$ is assumed to be zero, hence $\text{cov}(P, \varepsilon) = \text{var}(\varepsilon) > 0$ and $\rho_{P,\varepsilon} > 0$.

to avoiding the dependence with pricing error, while also keeping an eye on the above properties for indexing.

Their findings have been subject to many different critiques, the most serious one being from Perold (2007). He claims that the model is useless as the performance drag of traditional indexes can only be avoided if one has knowledge of the true fundamental ("fair") values. However, his reasoning is flawed as Dijkstra (2013)[5] points out in a commentary letter:

> The purpose of this note is to show that Perold's result is carried entirely by the log-uniform distribution. In other words, he has identified the singular situation, the unique distribution (and an improper one at that), that allows independence between value and error and between price and error. [...] The fundamental issue in fundamental indexation is therefore not the one identified by Perold, but the existence of fair value and mean reversion of errors.

The latter concern is exactly about the EMH characterization of Shleifer (2000), so Perold (2007) is implicitly distancing from efficient markets. However, as adressed above, the literature on this is divided and certainly no alternative accepted model exists. The claim of Arnott et al. (2005) is still alive, and, more importantly, is backed by empirical findings (see for example, Estrada (2006), Fuller et al. (2014), or Basu and Forbes (2014)).

As announced in Section 2.III, another line of hurdles stumbles upon the notion "index". Blitz and Swinkels (2008) argue that an index should be an equilibrium portfolio and that therefore the only proper index can be a cap-weighted one. Asness (2006) devotes an article to address the "incorrectness" of the word "index" for valuation-indifferent methods. However, the fundamental index of Arnott et al. (2005) does not suffer from many his raised objections, so it might be of interest to reason what happens if all investors hold the fundamental portfolio simultaneously. In this case, prices would change in accordance with the fundamental rules that are used in the index. Hence, the fundamental index holds an alternative pricing mechanism. The new equilibrium might even become a more stable one if, for instance, size is no longer correlated with the pricing error. From this point of view, it is not obvious if the statement of Blitz and Swinkels (2008) is even a critique. Also their other remarks that an index should be a buy-and-hold portfolio and it should not be subject to scrutiny (in defining its construction) are somewhat dubious claims. The first is always satisfied in equilibrium, and for the second, it is not clear why it is not satisfied by the fundamental index (as long as the rules are nicely written down).

I am certainly not the first in stating that any alternative portfolio in equilibrium just leads to changing prices. Estrada (2006) also states that: "and third, not all investors could link their portfolios to this benchmark; an attempt to do so would imply substantial changes in the prices of both stocks." But he overlooks the fact that price changes need not be substantial, and even if so, it is not clear which level of price changes is too high to be just. Valuation-indifferent indexes becoming the new equilibrium portfolio might therefore not be that problematic (at least not as obviously as mentioned). On the other hand, prices based on a selection of some fundamentals (as the fundamental indexes do) do not seem to incorporate enough information. This would be a serious problem with regards to Property (i) an (iii) in Section 2.III. The next section describes whether the problem persists for corporate bond markets.

Ultimately, Estrada (2006) finds that fundamental indexation outperforms cap-weighing but is, in its turn, inferior compared to a "simple" style portfolio. This is not at all worrisome as

[5] I am also aware of the commentary articles of Arnott and Markowitz (2008) and Treynor (2008), but those are nicely generalized by Dijkstra (2013).

a fundamental index is not created as a means of outperformance in general, but just to guide as a better benchmark/index than cap-weighing offers. In fact, the most papers I am aware of criticise the source of the outperformance, rather than the results themselves. The only case that I found with initially troubling empirical results comes from Blitz et al. (2010). They find that the outperformance is sensitive to the times at which the index is rebalanced, however they already provide a solution by suggesting to combine annually rebalanced portfolios rebalanced at different dates. Results are then robust again without increasing turnover.

In general, the outperformance of fundamental indexation is acknowledged but critics argue it to be merely the result of additional value and size exposures (as opposed for instance by Blitz and Swinkels (2008)). This would mean that the index is more risky after all. While it is uncertain whether size and value factors are risk premia or pricing anomalies, Arnott and Markowitz (2008) show that the discussion is the same as the market mis-pricing discussion. Mathematically they show that size and value premia are the result of market inefficiency, which, in turn, is the fundament of valuation-indifferent index outperformance.

Concluding, critiques on fundamental indexation either doubt technical assumptions and requirements that the literature has no consensus on, or stem greatly from a semantic discussion that does not lead to obvious objections. The most serious issue raised is whether a fundamental methodology satisfies the index properties discussed in Section 2.III. This is to be discussed further below. For now, empirical results in different equity environments lead to the same results; fundamental indexation outperforms in any of its characterisations that I found and have mentioned. Even the singular case of seemingly worrisome empirical results of Blitz et al. (2010) is relatively easily addressed. Performance in other asset classes is therefore not only interesting in providing additional insights, but also to see whether outperformance persists.

### 2.V Fundamental indexation for corporate bonds

Arnott et al. (2010) already discuss a transition of fundamental indexation to bond markets. As main arguments for the transition they state that a different asset class can provide additional merit to the benefits of bypassing prices in index construction, by checking whether outperformance persists. Second, they argue that the bond market is particularly interesting as it bypasses the size and value premia that only exist in equity markets. In our eyes these arguments still focus on added value towards equity research, while we see great opportunities for bond indexing in itself as well; whereas mis-pricing in equities is controversial, in bond markets much more consensus exists due to the so-called “bums-problem”. This term is used to describe the phenomenon that a debt-weighted index by definition allocates more assets to large debtors, simply because they have more bonds issued.

It is unclear, however, if prices do not fully incorporate for additionally issued debt and, to my knowledge, no direct tests exist in the literature. Factors that are commonly regarded to influence prices and returns are a bond’s exposure to (1) the interest rate term structure, (2) its issuer’s default probability, (3) its issuer’s credit quality, (4) currency fluctuations, and (5) optionality risks.[6] So, if valuation-indifferent indexes have outperformance after mitigating with respect to the known risk factors, this provides an indirect test on the bums problem.

In addition, analogously to equity markets, market inefficiency may lead to unsystematic mispricings of bonds (on top of the structural bums problem). Artificial demand, as a consequence

[6] This characterisation is much in line with the fixed income return attribution of Bank of America Merrill Lynch as stated in their bond index documentation

of allocation constraints of many investment and insurance companies, and limited liquidity in bond markets might even enlarge the gaps between price and fundamental value. Of course, the flip side of the coin is that these characteristics might make exploitation more difficult.

Previous research on indexing in fixed income markets can be distinguished in two groups: based on valuation-indifferent rules or based on optimisation methods. However, the latter group, as stated in Russo (2013), fall short in proxying for a market. Therefore, they fall out of our scope of what we regard as indexes.

The first valuation-indifferent methodology is the equal-weighted index taking even stakes in all available bonds. While performance is impressive, Kritzman et al. (2010) shows its shortcomings in illiquid markets. Moreover, we already argued in Section 2.III that the equal-weighing methodology is unlikely to be an index. Investigating its performance still yields an interesting interpretation, it shows the behaviour of the average bond.

The second group of indexes is based on fundamental approaches. In 2013 Citigroup launched a corporate bond index based on the white paper of Shepherd (2014).[7] The index uses a company's cash flow and long-term assets in assigning weights. Summarising, debt-weighted indexes in bond markets automatically invest/lend more money to higher debt firms. Taken all else equal, this means allocating assets to high leverage and low cash flow coverage[8] firms. Shepherd (2014) shows that an index on the two alternative factors yields better performance in corporate bond markets. De Jong and Wu (2014) is, by my knowledge, the only other paper (apart from RAFI) that examines a fundamental approach in taking just a company's sales figure. Again, outperformance is found and attributed primarily to avoiding the bums-problem.

Both methodologies, however, show that a fundamental approach requires multiple decisions in replacing the traditional constituent weights. This contrasts with the one-on-one replacement in equity markets. In general, decisions are made on:

1. Factors: which factors determine the alternative set of weights (and how).
2. Assignment: how are alternative weights distributed among the constituents, *e.g.* first on company level and then on bond level. Many papers do not assign weights to individual bonds, as this can lead to extreme weights (see Altman and Saunders (1997) for example).
3. Screening: is a screening applied or not.
4. Rebalancing and updating: how often are weights rebalanced and target weights updated.
5. Universe selection: which bonds constitute the set that are getting assigned weights.

Depending on the factors that are needed for the weight calculation and screening methodology, the bond universe gets affected. As soon as data is unavailable for one of the factors, one should decide whether this bond is neglected or define how to cope with these situations. This is not something that only alternative indexes face, but also the value-weighted index of Bank of America Merrill Lynch Global Investment Grade Index (ticker G0BC) needs a definition on its universe.

**Definition 2.2.** The full universe is defined as all constituents that are in The Bank of America Merrill Lynch Global Corporate Index, hereinafter referred to as G0BC-index. In short, eligible bonds are required to:

[7] `http://www.yieldbook.com/f/m/pdf/citi_indices/20131120_Citi_RAFI_Corp_rulebook.pdf`
[8] The ratio of a firms free cash flow and its debt

(i) be issued in a European or European denoted market,

(ii) have an investment grade rating (based on an average of Moody's, S&P, and Fitch),

(iii) at least 18 months to final maturity at time of issuance,

(iv) at least 1 year remaining term to final maturity,

(v) and satisfy the currency constraints.[9] △

The mathematical concepts of indexes in this thesis will be introduced in Section 4. For now, we regard the G0BC-index as the value-weighted index for corporate bond markets. Moreover, we regard its constituents as the broadest corporate bond universe possible. Hence, we will require any other index to assign weights over the bonds in this universe.

As illustration, De Jong and Wu (2014) take sales figures to define company weights. In more detail, a company's weight is its own sales figure divided by the sum of the sales of all companies in the universe. This weight is then distributed among the bonds of a company by their relative market value. With regards to frequency and timing, they choose to update company weights annually and at the beginning of the year without excluding firms (no screening). For the sake of completeness: the universe is then defined as the G0BC-constituents that have sales data, all other bonds are discarded in advance.

### 2.VI The approach of Shepherd (2014)

The rest of this section describes the approach of Shepherd (2014) qualitatively on the hand of the four decision points introduced above. In Section 4 we provide their mathematical representation.

#### Factors

As alternative weights Shepherd (2014) uses up to five year average cash flow and long term assets. Cash flow is taken as the result of operating activities plus depreciation and amortisation. Since depreciation and amortisation are a proxy for maintenance capital expenditures, the cash flow factor intuitively reflects the free cash flow a firm generates. For long term assets the sum of a company's equity and long term bonds are taken using the most recently reported figures. Reasoning here is that this number adequately reflects the assets that senior bond holders have a claim on. To obtain a single weight from the two factors, Shepherd (2014) suggests dividing a company's cash flow by the total cash flow of all companies, analogously for long term assets, and then taking the average of the two. As soon as one of the measures is unavailable due to data loss, then the other measure determines the full composite weight. To avoid negative weights both factors are first constrained to be minimally equal to 0. Because of this, the weights are normalised again to sum up to 1.

#### Assignment

As discussed above, directly applying alternative weights on the bond level can lead to extreme weights. Shepherd (2014) chooses to first assign the weights on company level, the bonds per company then sum up to that weight and are weighted by their relative face values.

[9] Bank of America Merrill Lynch requires minimum issue sizes per currency, for AUD at least 100 million, CAD 100 million, EUR 250 million, JPY 20 billion, GBP 100 million, USD 250 million. For exact rules see: http://www.mlindex.ml.com/GISPublic/bin/getdoc.asp?fn=G0BC&source=indexrules.

Screening

Shepherd (2014) moreover applies a screening method that takes place before the company weights are assigned. The rational is to filter out companies that are susceptible to downgrade. To do so they make use of three ratios: cash flow to assets, working capital to assets, and sales to assets. All numbers use up to 5 year averages. Leverage is defined as total assets over total debt and is used to magnify each of the ratios (by straight multiplication). The same can be achieved by dividing the cash flow, working capital, and sales numbers directly by total debt, but Shepherd emphasises the economic interpretation by defining the screening in this way. Next, z-scores are calculated per ratio and a final score per company is the average of z-scores. The screening then takes out the bottom 3%. For computational purposes, ratios higher than 100 are removed and the score is calculated over the remaining z-scores.

From the definition of an index, as defined in Section 2.III, the screening seems controversial. The vast exclusion of bonds from the index does not seem in line with an equilibrium bond holding, as this would imply a value of zero in the case that everybody holds the fundamental index. The weighing methodology itself should, from a technical point of view, take care of the bonds that are susceptible to downgrades. Fundamental values should, ideally, represent this — and price for this — rather than assigning a value of zero on the basis of a secondary weighing method. So, we disagree with the screening itself, but in order to fully examine the methodology of Shepherd (2014), we will include it. In Section 5 results without screening are also discussed.

Rebalancing and updating

Finally, a decision needs to be made on the frequency, order, and timing of the factor determination, bond weight calculation, and screening application. Shepherd (2014) filters out companies at the end of January[10] and calculates the composite weights for those companies. The company weights are then allowed to drift until next year, but on bond level the weights are updated at the end of each month including new issues for the screened companies. Hence, bonds of new companies are not included in the RAFI index until January next year.

Universe selection

As discussed above, the methodology of Shepherd (2014) requires data other than price in determining the weights per bond in the index. This means that not all bonds in the G0BC-index can be included and a different sub-universe is regarded. We will call this universe the "fundamental universe" and define it as follows.

**Definition 2.3.** The fundamental universe consists of all constituents of the Bank of America Merrill Lynch Global Corporate Index (ticker G0BC) for which data is available on:

(i) one of the factors that are needed for the weight calculation (so either cash flow data or long term debt data is needed),

(ii) *and* on total debt,

(iii) *and* on cash flow, working capital, or sales in order to obtain a score in the screening procedure. △

This definition concludes the overview of the academic arrival upon fundamental indexation for corporate bonds, and the introduction to the particular approach of Shepherd (2014).

[10] In the paper actually September is taken as main annual rebalancing month, but Shepherd (2014) concludes that January is most convenient in practice.

Summarising, we have seen that cap-weighing of constituents does not necessarily lead to mean-variance efficiency. Moreover, in our definition of an index in Section 2.III there is room for contenders, as cap-weighing is not the only proper index (at least not obviously). Fundamental indexation therefore is, in our eyes, a particularly interesting alternative and research on its performance on corporate bonds is valuable. The next section describes our methodology.

## 3. Construction of the bond dataset

In the previous section we explained that any index assigns weights to constituents (for example stocks or bonds) such that the requirements in Definition 2.1 are satisfied. A fundamental index uses company information to determine those weights. Therefore, this section describes how we have constructed a dataset with bond and company information.

The dataset is used in Sections 5, 6, and 7. As the sections regard different research questions, the precise methodologies will be described in these three sections.

### 3.I RAFI provided data

The bond data is provided by Research Affiliates and consists of the Bank Of America Merrill Lynch Global Corporate Index (ticker G0BC) constituents. The bond data spans the period of January 1997 until July 2015, consisting of end-of-month data. There are 41 different fields, including bond descriptives (ticker/sector/country/etc.) and pricing information (price/accrued interest/total rate of return/etc.) for, in total, 39670 different bonds.

Each security has a unique CUSIP code (US and Canadian securities) or Bloomberg identifier. Research Affiliates has added, per security, Bloomberg fields `ID_BB_ULTIMATE_PARENT_CO` and `ID_BB_COMPANY` that identify the bond's parent and issuer, respectively. Moreover, they added the begin of month weights for computational purposes.

### 3.II Data matching

In order to weigh bonds in accordance with their fundamentals, it is needed to find — per bond and per date — the company backing the debt. We use the following rules to do so:

- If the issuer is a financing vehicle, we take the parent.
- If the issuer is a child company, we take the obligor if it has data or its first parent with available data.

The rational here is that we select the companies that investors regard when assessing a bonds credibility in this way. However, we stumble into difficulties as there is no database that keeps track of backing companies in case of mergers, acquisitions, ticker changes, etc. (hereinafter referred to as "M&A" cases), but merely overwrite the ticker. Many different events potentially leading to ticker changes make things complicated as, *e.g.*, not every acquisition leads to a different company backing the debt. It is possible that debt is not, or only partially, taken over by the acquiring firm. Also for mergers or spin-offs there seems no one-on-one relationship between event and adequate response. We therefore conclude that the best way to go is to check the cases one-by-one.

Other approaches require taking on simplifying assumptions. *E.g.* Shepherd (2014) makes use of a bottom-up reasoning by gathering data of the issuing companies. If data is not available after a certain date they argue this must be an M&A case. They find the corresponding firms and repeat the same reasoning. This simplifying approach is elegant, but situations where debt is taken over by another firm, but both entities continue to exist are ignored as no data gap is identified.

On the other hand, manual checking seems a daunting task as it involves 39670 different bonds. Fortunately, we can first narrow down this number to include only unique couples of final ultimate parents and issuers.[11] We find that the securities are linked to 5778 different issuers, and 3026 different final parents. As the issuer and final parent, obviously, do not change we can regard the number of unique combinations of the two, leaving us at 6226 cases. The only assumption we make here is that companies with the same issuer and same final parent follow the same ticker trajectory (same order of obligor changes).

We are aware that the deductive reasoning requires the data to be increasingly accurate. *E.g.* if one CUSIP links to the wrong company then only 1 out of 39670 consists of wrong data, barely visible in statistical analysis. However, if 1 out of 6226 companies is linked to the wrong data, possibly much larger part of the dataset is flawed as many securities could link to this company. We therefore run numerous tests finding that for 3 bonds the CUSIP code was wrong, and 2311 bonds were linked to the wrong issuer/ultimate parent couple.[12]

The 2311 wrong bonds correspond to 114 unique combinations of issuer and parent, of which 7 were not in the set before. As a result we have 6233 cases to be checked for which we manually find 4278 different obligor histories (including 4259 different companies). For illustration, 347 obligor changes took place in the period of 1997-2015.

For all trajectories we extract data from Bloomberg on 40 different fundamentals, see the second column in Table 14 in Appendix E. Then, we calculate up to 5-year averages per fundamental and link both the averages and the latest observation to all bonds, per date. All fields include all the data that is available at the beginning of the month, so it can be used when constructing fundamental weights.

As stated, the used methodology has an implicit assumption that securities linked to the same issuer-parent combination follow the same obligor history. We test whether this is fair by going through the data and comparing the historic tickers from the `TICKER` field (this field can be used for comparison but, unfortunately, does not include enough information to use it for data extraction). In a control group of 400 cases we find no susceptible differences in ticker histories (some bonds report a ticker change 1 month later but we do not deem this problematic) and conclude that the data matching is satisfactory.

### 3.III Data description

As the final result, we have a dataset with 1783760 rows and 121 columns (the 41 RAFI provided columns, and the most recent and 5-year-averages of 40 fundamentals). Each row corresponds to one bond on a certain date, and the columns contain descriptive data, pricing data (both RAFI provided), and fundamental data (manually matched). The fundamental data however is incomplete either due to the company being private/non-listed, or in some cases due to complex debt structures, *e.g.* when two companies split the debt of a bond and it is unclear which fundamentals should be taken. In our extensive data matching we found that we were able to find fundamentals for 33389 bonds of the 39670, corresponding to a data loss of 16%. Hence, the resulting bonds are the ones that satisfy the requirements of Definition 2.3. For comparison, De Jong and Wu (2014) report a data loss of 4% by only taking sales figures, and Arnott et al. (2010) find the same data loss of 16%.

Figure 1 gives an idea what the data loss entails. On the horizontal axis all unique bonds are numbered after sorting on the G0BC-index inception date, and then grouped per country; US first, then Canada, Australia, Europe, and Japan. The vertical axis corresponds to the dates in

[11] We use the fields `ID_BB_COMPANY` and `ID_BB_ULTIMATE_PARENT_CO` from the RAFI data in order to do so.

[12] After checking this with RAFI, they made an assumption that the first 6 characters of a CUSIP identify the same company to minimise data downloads.

**Figure 1:** A contour plot of the fundamental data availability. The vertical axis corresponds to the dates in our sample (1997-2015), the horizontal axis shows all unique bonds. In grey-scale the darker areas correspond to regions of bonds that have nearly 100% data availability; the grey areas show the data losses; white colour indicates that the bonds are not in the G0BC-index at those points in time. In colour same is expressed by blue, green, and red respectively. In either case the right bar displays the colour scale.

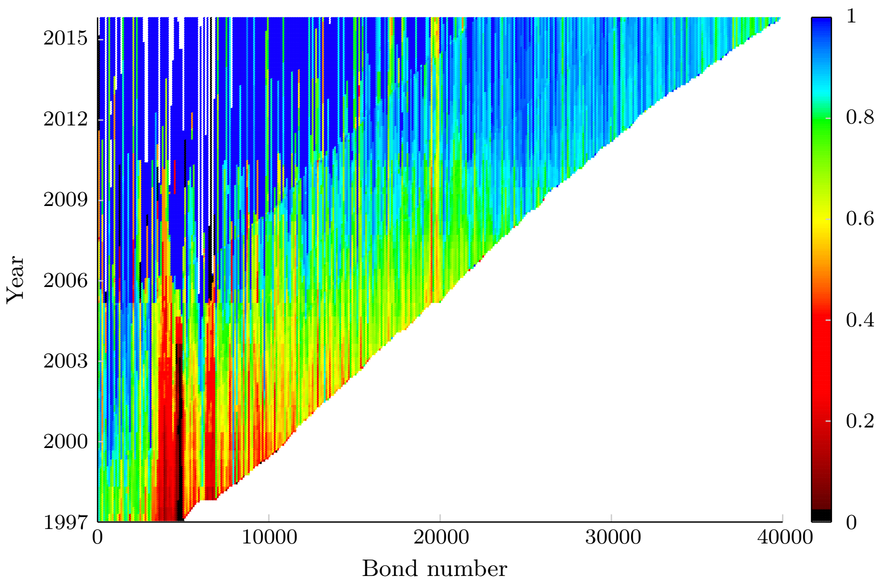



**Table 1:** The descriptive statistics of the bonds in the full BofA Merrill Lynch Global Corporate Bond Index and the ones matched to fundamental data, over the full period January 1997 - June 2015. For the numbers first averages are taken over the constituents per date (begin of month) and then the average, minimum, and maximum are displayed.

| | Full bond universe | | | Fundamental universe | | |
|---|---|---|---|---|---|---|
| Descriptives | Mean | Min. | Max. | Mean | Min. | Max. |
| *Panel A: Equal-weighted:* | | | | | | |
| Coupon | 5.61 | 4.17 | 7.36 | 5.77 | 4.16 | 7.77 |
| Duration | 5.50 | 5.10 | 6.42 | 5.60 | 5.12 | 6.37 |
| Yield | 4.70 | 2.40 | 8.18 | 4.90 | 2.37 | 7.96 |
| OAS (bps) | 139 | 45 | 611 | 144 | 52 | 590 |
| *Panel B: Debt-weighted:* | | | | | | |
| Coupon | 5.56 | 4.22 | 6.95 | 5.80 | 4.23 | 7.75 |
| Duration | 5.40 | 4.85 | 6.48 | 5.72 | 4.85 | 6.43 |
| Yield | 4.53 | 2.38 | 7.76 | 4.86 | 2.35 | 7.70 |
| OAS (bps) | 133 | 40 | 497 | 139 | 47 | 493 |

our sample, January 1997 till July 2015. Dots are drawn for each bond when it occurs in the dataset, due to the sorting this results in a diagonal line beneath which no dots are drawn. More precisely, the colour of a dot displays the data loss in this nearby region (as there are about 40,000 bonds displaying individual dots per company would become unreadable). The bar on the right shows per colour the successful data matching percentage it represents, as stated the white colour indicates the bond was not in the G0BC-index. To clarify, in January 1997 there were 4970 bonds in the index (hence the are no dots at 1997 for the approx. 35,000 other bonds).

**Table 2:** The relative exposures of constituents in the full BofA Merrill Lynch Global Corporate Bond Index and the fundamental universe towards currencies, sectors, and composite ratings are given. The constituents are either equally-weighted or relative to their outstanding debt. The numbers are the average of index weights over time, where the period stretches January 1997 till June 2015 and has begin of month data.

| Descriptives | | Full bond universe Equal | Full bond universe Debt | Fundamental universe Equal | Fundamental universe Debt |
|---|---|---|---|---|---|
| Currency | Australian Dollar | 2.56 | 0.93 | 1.90 | 0.76 |
| | Canadian Dollar | 7.63 | 3.50 | 6.60 | 3.06 |
| | Euro[1] | 15.30 | 22.76 | 13.46 | 20.59 |
| | Great British Pound | 7.31 | 7.01 | 6.40 | 6.14 |
| | Japanese Yen | 11.31 | 10.63 | 8.47 | 5.78 |
| | United States Dollar | 55.02 | 54.19 | 62.57 | 63.11 |
| Sector | Financial | 35.11 | 38.82 | 39.05 | 43.23 |
| | Industrial | 53.58 | 48.58 | 52.43 | 45.92 |
| | Utility | 11.28 | 12.44 | 8.43 | 10.43 |
| | Other[2] | 0.00 | 0.16 | 0.00 | 0.04 |
| Rating | AAA-AA3 | 26.09 | 29.39 | 21.52 | 24.70 |
| | A1-A3 | 41.04 | 41.19 | 43.06 | 43.69 |
| | BBB1-BBB3 | 32.87 | 29.32 | 35.41 | 31.61 |

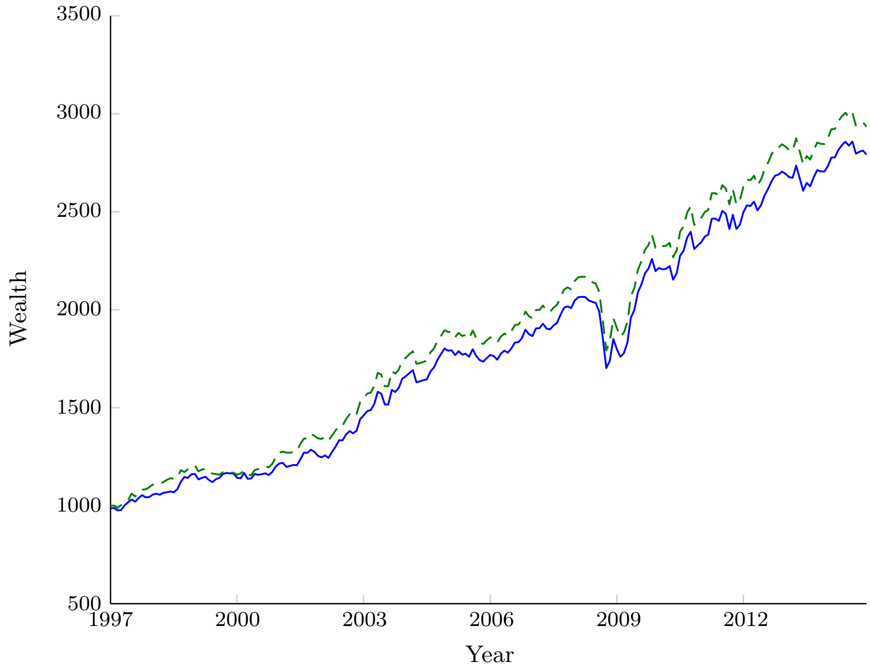
[1] This currency includes the Dutch Guilder, French Frank, German Mark, Italian Lira, and Spanish Peso currencies that appear in the data until December 2004.
[2] Includes the Bloomberg 'Covered', 'Quasi&Foreign Government', and 'Securitized' classes.

The US bonds are matched in about 80% of the cases while Japan is almost fully excluded. Over time we see that more data is available (more blue and green colours) and also newly included bonds have higher matching rates.

Still, the overall data loss has a considerable size (16%) and therefore a transition from the entire bond universe to the ones having fundamental data possibly affects average bond characteristics. Table 1 compares the characteristics of both sets of bonds. By equally-weighing the bonds in the index this gives a comparison of how the average bond in the universe looks like, see panel A. The average bond in the fundamental universe pays a higher coupon, and has higher duration, yield, and option adjusted spread (OAS). The latter three figures are common measurements of the riskiness of a bond. Hence, it means that the average bond in the fundamental universe has more exposure to common bond risk factors. Purely based on this, one would expect higher returns to compensate for the more risky universe. Our finding contrasts with the results of Shepherd (2014), who excludes all private companies (regardless if data for those companies was found). Hence, as stated, this difference in universe construction provides a nice robustness check to their results.

Panel B shows that if bonds are weighted by their debt size the risk exposure shifts persist. This indicates that for the debt-weighted index higher returns (and more volatility) are expected solely by the transition of universes. Outperformance of any alternative index in the new universe should therefore be mitigated with a possible improvement in performance of debt-weighing. By regarding differences in equal-weighing or debt-weighing, one gathers insights in the difference in bond characteristics for large issues compared to small issues (on an average level). With both reasonings in mind, we can focus on the described, and remarkable, finding that the universe becomes slightly more risky by only including companies with fundamental data. At first thought,

**Figure 2:** The solid line displays how a initial money holding in the BofA Merrill Lynch Global Corporate Index has evolved in the period January 1997 - June 2015. The dashed line shows how the same initial holding would have evolved by only debt-weighing the bonds for which fundamental data was found.




one might expect the universe to become less risky as larger firms (on average) are selected, since these are the ones that are more likely to report figures. However, we find in the data for the fundamental universe (by comparing figures between panels A and B) that the larger companies are also the ones that issue longer term bonds (increasing riskiness) and ones with a higher coupon, this combination gives the observed changes in the average universes. To clarify, larger bond issues are weighted more in the debt weighted index (with respect to a "1/N" weight) and we see that the duration increases from 5.60 to 5.72 by taking size into account, the (weighted) average coupon increases slightly to 5.80.

Table 2 investigates whether the fundamental universe is also more risky because of concentration shifts in currencies, sectors, and composite ratings. As an example, the average bond has 55.02% exposure towards the United States Dollar, while the debt-weighted average is slightly less exposed with 54.19%. All changes in exposure have the same direction regardless the weighing scheme. In general, for currencies a concentration towards United States Dollars occurs by switching to the fundamental universe, while on sector level more Financials are selected. Finally, the lower rating classes are more presented in the fundamental universe. This latter shift clearly explains the more risky universe, while the more concentrated sector and currency exposures might indicate less diversification gains, this is analysed further in Section 5.

Figure 2 shows how a debt-weighted index performs in the different investment universes. As suspected, the change in bond selection leads to a higher final money holdings, a more detailed comparison is given in Table 3 in Section 5.I. To conclude, care should be taken to compare results of fundamental indexes with the BofA Merrill Lynch index as the universe itself influences returns and volatilities.

## 4. Formalisation of the indexation methods

In Section 2 multiple ways to reconstruct a debt-weighted index are discussed. This section will formalise the definitions of the indexes (or portfolios) that are regarded in the rest of our thesis:

- the debt-weighted index,
- the debt-weighted index in the fundamental universe (as defined in Definition 2.3),
- the fundamental index as defined by Shepherd (2014),
- and the equal-weighted portfolio.

So, in this paper we start off by comparing one valuation-indifferent approach, equal weighing, and the fundamental indexation approach as described by Shepherd (2014).[13] Both approaches are compared with the traditional, value-weighted, indexes over the full investment horizon, and the subset of bonds for which we were able to match fundamental data as described in Section 3.II. We previously stated the fallacy of equal-weighing as index, however, it does provide insights in how the average bond has behaved over time and therefore remains interesting to regard as portfolio.

Before we introduce the indexation approaches mathematically, we summarise the main results of Section 2 that one should keep in mind. Section 2.V describes the transition of indexing in equities to bonds. The most important difference is that the bond universe changes much more rapidly than for equities. Expiring bonds and new issues occur more often than stock splits and stock issues. As a consequence, even a traditional debt-weighted index in bond markets is less passive in the sense that it requires more trades/turnover and hand-in-hand more decision making (*e.g.* when to rebalance, and how to update the target weights). Another difference is that for stocks an alternative weighing scheme is a one-on-one replacement of weighing constituents, while for bonds this might lead to complications. While it is, of course, possible, often different methodologies are applied to avoid occurrence of extreme weights. If a company both has high fundamentals and relatively many bonds outstanding the reweighing directly on bond level indeed can skew indexes considerably. Any alternative index needs additional decisions, hence also fundamental indexation.

Note, that we deviate slightly from the results of Shepherd (2014) already as we took on different decisions in the data matching procedure, see Section 3.II. Additionally, we slightly adjust the long term assets definition (as stated in Section 2.VI) by regarding all long term obligations instead of just bonds. There are three reasons why we do this. First the change is only marginal and has similar economic interpretation; the number reflects the available assets of a company to prevent a default and therefore also assesses the 'debt service capacity'. Moreover, even though a large part of the bonds (80%) is senior in both investment universes (fundamental and full), the long term debt measure seems a more consistent one. Third, the alternative measure provides a nice robustness check. The rest of this section describes how our slight adaptions to the methodology of Shepherd (2014) are defined mathematically.

### Mathematical index representation

This section will describe the mathematical concepts of the index definitions as introduced above. In its most general form an index can be represented by a function $\varphi$ with as range the set of feasible target weights, $\mathbb{W}$. An index moreover defines a set of factors per constituent, $\Omega$,

[13] The Citi RAFI indexes are based on this paper, but use a slightly simplified implementation.

that, together with time, spans the domain. The regarded time horizon is denoted by the set $\mathcal{T}$, which is the continuous interval between 0 and $T$. We will use a subscript $t$ throughout to denote a variable's dependence on time (convention for continuous time is a bracket form $(t)$, but for the sake of readability we will deviate from this).

The set $\mathcal{B}_t$ is used to denote the constituents that are present at a certain date and that demand weight assignment. Its cardinality, the number of constituents, is denoted by $N$ and defines the dimension of the information set (partly) and the target weights (fully). The other dimension of the information set is dependent on the number of regarded factors $F$ in $\Omega$. The index $i$ will be used to refer to the $i$-th element in $\mathcal{B}_t$ and target weights, and to the $i$-th row of the information set. At any time $t$ we will let $\mathcal{B}_t$ consist of the bonds that satisfy the requirements as stated in Section 3.I.

As discussed in section 2.III there are certain properties that the target weights should satisfy in addition to being non-negative and summing up to one. However, those concepts are unclear to define mathematically. For instance, one of the properties is that an index should be a good proxy for its underlying market. As the "true" market development is unknown, defining maximally allowed deviations becomes vague. Ultimately, the index mapping $\varphi$ should map to weights that are within a set $\mathcal{W}$ of allowed weights, where, for now, its precise shape is unknown and debatable. We will assume throughout that our mapped weights will satisfy $\mathcal{W}$ and continue with the formulations of our indexes. The idea is to give additional insights in the optionality that a decision-maker faces in the transition from a traditional bond index to an alternative.

As final note, we will make use of a sub-mapping $f$, that first assigns a non-negative value to each constituent that can be viewed as a factor score, the index function $\varphi$ then turns them into weights. Definition 4.1 then provides our definition of an index.

**Definition 4.1.** Any index defines a mapping $\varphi$ for $N$ constituents, contained in the set $\mathcal{B}_t$, and depending on the factors in $\Omega = \{\omega_1, \ldots, \omega_F\}$ as:

$$
\begin{aligned}
&\varphi \circ f \ : \ \mathcal{T} \times (\mathit{\Omega} \times \mathcal{T}) \to \mathbb{W}, \text{ where } f : \mathcal{T} \times (\mathit{\Omega} \times \mathcal{T}) \to \mathbb{R}_+^N,\ \varphi : \mathbb{R}_+^N \to \mathbb{W}, \text{ and:}\\
&\quad \mathit{\Omega} \in \mathbb{R}^{N \times F},\\
&\quad \mathcal{T} = [0, \ldots, T],\\
&\quad \mathbb{W} = \left\{ w \in [0,1]^N : \textstyle\sum_{i=1}^N w_i = 1 \right\} \cap \mathcal{W},\\
&(t, \Omega_t) \mapsto \varphi(f(t, \Omega_t)) = w^T.
\end{aligned}
$$

So, a factor mapping $f$ has as its domain the Cartesian product (all combinations) of time and an array of factor values per constituent evaluated at time $t$. As range the mapping $f$ assigns a real value in $\mathbb{R}_+^N$ to each constituent. The index mapping $\varphi$ is composed with $f$ and turns the values into target weights $w^T$ between 0 and 1, such that they sum up to 1. △

In the way that Definition 4.1 is formulated it is allowed to take any set of factors in $\mathit{\Omega}$, as long as it has some value for each of the constituents and it is consistent through time (the regarded set of data may not change through time, the function form itself can). A factor mapping then assigns a non-negative value for a constituent $i$, based on the time period $t$ and the information on bond $i$ as contained in the $i$-th row of $\Omega_t$. Finally, the index turns them into weights.

A well-defined index therefore specifies the regarded constituents and required data, the frequency of rebalancing, the needed information, a non-negative factor mapping, and finally a distribution function assigning the weights. An illustration is given by means of the G0BC-index, which is updated monthly. The rebalancing dates are included in $\mathcal{F}$, and hence $\mathcal{F} \subseteq \mathcal{T}$. We will allow different choices for the rebalancing frequency in Definition 4.2 of debt-weighted indexes. So, the G0BC-index is in fact a specific choice of debt-weighing indexes. Recall that the set $\mathcal{B}_t$ holds all constituents based on the G0BC-rules at all times. So, the ones that are regarded by the G0BC-index at some date $t$ other than the rebalancing dates are not simply $\mathcal{B}_t$, but the ones of the last rebalance date.

**Definition 4.2.** In line with Definition 4.1 a debt-weighted index takes $\Omega = \{p, n, a\}$ and assigns target weights to each of the constituents, $i \in \mathcal{B}_t$, in the following way:

$$\varphi_i^D(f^D(t,\Omega_t)) = \begin{cases} \dfrac{f_i^D}{\sum_{j=1}^N f_j^D} & ,\forall i \in \mathcal{B}_{\underline{t}} \\ 0 & , \text{otherwise} \end{cases} \quad , \text{ where:}$$

$$\underline{t} = \max\left(\tau \in \mathcal{F} \ : \ \tau \leq t\right),$$

$$f_i^D(t,\Omega_t) = p_{\underline{t},i} \cdot n_{\underline{t},i} + a_{\underline{t},i}.$$

Hence, a debt-weighted index takes price $p$, the number of issues $n$, and accrued interest $a$ as factors in $\Omega$. All are $N \times 1$ vectors with on the $i$-th index the data of bond $i$ as ordered in $\mathcal{B}_t$. The $\underline{t}$ is defined as the first moment prior to $t$ that was a rebalancing moment (and hence included in $\mathcal{F}$). A value is assigned to each bond by the function $f$, note that the resulting values are indeed non-negative. Finally, the index function $\varphi^D$ normalises them assuring that all weights are between 0 and 1, and sum up to one, so that it is indeed in $\mathbb{W}$. △

Definition 4.2 clarifies what is meant with target weights. Newly issued bonds between two rebalance dates will receive a zero weight, while expired ones can have a positive weight even though they no longer exist. We are not interested in the actual weight movements due to market developments, but solely focus on the behaviour of the index itself.

The first difference in the fundamental index of Shepherd (2014), as described in the previous section, is that it can only include the bonds in $\mathcal{B}_t$ that have fundamental data (also see Section 2.VI). We let $\mathcal{U}_t$ be the set of bonds in $\mathcal{B}_t$ that are in the fundamental universe, its size denoted as $N^u$. The debt-weighted index in the fundamental universe is defined exactly the same as Definition 4.2 but takes $\mathcal{U}_t$ as the constituents.

The second difference is that Shepherd (2014) applies an annual screening on those constituents. As a consequence, only bonds of the companies with fundamental data that survived the screening can be selected throughout the rest of the year. As discussed before, this screening is not uncontroversial for an index methodology. We will therefore define it separately from the fundamental index strategy in Definition 4.3. Here, we require the screening moments to be a subset of the rebalancing moments and denote this as $\mathcal{G} \subseteq \mathcal{F}$.

**Definition 4.3.** The screening, applied at periods $t \in \mathcal{G}$, on the $N^u$ constituents contained in $\mathcal{U}_t$, and their corresponding companies in the array $\mathbb{C}$, is defined as:

$$\mathcal{S} = \left\{ i \in \mathcal{C} : \text{card}\left( \bigcap \left( j \in \mathcal{C} \ : \ s_i > s_j \right) \right) \geq \varepsilon \cdot N^c \right\}, \text{ where:}$$

$$\mathcal{C} = \{\mathbb{C}_i\}_{i \in \{1,\ldots,N^u\}}, \text{ and } N^c = \text{card}(\mathcal{C}),$$

$$\mathcal{I} = \{wc, cf, s, d\},$$

$$r_i(\mathcal{I}) = \left\{ \frac{wc_i}{d_i}, \frac{cf_i}{d_i}, \frac{s_i}{d_i} \right\},$$

$$z_{i,j}(r_{i,j}) = \frac{r_{i,j} - \mu_j^r}{\sigma_j^r},$$
$$s_i(z_{i,j}) = \tfrac{1}{3}\textstyle\sum_{j=1}^{3} z_{i,j}, \forall i \in \mathcal{C} \text{ and } \forall j = \{1,2,3\}.$$

The variables $wc$, $cf$, $s$, and $d$, contained by $\mathcal{I}$, are $N^c \times 1$ vectors of working capital, cash flow, sales, and total debt data per company in $\mathcal{C}$, the unique companies in $\mathbb{C}$. The variable $\mu^r$ contains the column averages of $r$, $\sigma^r$ the standard deviations. The cardinality of a set is denoted by card$(\cdot)$. △

The first operation makes sure that only the unique values of $\mathbb{C}$ are included in $\mathcal{C}$, and hence consists of the unique companies present at time $t$. The screening then takes working capital, cash flow, and sales ratios (over total debt) as factors. For each factor a $z$-score is calculated, and as final score per company, the average is taken. To clarify, the construction of $\mathcal{S}$ can be read as the set of companies in $\mathcal{C}$ for which more than $\epsilon$ percent has a lower final score (the cardinality of the set of companies with a lower score is then more than $\epsilon \cdot N^c$). Shepherd (2014) takes $\epsilon$ equal to 3%. For the sake of clarity the screening mapping is only defined on moments that are in $\mathcal{G}$, and the resulting set $\mathcal{S}_t$ of screened companies is used.

Apart from the screening choice, the fundamental weighing approach first distributes weights among companies, and then over the bonds per company. For clarity we therefore let the function $f$ be a composite function in the following definition:

**Definition 4.4.** Using the notions of Definition 4.1 and the set of screened companies $\mathcal{S}_t$, we can define the mapping $\varphi_i^F$, for every $i \in \mathcal{U}_t$, based on the information in $\Omega^F = \{CF, LTA, FV, \mathbb{C}, r\}$ as:

$$\varphi_i^F(f^F(t,\Omega_t)) = \begin{cases} \dfrac{f_i}{\sum_{j=1}^{N} f_j}, & \forall i \in \mathcal{U}_{\underline{t}} \\ 0 & \text{, otherwise} \end{cases} \text{, where:}$$
$$\hat{\underline{t}} = \max\left(\tau \in \mathcal{F} \ : \ s \le t\right),$$
$$\underline{t} = \max\left(\tau \in \mathcal{G} \ : \ s \le t\right),$$
$$g_i^F(t,\Omega_t) = (1 + r_{t,i}^{\hat{\underline{t}}})\left(\frac{CF_{\hat{\underline{t}},i}}{\sum_{j\in\mathcal{C}} CF_{\hat{\underline{t}},j}} + \frac{LTA_{\hat{\underline{t}},i}}{\sum_{j\in\mathcal{C}} LTA_{\hat{\underline{t}},j}}\right)$$
$$f_i^F(t,\Omega_t) = \begin{cases} g_i^F \cdot \dfrac{FV_i(\underline{t})}{\sum_{j\in\mathbb{C}:\mathbb{C}_j=\mathbb{C}_i} FV_j(\underline{t})} & \text{, if } i \in \mathcal{S}_t \\ 0 & \text{, otherwise} \end{cases}$$

The $\Omega^F$ contains data on the 5-year average cash flow data $CF$, the long term asset data $LTA$, the face value $FV$, company $\mathbb{C}$, and return per bond in $\mathcal{U}_t$. The factors $CF$ and $LTA$ are constrained to be minimally 0 in order to satisfy the non-negative range requirement of the mapping $f$. The $\underline{t}$ is used to denote the first moment prior to $t$ that is in $\mathcal{F}$, $\hat{\underline{t}}$ the first preceding one that is in $\mathcal{G}$. The $r_t^{\hat{\underline{t}}}$ is the return per bond at time $t$ since last screening date. △

Before, we argued in words that the fundamental-weighing approach of Shepherd (2014) requires considerably more decision-making than the debt-weighted index. This can now also be seen mathematically. Most opposing is that the fundamental index, apart from a screening, requires two different rebalancing frequencies. Shepherd (2014) calculates company target weights only once per year, as indicated by the set $\mathcal{G}$ and index $\hat{\underline{t}}$. Then, on monthly basis the company weights are allowed to drift by market movements and are distributed among the bonds issued by those companies that are still present (or newly issued).

Summarising a decision-maker should specify: which factors are regarded in the company weight assignment and which ones on bond level, in what way they define factor scores, when they are updated, and whether they are allowed to drift. For the screening one should choose: when to apply the screening, the factors for screening, how they lead to a score per company, and how many are discarded. The next section will investigate first whether the specific choices of Shepherd (2014) outperform the debt-weighted index, moreover it is examined whether changes in the methodology lead to differences.

## 5. Analysis of the RAFI findings

This section compares the results of the indexation methods that are discussed in Section 4. First, the results over January 1997-December 2014 are discussed and possible drivers of the performance are analysed. To see whether the found performances are just results by chance we check its robustness by regarding different environments and small changes to the index construction definitions. In this section we assume that the fundamental indexation approach of Shepherd (2014) satisfies all requirements of Definition 2.1. The next sections will examine whether this assumption is likely to hold.

### 5.I Index performance

Table 3 shows the performance of the value-weighted index and an equal-weighted portfolio in the entire bond universe; for the fundamental universe moreover the RAFI method performance is displayed. To clarify, the value-weighted index in the full universe is the G0BC-index that was introduced above. A graph similar to figure 2 can be found in appendix A, the solid line shows how an initial money holding has evolved for the G0BC-index, the dashed line for the debt-weighted counterpart in the fundamental universe, and the dotted line for the fundamental index. First, we see that the RAFI methodology indeed outperforms the G0BC-index with 55 basis points higher return and 30 basis points lower volatility, leading to a Sharpe ratio of 0.69 versus 0.57. A higher return is to be expected as the fundamental index has higher exposure to risk factors; the average held bond trades at a higher option adjusted spread, 144 versus 134, and a yield of 4.86 compared to 4.59. However, if this was the sole reason for higher returns it should go along with an increase in volatility which is not the case, the drop in average duration might indicate an offset in risk exposures. Turnover figures indicate that the alternative indexation is indeed more active but remains acceptable in comparison with the G0BC-index, also tracking error is only marginal and from the point of view of market reflection this is to be expected. Still, the debt-weighted index shows a considerable drop in turnover of about 9% by the universe transition. So, it remains interesting if changes in the methodology can get the fundamental index turnover closer to this figure if one desires this. This will be examined later.

For now, we see that the debt-weighted index in the fundamental universe is also outperformed, albeit less severe. A 26 basis point higher return and 14 basis point decrease in volatility translates into an increase in Sharpe ratio of 0.05. In comparison with the debt-weighted index in the fundamental universe yield figures are roughly similar and an offset between duration and option-adjusted-spread can be observed. Hence, no clear increase in risk exposures can be seen from this table.

The aforementioned outperformances are dubbed so merely by looking at one sample's result figures and observing higher returns with lower volatility. A more reliable source can be obtained from a Jobson and Korkie (1981)-test. The underlying assumed limiting distribution is a decent approximation if market returns are stationary and ergodic (since it requires returns to be independently and identically distributed). The test, however, fails to reject statistical insignificance of the difference in Sharpe ratios even at confidence intervals of just 64%. This means that if both Sharpe ratios are results of the same limiting distribution, that is, with the same mean and variance, a higher or equal difference in Sharpe ratios occurs 36% of the time. In fact, none of the indexes' Sharpe ratio is statistically distinct. As stated in Section 2.III this is not problematic as alternative indexes aim to be better market proxies and possibly performance improvements

**Table 3:** This table reports the historical performance of the indexing methodologies as discussed in Section 2.VI over the period January 1997 - December 2014. The full bond universe refers to the constituents of the Bank of America Merrill Lynch Global Corporate Index (G0BC), the debt-weighted (DW) column displays its performance. For the full universe also an equal-weighted (EW) portfolio is reported. Finally, the fundamental index (FW), as described in Section 2.VI, exists only in the fundamental universe and is compared with both a debt-weighted and equal-weighted counterpart. All numbers are annualised, except for the yield, duration, and OAS figures; those are the average over time of the average held bond per index. Average buys and sells are reported monthly.

| | Full bond universe | | Fundamental universe | | |
|---|---|---|---|---|---|
| Descriptives | DW* | EW | DW | EW | FW |
| Annual return | 5.87 | 6.33 | 6.16 | 6.51 | 6.42 |
| Volatility | 6.34 | 5.99 | 6.18 | 5.87 | 6.04 |
| Sharpe ratio | 0.57 | 0.68 | 0.64 | 0.72 | 0.69 |
| Yield | 4.59 | 4.76 | 4.93 | 4.98 | 4.86 |
| Duration (years) | 5.38 | 5.48 | 5.50 | 5.58 | 5.34 |
| OAS[1] (bps) | 134 | 140 | 139 | 144 | 144 |
| Tracking error[2] | - | 1.44 | - | 1.31 | 0.82 |
| Turnover[3] | 29.47 | 32.51 | 20.79 | 36.36 | 32.51 |
| Monthly buys | 2.63 | 2.78 | 2.24 | 3.55 | 4.76 |
| Monthly sells | 3.12 | 3.31 | 2.10 | 3.00 | 4.22 |

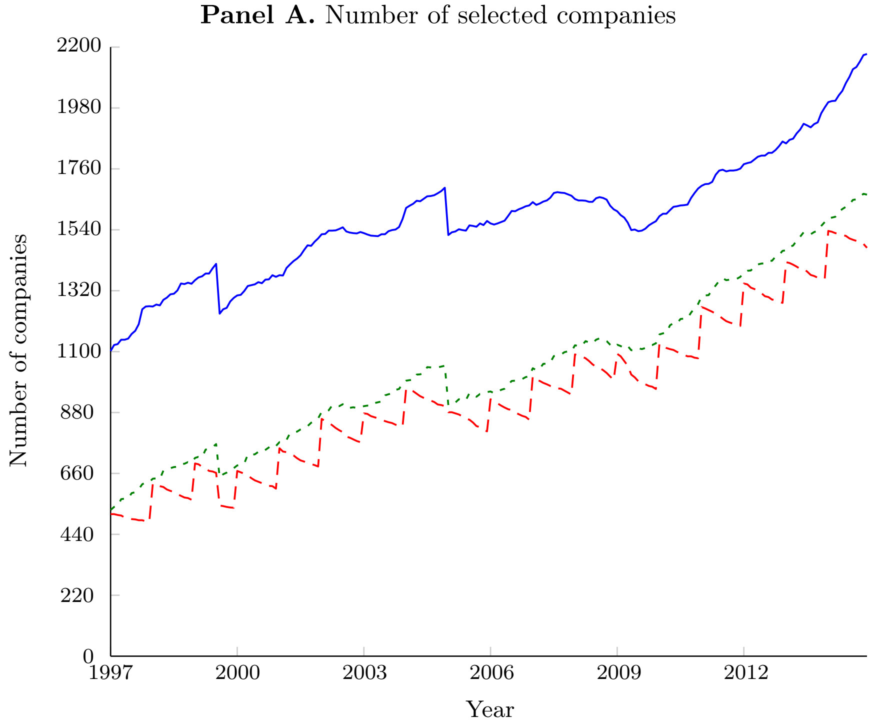
[1] The option adjusted spread is measured versus a basket of duration-matched government bonds.
[2] Calculated as the standard deviation of the difference in annual portfolio returns with respect to the benchmark (debt-weighted portfolio in the same bond universe).
[3] The lesser of monthly purchases and sells divided by the net asset value averaged over time and then annualised.
* Results shown here for the BofA ML GCI can differ (minimally) from the reported figures by Bank of America due to rounding errors in portfolio weights.

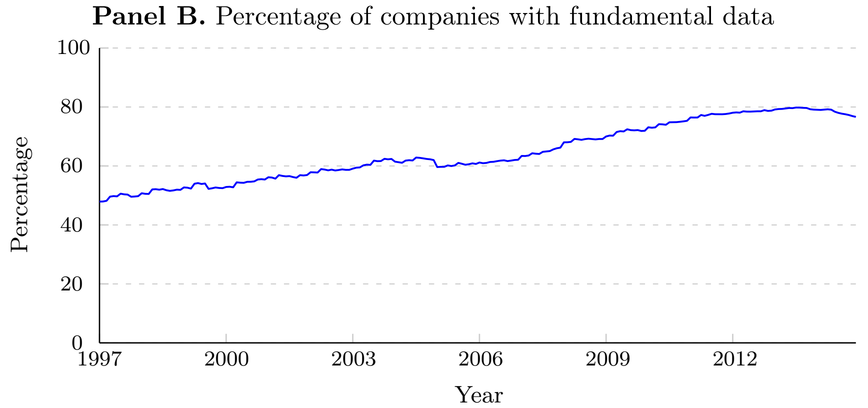
are only able to be small (especially if debt-weighing is already nearly efficient). The supposed outperformance however does need to be persistent and robust. Section 5.II will examine if this is the case, the rest of this section describes in more detail the found results and discusses at this point what drives the outperformance as reported in table 3. After all, if the outperfomance is entirely dragged by the shift in bond universes, the fundamental index would add no value.

#### Universe transition

As suggested in Section 3.III, the shift in investment universes might well affect risk/return measures. We see that this is indeed the case as a debt-weighted strategy on the smaller subset of bonds has better results. Figures 3 and 4 show the details on the total available companies and bonds (solid line) in the G0BC-index, and the ones having fundamental data per date (dotted line). The dashed line indicates the ones selected after the screening. Every January the included companies for the coming year are determined, hence throughout the rest of the year the number of included companies can only decrease. There are two sudden (large) drops observable in the available companies (as well as in the available bonds) at the end of July 1999 and December 2004. At those dates changes were made to the G0BC-index rules, such that the constituent selection became more strict; in 1999 the minimal debt issue size for US Dollar denominated bonds rose to 150 million from 100 million, in 2005 all European non-Euro currency bonds were dropped in addition to US Dollar denoted bonds with a face value of 200 million or less. Moreover, we see a gradual decrease in 2008 due to downgrades and defaults in the credit crisis. Overall, more bonds and companies have made it into the index and data is better available over time.

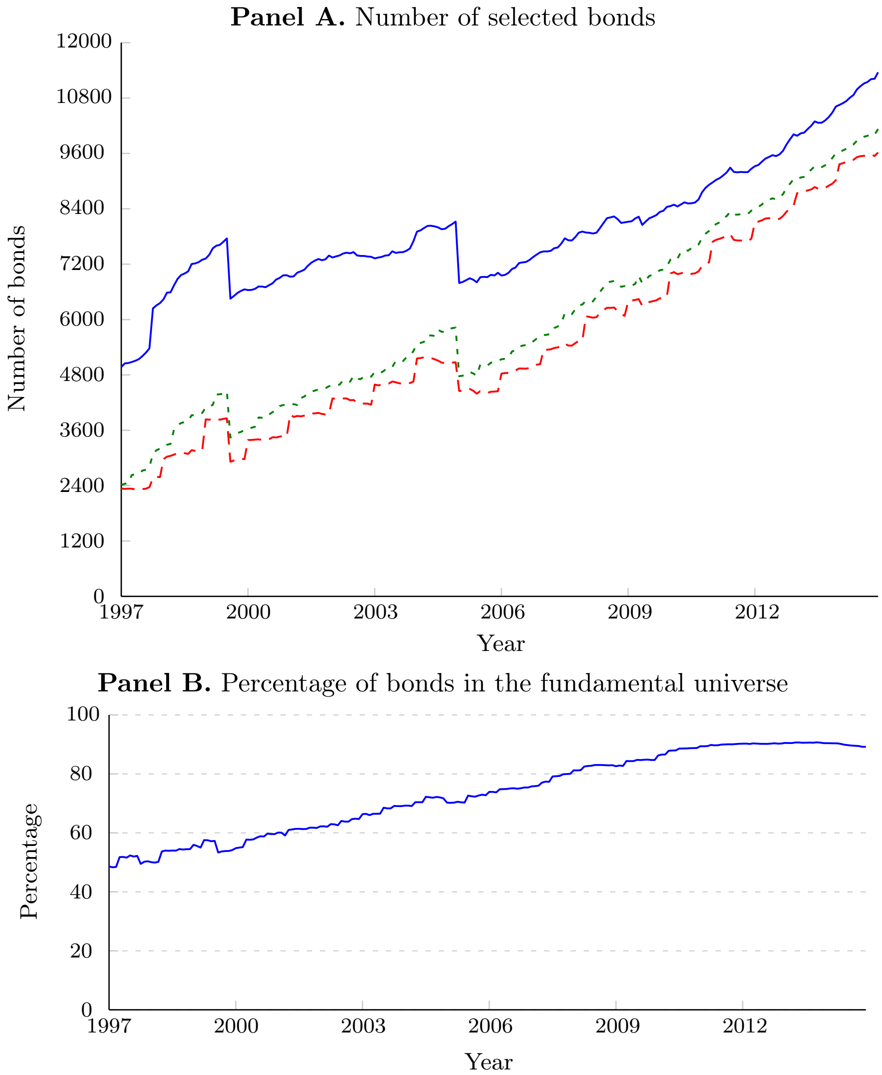
**Figure 3:** Panel A of this figure displays the total number of eligible companies in the benchmark over time (solid line). The dotted line shows the number of companies that are linked to fundamental data, the dashed line shows for the fundamentally-weighted index how many companies are selected after the annual screening procedure. Panel B shows the percentage of companies with fundamental data.




The companies that are selected, and the bonds linked to them, are a selection from the entire universe. Roll (1977) already warned that any test addressing market inefficiency is troubled by the incompleteness of alternative indexes. We acknowledge this problem, but try to solve this in two ways. First, we have included as many bonds as we found data for, in contrast to Shepherd (2014) this means we also include private companies in the universe. Even though we reported a similar data loss, this means that our subset reflects the full bond universe better (as we do not exclude bonds with a certain characteristic). Secondly, we argue that, as it is impossible to regard the fundamental index performance in the full universe, the opposite is possible: regarding a debt-weighted index in the fundamental universe. Any found results can still be influenced by the transition of universes, but if differences are persistent to different settings they seem more

**Figure 4:** Panel A of this figure displays the total number of selected bonds in the benchmark over time (solid line). The dotted line shows the number of bonds that are linked to companies with fundamental data, the dashed line shows for the fundamentally-weighted index how many bonds are selected after the annual screening procedure. Panel B shows the percentage of bonds that are issued by companies with fundamental data.



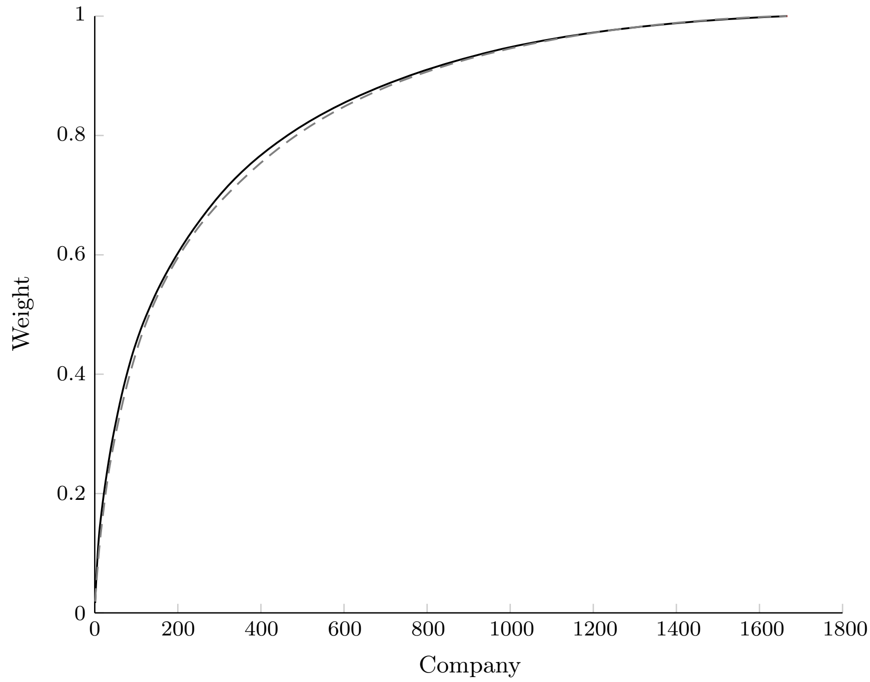

likely to be the result of the different weighing schemes. This is examined in Section 5.II.

Magnitude tilts

It is relevant to examine the distributed weights of the fundamental and debt-weighted indexes. For instance, if the fundamental weighing methodology results in weights that are close to the "1/N"-weights, this would mean it suffers from the same issues as the equal-weighted index. Contrarily, if the index assigns substantial weight to only a small portion of bonds, it arguably more resembles an investment strategy (like the style portfolios for instance) than an index.

Figure 5 therefore shows the cumulative weights at the beginning of December 2014 for the debt-weighted and fundamental-weighted indices, both in the fundamental universe. The weights are first sorted on size, resulting in monotonically increasing graphs. No size bias appears, that is, a similar distribution of weights is allocated to different bonds. However, the figure is for one

**Figure 5:** A plot of the cumulative weights as assigned by the debt-weighted index (solid line) and the fundamental-weighted index (dashed line), weights are first ordered in descending order.




period only, by regarding the weights on a monthly basis we find that the maximum allocated weight is 8.74% on company level for the fundamental index versus a maximum of 4.78% of the debt-weighted index. This raises doubt whether the pattern in figure 5 is consistent over time. Rather than drawing similar figures per date we take the sum of squared weights as measure of weight dispersion and find its average to be 0.61% for the debt-weighted index versus 0.62%. Also, minima and maxima are close to each other 0.34% vs 0.33%, and 0.9% vs. 1.1% respectively. Eventually, we conclude that the outperformance can not be attributed to selective weight allocation.

Exposures

The last section showed that the weighing methodologies allocate similar weights when the entire universe is regarded. It is still possible that the distributed weights behave differently over time, namely when certain sections structurally receive more/less weight *within* the universe. For example, figure 6 investigates how exposures to different industries, on second sector level as classified by Bloomberg, compare. The dashed lines indicate financials, the dotted area corresponds to utilities, and the blank area industrials. As a first observation, we now see that, while weights are distributed in similar magnitude and spread, over time they appear more stable than for the debt-weighted counterpart. This is for two reasons, first the weights are based on fundamentals that are only updated once every quarter (or even annually), and two, because of the annual weight updating and allowing target weights to drift in between. In fact, Shepherd (2014) examined the different frequencies of rebalancing and the effect on performance. He concludes that because of mean reversion and practical convenience annual rebalancing seems the most appealing frequency. Section 5.II will therefore investigate the performance of debt-weighted indexing with annual company weight allocation as well as fundamental index performance under

**Figure 6:** The figure displays the historic allocations to financials (diagonal lines), industrials (blank fill), and utilities (dotted area) for the fundamental-weighted index (panel A) and the debt-weighted index (panel B).

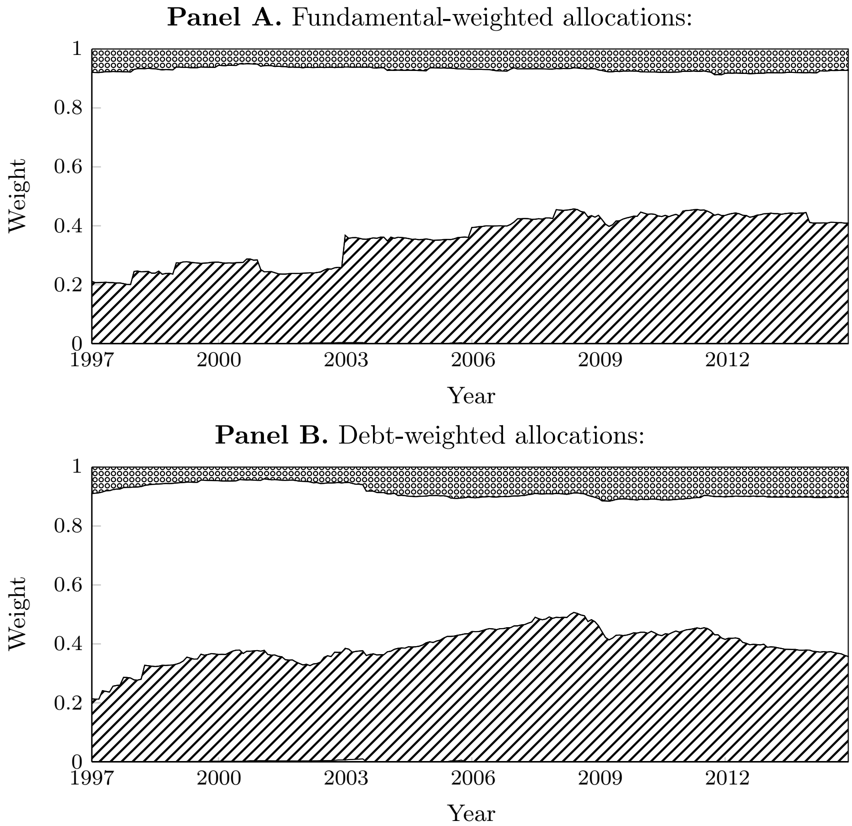
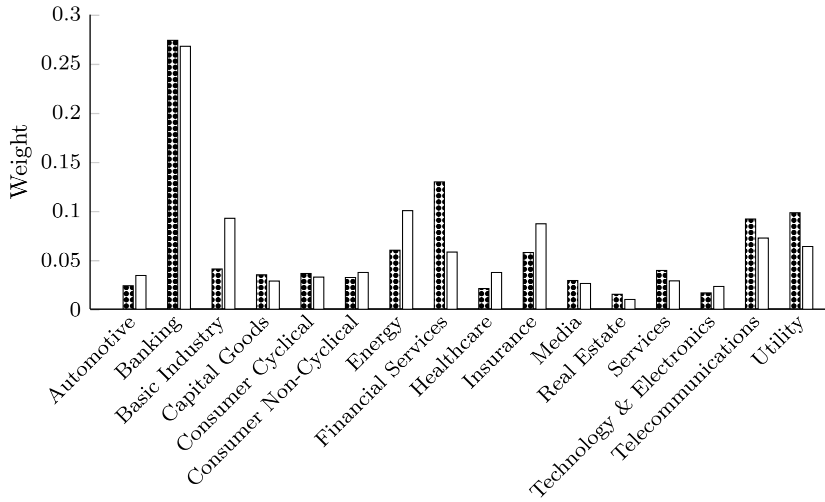


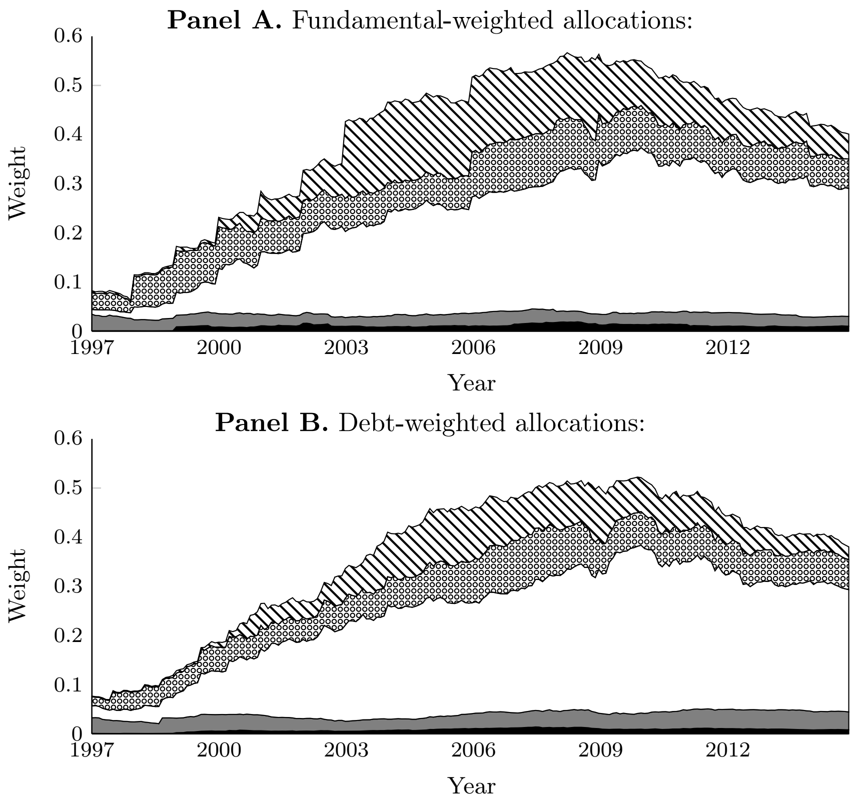

**Figure 7:** Detailed overview of the weights allocated to different sectors in January 2007 (after rebalancing) for the fundamental-weighted index (blank bars), and the debt-weighted index (filled bars).




different rebalancing frequencies.

In more detail figure 7 shows the level three sectors from Bloomberg (16 sectors are distin-

**Figure 8:** Panel A shows the currency exposures over time for the fundamental-weighted index, panel B for the debt-weighted index (for the fundamental universe). The black area shows the weight allocated to Australian Dollar, grey to Canadian Dollar, blank to Euro (before 2005 it includes obsolete European currencies), dots to Great British Pounds, and striped to Japanese Yen. Remaining weights are allocated to United States Dollar denoted bonds.




guished) at one moment in time, namely the post-rebalancing weights of January 2007. This period is chosen as it is the final rebalancing moment preceding the credit crisis in 2008. Whereas figure 6 did not show a clear drop in financial allocation, it can now be seen that especially financial services receive less weight, while basic industries, energy, and insurance are the largest overweights of the fundamental index. The interpretation of this is ambiguous, either the index returns have gained from this beneficial allocation (obviously, bonds issued by the financial sector had very bad returns in the following period) by chance and therefore make the index performance *seemingly* look attractive. Or, the fundamental values are indeed better addressed by the fundamental index. In this case, one could argue that market mis-pricing existed preceding the financial crisis (by overvaluing this sector) and that the fundamental index indeed bypassed this.

Figure 8 shows the exposures per index on currency level. Again we observe sudden increases/decreases in allocation at rebalancing moments for the fundamental index. Relative changes per currency are similar for both indexes, that is exposure to Euros (white area) first increases until mid 2010 and decreases afterwards for both indexes. Possibly outperformance is driven by a larger exposure to Japanese Yen, but this is again subject to the same ambiguity as mentioned above.

Table 4 shows whether outperformance can be attributed to higher exposure towards known bond risk premia. TERM denotes the term structure premia (longer maturities/durations earn

**Table 4:** An overview of the results of three linear regression models: $R_i^e = \alpha + \beta_1\text{TERM} + \beta_2\text{DEF} + \varepsilon$. TERM denotes a factor of average returns on 7-10 year duration Treasuries in excess of the one-month US T-bill. DEF is the default premium factor constructed as the returns on 7-10 year US corporate bonds minus 7-10 year US treasuries. The regression is done on monthly basis, $R_i^e$ are the regressed excess returns; per row is displayed which returns are taken.

| Dependent variable | Intercept | t-statistic | TERM | DEF | $R^2$ |
|---|---|---|---|---|---|
| Debt-weighted | 0.028% | 0.48 | 0.80*** | 0.92*** | 0.78 |
| Fundamental | 0.055% | 0.88 | 0.78*** | 0.86*** | 0.74 |
| Difference | 0.027%*** | 1.94 | −0.02*** | −0.06*** | 0.17 |

Statistical significance levels are displayed as *,**,*** for 90%, 95%, and 99% respectively.

higher returns) and is constructed as the average return on 7-10 year US treasuries minus the one-month T-bill; DEF captures the default premium by regarding the additional returns earned on 7-10 year US corporate bonds with respect to 7-10 year US treasuries. The index returns are regressed on a constant, and the TERM and DEF factors to filter out the systematic risk. A debt-weighted portfolio has more exposure towards the default premium and similar term structure risk; this is in line with the results of RAFI. As they correctly pointed out, this might indicate more efficient market exposure without taking additional risk. While the intercepts on individual level are not statistically significant, a test of the difference in returns shows that the variability is explained less and the difference of 27 basis points is significant even at 95% confidence level. This is a nice result, and emphasises that a fundamentally-weighted index might proxy the market better.

Concluding at this point, the outperformance of RAFI's fundamental-weighing approach seems to stem mostly from differences in sector and currency exposures. The allocations differ in two ways: in absolute terms (Japanese Yen have a higher stake in the fundamental index throughout the period) possibly due to the alternative weighing; and in relative terms (over time) seemingly due to the rebalancing moments (as variability in weights is then observed mostly). Is the fundamental index superior in allocating debt, or are the results merely by chance as only one sample is regarded? This is what the next section will scrutinise.

### 5.II Index robustness

As the first set of robustness checks, slight changes in the RAFI methodology are regarded. Table 5 shows the difference in the annualised return ($R^a$), volatility ($\sigma(R)$), Sharpe ratio (SR), and yield, duration, option adjusted spread figures with respect to the debt-weighted benchmark in the fundamental universe (as reported in table 3). Most strikingly, we see that while none of the Sharpe ratios is statistically different from 0.64 (the debt-weighted index ratio), in all cases the annualised return is improved while only monthly rebalancing leads to a higher volatility. The latter finding is also reported by De Jong and Wu (2014) and shows that outperformance is mainly due to mean-reversion in bond prices as a consequence of the bums-problem. In addition we still find improvements in the Sharpe ratio, meaning that also mis-pricing in corporate bonds seems to exist. In comparison to the results of Shepherd (2014) we find that mean-reversion in bond prices seems to be around two years, while they find 18 months as the average period it takes for the structural overpricing/underpricing in debt-weighing to disappear. Again, this seems rather consistent.

Panel B shows the effects of rebalancing in different months, a critique of Blitz et al. (2010)

**Table 5:** This table displays the performance of slight adaptions to the fundamental-weighted index methodology. $R^a$ denotes average annualised return, $\sigma(R)$ its volatility, and SR the Sharpe ratio. Then yield, duration, and option adjusted spread figures are displayed. Each row shows the basis points difference with the reference (debt-weighted in the same universe) portfolio. For the reference portfolios only OAS is in basis points, rest are percentages.

| Universe changes | $R^a$ | $\sigma(R)$ | SR[†] | yield | Dur. | OAS |
|---|---|---|---|---|---|---|
| *Full universe* | | | | | | |
| Debt-weighted | 6.17 | 6.18 | 0.64 | 4.93 | 5.50 | 140 |
| Fundamental | 24 | -14 | 5 | -5 | -16 | 4 |
| *Panel A: Rebalance frequency* | | | | | | |
| Monthly | 11 | 1 | 1 | -5 | -20 | 8 |
| Quarterly | 22 | -4 | 4 | -7 | -16 | 4 |
| Semi-annually | 20 | -12 | 4 | -7 | -15 | 4 |
| 18-monthly | 28 | -16 | 6 | -6 | -13 | 4 |
| Bi-annually | 28 | -24 | 7 | -6 | -14 | 3 |
| 30-monthly | 24 | -14 | 5 | -5 | -11 | 0 |
| *Panel B: Rebalance month* | | | | | | |
| March | 31 | -8 | 5 | -5 | -16 | 4 |
| July | 24 | -19 | 5 | -7 | -14 | 2 |
| September | 12 | -14 | 3 | -6 | -14 | 3 |
| December | 40 | -26 | 9 | -3 | -13 | 6 |
| *Panel C: Screening percentage* | | | | | | |
| no screening | 22 | -1 | 3 | -7 | -21 | 5 |
| 1% | 22 | -7 | 4 | -7 | -19 | 5 |
| 5% | 25 | -19 | 6 | -7 | -15 | 3 |
| 10% | 26 | -21 | 6 | -7 | -9 | 2 |
| 15% | 26 | -9 | 5 | -4 | -4 | 3 |
| *Panel D: Bond weighing* | | | | | | |
| Equal weighing | 19 | -15 | 6 | -3 | -17 | 7 |
| Issue size | 35 | -5 | 6 | -5 | -16 | 6 |
| Debt-weighted | 17 | -14 | 4 | -6 | -16 | 4 |

[†] Statistical significance of Sharpe ratio difference with the G0BC-index are displayed as *,**,*** for 90%, 95%, and 99% respectively. Results from a Jobson and Korkie-test.

asserts that outperformance vanishes if a different month is taken for rebalancing. We do not find this result, while performance fluctuates, any month (also the seven unreported months) lead to improved Sharpe ratios. The sensitivity towards certain months we feel are the result of calendar sensitivities like window-dressing in December and the publication of new fundamentals per quarter. As we do not find a clear dependence on calendar time, and the Sharpe ratio is improved by at least 3 basis points, it does not seem necessary to spread out the rebalancing throughout the year as suggested by Blitz et al. (2010). Moreover, while it is shown to only increase turnover marginally it does incur more rebalancing moments making the indexation more active.

A more technical decision in the fundamental methodology is the cut off level of 3% in the annual company screening. Panel C shows that even if no screening is applied the Sharpe ratio

**Table 6:** This table displays the performance fundamental-weighted index and a debt-weighted index in different bond universes. $R^a$ denotes average annualised return, $\sigma(R)$ its volatility, and SR the Sharpe ratio. Then yield, duration, and option adjusted spread figures are displayed. Each row shows the basis points difference with the reference (debt-weighted in the same universe) portfolio. For the reference portfolios only OAS is in basis points, rest are percentages.

| Universe changes | Index | $R^a$ | $\sigma(R)$ | SR† | yield | Dur. | OAS |
|---|---|---|---|---|---|---|---|
| *Panel A: Per sector* | | | | | | | |
| Financials | DW | 5.74 | 6.98 | 0.51 | 4.80 | 4.63 | 143 |
| | FW | 29 | 20 | 3 | 7 | 1 | 21 |
| Industrials | DW | 6.56 | 5.85 | 0.74 | 5.05 | 5.98 | 139 |
| | FW | -2 | -18 | 2 | -26 | -36 | -12 |
| Utilities | DW | 6.69 | 6.15 | 0.72 | 4.67 | 6.44 | 118 |
| | FW | 28 | 34 | 1 | 44 | 38 | 11 |
| *Panel B: Currency inclusion* | | | | | | | |
| Non-Japan | DW | 6.39 | 6.45 | 0.65 | 5.16 | 5.57 | 147 |
| | FW | 40 | -1 | 6 | 8 | -9 | 8 |
| US only | DW | 6.40 | 5.30 | 0.78 | 5.34 | 5.95 | 157 |
| | FW | 36 | -34 | 12 ** | 1 | -20 | 4 |
| Europe only | DW | 6.18 | 5.09 | 0.76 | 4.97 | 5.74 | 142 |
| | FW | 27 | -26 | 9 * | -9 | -26 | 3 |
| *Panel C: Economic environment* | | | | | | | |
| 1999-2000 (rising interest rates) | DW | 3.63 | 3.83 | -0.34 | 7.30 | 5.71 | 131 |
| | FW | 31 | -8 | 7 | 0 | -4 | 0 |
| 2001-2002 (falling interest rates) | DW | 10.41 | 4.42 | 1.18 | 6.18 | 5.45 | 174 |
| | FW | 43 | -5 | 13 ** | -5 | -3 | -5 |
| 2003-2006 (rising interest rates) | DW | 4.74 | 4.83 | -0.07 | 4.96 | 5.74 | 101 |
| | FW | 4 | -9 | 1 | -3 | -9 | -5 |
| 2007-2009 (falling interest rates) | DW | 5.50 | 8.83 | 0.11 | 6.19 | 5.99 | 275 |
| | FW | 158 | -93 | 19 ** | 5 | -29 | 9 |

† Statistical significance of Sharpe ratio difference with the G0BC-index are displayed as *,**,*** for 90%, 95%, and 99% respectively. Results from a Jobson and Korkie-test.

is improved by 3 basis points. This means that just the switch to fundamental weights (in combination with annual updating) leads to better results, and the indexation is not merely a bond selecting strategy. We find that in our (wider) universe than regarded in Shepherd (2014) possibly higher cut-off percentage is desired; a tipping point in improvement is around 10% levels, but only if one accepts a screening procedure in the first place. As argumentation, the riskier universe possibly leads to more companies to be discarded in the screening procedure because more bonds have a score below the critical value that Shepherd (2014) found in his less risky universe.

Finally, panel D shows what happens on the bond level if the company weights are no longer assigned by their relative face value, but by equal distribution, or by relative issue size, or by debt-weighing (meaning that the bond weights of the debt-weighted index are used). The first two alternatives lead to even higher Sharpe ratio improvements. These are precisely the two methods that bypass prices completely, possibly indicating that on bond level avoiding the market mis-pricing adds value.

Table 6 investigates whether outperformance persists in different sub-universes. First, panel

A shows if fundamental weighing within sectors is beneficial. We see that on Sharpe ratio level this is the case, but volatility levels are increased substantially in both the financial and utilities industries. It seems that while improvements occur both within and among sectors, the reduction in volatility (due to selection of more credit-worthy bonds) is only obtained when fundamental weights are spread over multiple sectors. This stresses that fundamental indexation is not a selection methodology, it works best over the entire universe.

The second panel is much in line with figure 8. The fundamental index has substantial higher exposure to Japan, exclusion still leads to higher returns and Sharpe ratio but the volatility decrease diminishes. US and Europe only seem interesting as also the debt-weighted index shows better performance, overall we believe that a broader universe leads to more robust performance.[14] Moreover, much of the differences in performance are the result of intra-developments (exchange rates etc.) that are hard to predict.

On the sector and currency levels one can make decisions on certain inclusions/exclusions, more interesting therefore is the performance within different market environments; as unfortunately one cannot control market behaviour. Panel C shows the performance in different market environments by distinguishing periods of rising or falling interest rates, because this is the main driver of bond returns. While total behaviour differs greatly (both indexes perform much better/worse in falling/rising interest rate universes respectively), the relative behaviour remains better for the fundamental index. Still, outperformance greatly stems from worsening interest rates; both in the period of 2001-2002 and 2007-2009 outperformance results even statistically significant Sharpe ratios. Moreover, in all of the periods the yield, duration, and option adjusted spread of the average held bond are lower in general; or seemingly offset by the persistent decreases in duration (as volatility also decreases in all cases).

Overall results are impressive; the methodology gives persistent outperformance despite all examined adjustments (some of which are quite radical). Moreover, the results are resistant to different market environments, that is the results fluctuate but compared to a debt-weighted alternative remain superior. The next section will investigate whether the returns can be replicated by investors using smaller portfolios.

[14]This can be checked by testing different periods on the sub-universes of Europe or US bonds only. We find that performance indeed fluctuates more than in the entire universe.

# 6. Replication of the RAFI index

In this section, we examine the performance of sub-selections of the fundamental index. As stated in section 2.III an index requirement is that matching its holdings should lead to similar results. Moreover, the performance of the sub-selections is of particular interest to investors as these are the more likely portfolios they can hold. Throughout this section we will use the notion "benchmark" to refer to the index, or portfolio,[15] that is to be replicated.

As the number of bonds and companies are large, common sense in the selection procedure becomes hard. For small numbers logical reasoning can exclude certain bonds or pin-point ones that are certainly up for inclusion. Such methods lose their appeal when the number of candidates increases. We therefore examine an option of defining a selection method that automates the procedure. The approach builds on a stratified sampling methodology to determine the characteristics to be matched; the bonds are divided into different subgroups and their properties are mimicked.

The sampling procedure will then optimise at each point in time the set of bonds on the basis of their characteristics (performance criteria), and as inter-dependencies the selected bonds should satisfy the turnover and trading constraints. Purely the matching optimisation might lead to highly different bonds in the subset per month, so the trading restrictions in fact prevent selling huge positions to buy many new bonds. While this is nice, necessity of those constraints would also suggest that performance depends on the initial bond selection (as we cannot adjust it to be too different from last period's holdings). In that way we see the optimisation as two-fold; the interlinked optimisations at each moment after initialisation, and the initial bond selection (that puts restrictions on all following sub-samplings).

## 6.I The model

One should watch out to desire a full optimisation over both parts. It might seem as the desirable action at this point, however the approach is highly dependent on the sample and added value can be scrutinised. Instead, we will first examine straight optimisation in the first period (so ignoring the possible implications of selecting a bond on future restrictions) and then optimising each period after with the turnover constraints. After we will discuss two ways that possibly increase robustness of the results. For now, we focus on a practically implementable method that shows how investors can use the fundamental index as benchmark portfolio.

### Optimisation objective

The objective in a stratified sampling approach is to minimise the differences in characteristics of a selection with respect to the full index. The idea is that tracking error will be minimised in this way. This aim sidesteps the estimation sensitivities that are faced by direct Sharpe-ratio maximisation methods (or other return maximising strategies), as discussed in section 2.I. Question remains, however, if the resulting tracking error is acceptable and robust to other facets.

### Matching criteria

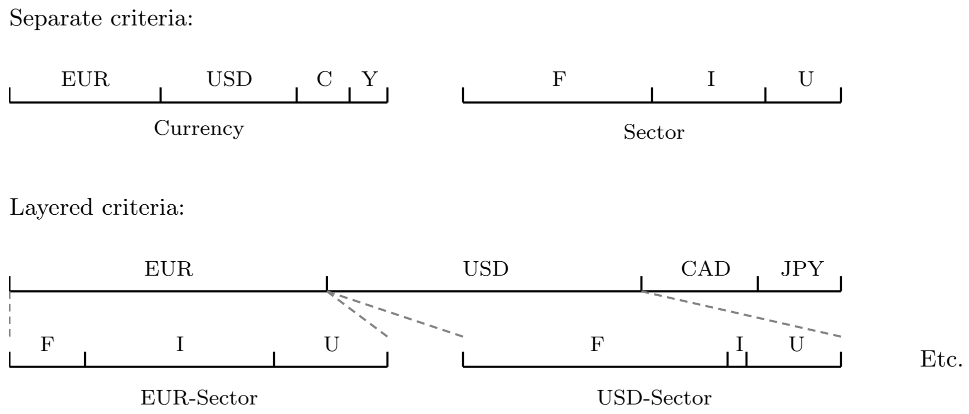
In order to determine the matching characteristics we use implications of the results in section 5.II and a risk analysis. We saw that the fundamental index is particularly good in allocating funds in broad indexes; distributing funds more efficiently on the currency and sector levels. This is something that should be resembled by our sample portfolio. Moreover, we want to control

[15] Our methodology will allow replicating any portfolio, not only indexes.

for differences in duration as it is a main driver of index performance. Combined they provide a nice basis for the homogeneous subgroups, called “strata”.

In general, characteristics can be aimed at separately or in a “layered” fashion. The diagram in figure 9 shows an illustration for the currency and sector distinction; a layered approach is more precise as it sets targets on sectors per currency, while isolating the characteristics sets no restrictions on the within exposures (leaving it possible to *e.g.* select no financials in Euro currency and compensate for this in another currency). The trade-off here is between precision (of matching the index) and restrictiveness (in meeting other constraints, like weights and transaction). In the end, one needs to keep in mind the importance of reaching a requirement in matching the index performance/behavior.

**Figure 9:** Graphical representation of sampling criteria. The non-exclusive strata show how an index can be matched separately on currency exposures (EUR for euro, USD for United States Dollar, CAD or C for Canadian Dollar, and JPY or Y for Japanese Yen) and sector weights (F for financials, I for industrials, and U for utilities). The layered strata show that the sector exposures *per* currency are distinguished.




For the interest rate sensitivities we will apply a layered approach on maturity and (weight-average) duration. Rather than regarding all different maturities we make buckets of 1 to 2 years maturity, and similarly so for 3, 5, 7, 10, 20, 30 years and a last bucket with all maturities over 30 years. Subsequently the durations are averaged by their index weights per basket. So, we want the duration contribution per maturity bucket to be matched. We do the same for currency and sectors as changes in interest rates affect those differently.

Finally, we also want the sector exposures and currency exposures to be matched, separately. We suspect that diversification benefits are mimicked in this way while asserting a similar risk profile.

### Weight constraints

We regard long-only bond portfolios and take $M^+$ as the maximum number of bonds to be selected and $M^-$ the minimum. We acknowledge illiquidity problems for corporate bond markets and therefore put a 5 basis point threshold on the minimum amount to be allocated to an issue with respect to its face value; the maximum stake is 2%. We suspect that this will already make sure that weights are allocated with similar magnitudes and skewness as the benchmark. Still, the next section will show how one should proceed if constraining the absolute difference between portfolio weights with respect to the benchmark is desired.

In the literature we find more suggestions for additional weight constraints (*i.e.* Israel et al. (2015) restrain the maximum number of bonds per company to be less or equal to one) but also keep in mind the famous words of Simon (1965): "If you let me determine the constraints, I don't care who selects the optimisation criterion". In his book, Simon (1965) warns for the fact that, if the feasible set only allows few solutions, the objective only adds marginal value. Therefore we will not further restrict the weights at this point.

Transaction constraints

We will regard the transaction costs as a constraint and an objective. Pure minimisation gives no insight in the trade-off between characteristic matching or trading costs, leaving it possible that index is matched nicely but against undesirable costs. By setting a maximum turnover constraint of 5% we make sure that the first concern is avoided while we leave room for (much) lower turnover in attaining the portfolio risk characteristics by adding this desire in the objective function. Moreover, there is a minimum trade size restriction to avoid relatively high transaction costs on small trades (however, we do not model the transaction costs explicitly).

### 6.II Mathematical representation

We start off by regarding a generalised version of the selection problem. The idea is to show its complexity and the particular difficulties one runs into. In its most general form we reason any benchmark replication optimisation is constructed by a selection of constraints of the following form:

**Problem 6.1.** The problem of selecting at most $M^+$, and at least $M^-$, bonds out of $N$ at periods $t = 1, \ldots, T$ can be described mathematically as:

$$\begin{array}{rll}
\underset{w}{\text{minimise}} \; CD(w) + \tau \cdot TC(w,t) & \text{subject to:} & \\
w_i \geq \alpha \cdot \overline{w}_i^{BM} & \forall i : w_i \geq 0 & \text{(minimum weight)} \\
w_i \leq \beta \cdot \overline{w}_i^{BM} & \forall i & \text{(maximum weight)} \\
\sum_{i \in I_j}^{N} |w_i - \overline{w}_i^{BM}| \leq \gamma & \forall j : w_i \geq 0 & \text{(absolute differences)} \\
p_{i,t} \cdot |w_i - \hat{w}_i| \geq \delta & \forall i \text{ and } t \geq 2 & \text{(minimum trade)} \\
\sum_{i=1}^{N} |w_i - \hat{w}_i| \leq \epsilon & \forall t \geq 2 & \text{(turnover constraint)} \\
\sum_{i=1}^{N} \mathbb{1}_{(w_i > 0)} \leq M^+ & & \text{(maximum selection)} \\
\sum_{i=1}^{N} \mathbb{1}_{(w_i > 0)} \geq M^- & & \text{(minimum selection)} \\
w_i^{BM} / (\sum_{i=1}^{N} w_i^{BM} \cdot \mathbb{1}_{(w_i > 0)}) = \overline{w}_i^{BM} & & \text{(normalised weights)} \\
\sum_{i=1}^{N} w_i = 1 & & \text{(full investment)} \\
w_i \in [0,1] & \forall i. &
\end{array}$$

Here $w$ and $w^{BM}$ are $N \times 1$-vectors of sample portfolio and benchmark weights respectively. A subscript $i$ denotes the bond number, $j$ is used to distinguish bonds per company (or some other criteria). The set $I_j$ holds the indexes $i$ of group $j$. The $\hat{w}_i$ are the sampled weights before rebalancing, and hence its constraints are only included in periods after $t = 1$. The selection constraint and weight normalisation use an indicator function that equals one as soon as the $w_i$ is larger than zero. In the optimisation objective $CD(w)$ (abbreviation of 'characteristic deviations') is a function of deviations from the benchmark characteristics by choosing $w$, and can be interpreted as the disutility that a policy maker faces. The $TC(w,t)$ denotes the transaction costs depending on $w$ (and $t$ as it is only included after starting date). Their exact definitions will follow, $\tau$ specifies the relative importance. Along with $\alpha, \beta, \gamma, \delta, \epsilon$ they form the parameter set to be chosen by the decision-maker. △

To clarify, at the starting period $t = 1$ we minimise the characteristic differences of the sample selection corresponding to the weights $w$. The first restriction states that any weight should be larger than $\alpha$ times the target weight, as soon as it is selected. Similarly the maximal weight and absolute different constraints can be read. For the absolute differences bonds are divided in $j$ groups, the $i$-indexes are contained in $I_j$. If $j$ distinguishes companies the constraint restrict the absolute differences on company weights to be at most $\gamma$. At period $t \geq 2$ we include the minimal trade and turnover constraints and minimise transaction costs. In all cases we should select at most $M^+$ bonds (selection constraint) leading to a lower market exposure than the index. This means that relative weights are higher (the average stake per bond is higher) so all benchmark weights should be rescaled. This is done in the normalised weights constraint; all weights are divided by the sum of selected weights.

In order to ensure feasibility of the problem we need to undertake some steps. While most of the appearing non-linear constraints can be linearised, it is the normalisation of weights that leads to some problems. The idea is clear; if we select only, say, $M$ bonds, the benchmark weights should be reweighed in order to enable comparison. While it is clear why, convexity is not assured as selecting different bonds can change the shape of the optimisation region. Hence, corner points move and the problem possibly becomes unsolvable for branch and bound techniques (if no convergence exists).

A possible solution would be to linearise the constraint by applying Taylor approximation or log-replacements. Both approaches incur precision loss, which is undesirable as bond weights are small and sensitive to small deviations. Another approach is to state a desired level of index exposure prior to the minimisation so that all normalised benchmark weights are fixed. This seems promising, especially as we can decrease the desired exposure level and see the trade-offs between higher exposure versus higher penalties (as portfolio characteristic restrictions are harder to attain). An illustration of this is given in problem 6.2. Here, we took the absolute differences (in comparison with the benchmark weights) constraint on bond level. Moreover, we linearise all other constraints in order to obtain a mixed integer linear problem (MILP).

**Problem 6.2.** The problem of selecting $M$ out of $N$ bonds while demanding a market exposure of $W^o$ at periods $t = 1, \ldots, T$ is described by the following mixed integer linear model:

$$\underset{w,V}{\text{minimise}} \;\; p'\Delta^c + \tau \cdot \iota_N' \Delta^t \qquad \text{subject to:}$$

$$
\begin{aligned}
w_i - c_i &\leq 0 && \forall i && (c') \\
-w_i + c_i &\leq 1 - w^L && \forall i && (c'') \\
-w_i - (1 - c_i) &\leq -\alpha \cdot \overline{w}_i^{BM} && \forall i && \text{(minimum weight)} \\
w_i &\leq \beta \cdot \overline{w}_i^{BM} && \forall i && \text{(maximum weight)} \\
-(w_i - \overline{w}_i^{BM}) - 2(1 - c_i) &\leq \Delta_i^w && \forall i && (i) \\
w_i - \overline{w}_i^{BM} &\leq \Delta_i^w && \forall i && (ii) \\
-c_i &\leq \Delta_i^w && \forall i && (iii) \\
-(w_i - \overline{w}_i^{BM}) - 2(1 - c_i) &\leq -\Delta_i^w && \forall i && (i^-) \\
w_i - \overline{w}_i^{BM} &\leq -\Delta_i^w && \forall i && (ii^-) \\
-c_i &\leq -\Delta_i^w && \forall i && (iii^-) \\
\Delta_i^w &\leq \gamma && \forall i && \text{(absolute differences)}
\end{aligned}
$$

$$
\begin{array}{rll}
-z_i + w_i - \hat{w}_i \leq 0 & \forall i \text{ and } t \geq 2 & (z') \\
z_i - (w_i - \hat{w}_i) \leq 1 - w^L & \forall i \text{ and } t \geq 2 & (z'') \\
-(w_i - \hat{w}_i) - \Delta_i^t \leq 0 & \forall i \text{ and } t \geq 2 & (iv) \\
w_i - \hat{w}_i + \Delta_i^t - 2z_i \leq 0 & \forall i \text{ and } t \geq 2 & (v) \\
-(w_i - \hat{w}_i) + \Delta_i^t - 2(1 - z_i) \leq 0 & \forall i \text{ and } t \geq 2 & (vi) \\
w_i - \hat{w}_i - \Delta_i^t \leq 0 & \forall i \text{ and } t \geq 2 & (vii) \\
-p_i \Delta_i^t - (1 - z_i) \leq -\delta & \forall i \text{ and } t \geq 2 & \text{(minimum trade)} \\
\sum_{i=1}^N \Delta_i^t \leq \epsilon & t \geq 2 & \text{(turnover constraint)} \\
\sum_{i=1}^N c_i \leq M^+ & & \text{(maximum selection)} \\
-\sum_{i=1}^N c_i \leq -M^- & & \text{(minimum selection)} \\
\sum_{i=1}^N w_i \cdot \xi_{i,k} \leq (1 + \Delta_k^c)\overline{\xi}_k & \forall k & \text{(maximum distance)} \\
-\sum_{i=1}^N w_i \cdot \xi_{i,k} \leq -(1 - \Delta_k^c)\overline{\xi}_k & \forall k & \text{(minimum distance)} \\
\sum_{i=1}^N w_i = 1 & & \text{(full investment)} \\
\sum_{i=1}^N w_i^{BM} c_i = W^o & & \text{(demanded exposure)} \\
c_i, z_i \in \{0, 1\} & \forall i & \\
\Delta_i^w, \Delta_i^c, \Delta_i^t, w_i \in [0, 1] & \forall i &
\end{array}
$$

Here $\overline{w}_i^{BM} = w_i^{BM}/W^o$ are constructed outside the optimisation. The optimisation takes place over the weights $w$ and the corresponding auxiliary variables $V = \{\Delta^w, \Delta^t, \Delta^c, c, z\}$. The variables $\Delta^w$ and $\Delta^t$ are $N \times 1$ vectors with deviations in weights with respect to the benchmark and with respect to the ante-rebalanced weights respectively. The vector $\Delta^c$ is of length $C$, the number of characteristic requirements, and contains their deviations. Per criterion $\xi_{i,k}$ denotes the contribution of bond $i$ towards the portfolio characteristic; $\overline{\xi}_k$ is the target value and $\Delta_k^c$ the deviation. The penalties towards violations are given by $p$ ($C \times 1$) and a scalar $\tau$. The variables $c$ and $z$ are binary; $c_i$ equals one when a bond is selected (hence its weight is positive), $z_i$ indicates whether a transaction is made for bond $i$. Finally, $\iota_N$ denotes a $N \times 1$ vector of ones. All other variables are the same as in problem 6.1. △

Problem 6.2 shows the MILP implementation of the original problem 6.1. The build up is as follows: per constraint in the original problem the auxiliary variables are introduced (along with their restrictions) that enable linearisation. First the maximum and minimum weight restrictions are regarded, needing a binary variable $c_i$ that is one if the bond $i$ is chosen. Similarly for the other constraints the auxiliary variables are introduced before the actual constraint is presented. Most replacements of non-linear constraints build on basic programming tricks, see appendix B for an example on the rebalancing weight differences. In order to arrive at the restriction of absolute differences with benchmark weights we make use of the following proposition:

**Proposition 6.1.** The constraints $i$, $ii$, and $iii$ combined with the definitions of the control variables and their negative counterparts $i^-$, $ii^-$, and $iii^-$ in Problem 6.2 are sufficient to have that:

$$
\Delta^w \equiv \left| w_i - \overline{w}_i^{BM} \right| \cdot c_i = \begin{cases} -(w_i - \overline{w}_i^{BM}), & \text{if } w_i \leq \overline{w}_i^{BM} \text{and } w_i > 0 \\ w_i - \overline{w}_i^{BM}, & \text{for } w_i > \overline{w}_i^{BM} \text{and } w_i > 0 \\ 0, & \text{otherwise} \end{cases} . \tag{1}
$$

*Proof.* See appendix C. □

The maximum and minimum distance constraints are constructed in such a way that a particular entry of $\Delta^c$ contains the percentage absolute deviation of this benchmark characteristic. To clarify, if $\xi_1$, the first characteristic constraint, constitutes the financial sector exposure, this is represented by a $1 \times N$-vector of zeros with ones on the entries of the bonds that are issued by financials. The $\bar{\xi}_1$ is in this case the summation of benchmark weights for all financials, or mathematically $\bar{\xi}_1 = \xi_1 \cdot w^{BM}$. For the duration contribution a similar approach suffices, where the one-entries are replaced by the respective durations per bond. In other words, the $i$-th entry equals the duration of bond $i$ if it is in one of the buckets (maturity, sector, or currency), and zero otherwise.

The model provides a nice contribution to the literature by allowing to control the sampling on the weight level as well. As stated, only a slight modification allows to constrain weights on other levels as well by adding up selected weights per company (for example) and comparing with the index company weights. While elaborate and adaptive, the model requires a lot of computation time as the playing field is large and the number of feasible solutions is only limited. This is not a problem as practical implementation needs to run it only once per rebalancing moment, however at this point it is inconvenient for research purposes. We will stick to the problem as defined in Section 6.I, to be presented in Problem 6.3.

**Problem 6.3.** The problem of selecting at most $M^+$ bonds out of $N$, and at least $M^-$, at times $t = 1, \ldots, T$, with an initial money holding of $I$ is:

$$\underset{w,V}{\text{minimise}}\ p'\Delta^c + \tau \cdot \iota_N' \Delta^t \qquad \text{subject to:}$$

$$\begin{aligned}
w_i - c_i &\leq 0 && \forall i && (c')\\
-w_i + c_i &\leq 1 - w^L && \forall i && (c'')\\
-w_i \cdot I &\leq -\alpha \cdot FV_i \cdot c_i && \forall i && \text{(minimum position)}\\
w_i \cdot I &\leq \beta \cdot FV_i && \forall i && \text{(maximum position)}\\
-z_i + w_i - \hat{w}_i &\leq 0 && \forall i \text{ and } t \geq 2 && (z')\\
z_i - (w_i - \hat{w}_i) &\leq 1 - w^L && \forall i \text{ and } t \geq 2 && (z'')\\
-(w_i - \hat{w}_i) - \Delta_i^t &\leq 0 && \forall i \text{ and } t \geq 2 && (iv)\\
w_i - \hat{w}_i + \Delta_i^t - 2z_i &\leq 0 && \forall i \text{ and } t \geq 2 && (v)\\
-(w_i - \hat{w}_i) + \Delta_i^t - 2(1 - z_i) &\leq 0 && \forall i \text{ and } t \geq 2 && (vi)\\
w_i - \hat{w}_i - \Delta_i^t &\leq 0 && \forall i \text{ and } t \geq 2 && (vii)\\
-\Delta_i^t \cdot I + \delta \cdot z_i &\leq 0 && \forall i \text{ and } t \geq 2 && \text{(minimum trade)}\\
\textstyle\sum_{i=1}^N \Delta_i^t &\leq \epsilon && t \geq 2 && \text{(turnover constraint)}\\
\textstyle\sum_{i=1}^N c_i &\leq M^+ && && \text{(maximum selection)}\\
-\textstyle\sum_{i=1}^N c_i &\leq -M^- && && \text{(minimum selection)}\\
\textstyle\sum_{i=1}^N w_i \cdot \xi_{i,k} &\leq (1 + \Delta_k^c)\bar{\xi}_k && \forall k && \text{(maximum distance)}\\
-\textstyle\sum_{i=1}^N w_i \cdot \xi_{i,k} &\leq -(1 - \Delta_k^c)\bar{\xi}_k && \forall k && \text{(minimum distance)}\\
\textstyle\sum_{i=1}^N w_i &= 1 && && \text{(full investment)}\\
c_i, z_i &\in \{0, 1\} && \forall i &&\\
\Delta_i^c, \Delta_i^t, w_i &\in [0, 1] && \forall i &&
\end{aligned}$$

This characterisation of the problem follows the variables as introduced in Problem 6.2. Constraints regarding the weight deviation from the benchmark have been dropped, and the minimum and maximum weight are now with respect to the face value $FV_i$ per bond $i$. The objective market exposure can therefore be dropped. △

The main difference of Problem 6.3 with respect to the previous characterisation is that we drop the desire to match the index weights on bond level as soon as a bond is selected (within certain precision), for the reasons mentioned above. Problem 6.2 reweighed all selected bonds in order to have that the relative benchmark weights of the selection were matched (within certain bandwidths). This reweighing is no longer needed and the computation time decreases drastically as only one optimisation per time moment has to be regarded. We conclude this section with the aspects of our problem design that are open to improvements, namely that:

1. The restriction on the minimally allowed position is made to prevent occurrence of such small positions that they become untradable in practice, and to proxy for the minimal allowed trade size per bond issue (that in practice differs per issue). However, by restricting this to be a constant fraction of the face value, certain bond issues are ruled out entirely, possibly leading to overly strict requirements.

2. Second, the trade increments per issue (the steps with which a trade size can be increased in excess of the minimal trade size) are ignored.

3. Also, the characteristic deviations are now taken relative to the benchmark value and the disutility function of the decision-maker simply adds them up. More adequate function forms can be used that are based on the true tradeoffs a portfolio-maker faces.

4. Similarly, the disutility of making transactions is the same for all bonds (by taking $\iota_N$ in the objective function), a possible improvement is found by taking relative transaction costs per bond into account.

5. With regards to the separate optimisations (the initial one and the subsequent ones) no steps are taken to anticipate on expected transaction costs. Possible improvement can be examined from including expected transaction costs in the bond selection procedure.

6. Finally, the optimisation results in the next section can possibly be further improved by applying the more precise (or, more strict) methodology of Problem 6.2.

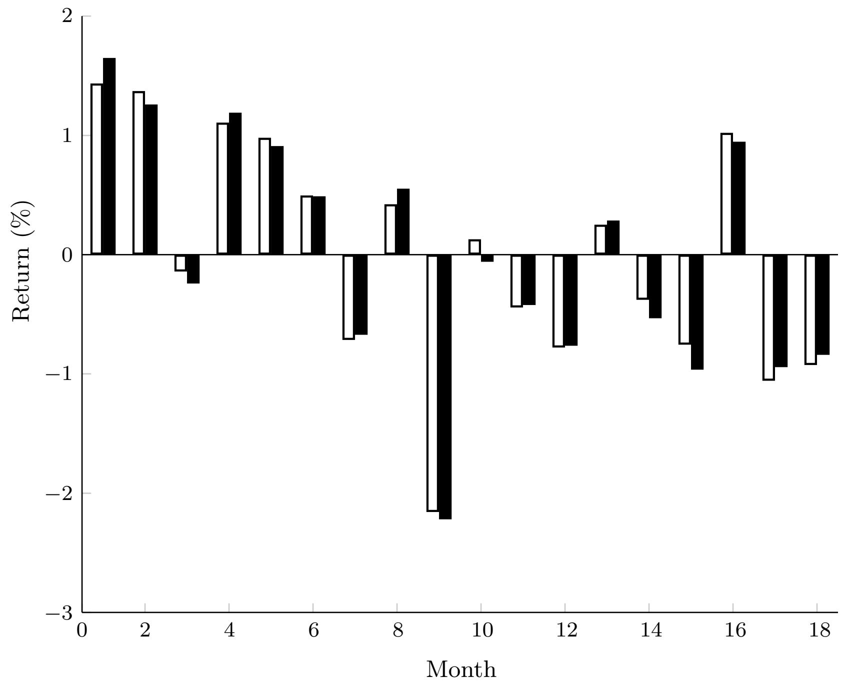
These remarks are left open for future research. Now, as a guiding model, we expect it to already provide useful results.

## 6.III Sampling portfolio performance

As the first set of tests, we run the model of Problem 6.3, that is also described in Section 6.I. We will start off by regarding the time period of January 2014 until June 2015 to see the optimisation performance. Most parameters have already been set in our description of the problem, summarising we take: (1) $\alpha = 5$ bps, the minimal weight of a bond relative to its face value, (2) $\beta = 2\%$, the maximal stake, (3) $w^L = 10^{-7}$, minimal weight as needed for computation purposes, (4) $\epsilon = 5\%$, the maximally allowed turnover per month, (5) $\tau = 2$ and $p$ a vector of ones, to penalise turnover twice as much as deviations in the characteristic matching in the objective function, and (6) $I = 200$ million as investable money.

One of the possible areas for improvements that we discussed in the previous section, is that the true tradeoffs between deviations from the benchmark characteristics and transaction costs are not adequately described by the linear approximation. At the same time, the true tradeoffs are likely to differ between policy-makers. By taking $p$ a vector of ones and $\iota_N$ in the optimisation objective, we specify a case where all characteristic deviations give the same disutility and also

**Figure 10:** The monthly total returns of the benchmark (the fundamental index of RAFI), blank bars, and the sampled portfolio as defined by Problem 6.3, filled bars.




no distinction is made in the transactions for different bonds. In this setting, the only remaining parameter is $\tau$ that compares the sum of character deviations with the sum of transactions. This parameter is clearly dependent on the preferences of a policy-maker. For example, if one is regulated strictly and deviations must be kept low, the $\tau$ is small. We choose a $\tau = 2$ as we expect this to be a reasonable value for the average policy-maker. The results will possibly indicate whether the value should be adjusted; if a policy-maker decides that the characteristic deviations are too high, a lower $\tau$ should be considered.

Remaining for a decision-maker is to determine $M^-$ and $M^+$, the minimum and maximum number of bonds to select. Ultimately, one would like the playing field to be large, enabling better objective values: lower transaction costs and higher matching of the benchmark characteristics. Hence, one would like $M^+$ to be large, and $M^-$ small. However, holding many bonds in combination with our minimal stake would require a large amount of investable assets available. So, this leads to the reasonable insight that the parameters should depend on the available investable amount.

We first find that for reasonable numbers of $I$, *i.e.* investable amounts of more than 100 million, the restrictions of the minimal and maximal allowed position are almost redundant. For instance, the face values of the bond issues present in January 2014 (with a positive weight in the benchmark) are such that we can select minimally 2 bonds and be fully invested, or at most 1778 if we have 200 million available. Hence, we will narrow down the number from practical argumentations. In order to obtain a portfolio that is practically maintainable we allow maximally 300 bonds, and minimally 150 to keep a reasonable playing field while preventing the selection of only few large issues (with the above considerations in mind).

We have now specified all parameters, and find that the playing field allows feasible solutions (as expected from the above reasoning) for the monthly optimisations on January 2014 till July

**Table 7:** Reporting of the simulated stratified sampling performance of the optimisation as described in Problem 6.3, for the period January 2014 till July 2015. In panel A the monthly results are shown, panel B shows the overall performance figures (annualised). The first two columns of panel A display the end of month returns of the benchmark, BM, and the stratified sample, SS. Then the transaction costs $\Delta^t$ of the stratified sample portfolio are displayed. The $\Delta^c$ column shows the corresponding sum of relative characteristic deviations, and $M$ denotes the number of selected bonds, $O$ shows the overlap in bonds selected between two months. In panel B the corresponding annualised returns and volatility are displayed, as well as the Sharpe ratio, Information ratio, tracking error, and total turnover respectively. Tracking error is measured as the standard deviation of the difference in returns between the benchmark and stratified sample. Total turnover is in the total percentage of traded bonds, including the ones that were sold due to expiration. All numbers are percentages except for the tracking error in basis points.

*Panel A: Monthly results*

| | Return | | Stratified sample | | | |
|---|---|---|---|---|---|---|
| Year | BM | SS | $\Delta^t$ | $\Delta^c$ | $M$ | $O$ |
| 2014-01 | 1.44 | 1.65 | 0 | 0 | 167 | – |
| 2014-02 | 1.37 | 1.26 | 5.00 | 42.40 | 150 | 144 |
| 2014-03 | −0.14 | −0.24 | 5.00 | 4.04 | 150 | 123 |
| 2014-04 | 1.11 | 1.19 | 5.00 | 19.78 | 159 | 130 |
| 2014-05 | 0.98 | 0.92 | 5.00 | 28.06 | 150 | 147 |
| 2014-06 | 0.50 | 0.49 | 5.00 | 174.16 | 150 | 142 |
| 2014-07 | −0.72 | −0.67 | 5.00 | 160.55 | 150 | 131 |
| 2014-08 | 0.42 | 0.55 | 5.00 | 142.41 | 150 | 138 |
| 2014-09 | −2.16 | −2.24 | 5.00 | 84.60 | 151 | 148 |
| 2014-10 | 0.13 | −0.05 | 5.00 | 98.82 | 150 | 138 |
| 2014-11 | −0.44 | −0.43 | 5.00 | 64.39 | 151 | 142 |
| 2014-12 | −0.78 | 0.76 | 5.00 | 55.03 | 151 | 142 |
| 2015-01 | 0.25 | 0.28 | 10.00 | 0 | 172 | 112 |
| 2015-02 | −0.38 | −0.54 | 4.39 | 0 | 170 | 151 |
| 2015-03 | −0.75 | −0.97 | 2.58 | 0 | 188 | 153 |
| 2015-04 | 1.02 | 0.94 | 2.87 | 0 | 199 | 176 |
| 2015-05 | −1.06 | −0.94 | 3.22 | 0 | 189 | 174 |
| 2015-06 | −0.93 | −0.84 | 5.00 | 42.90 | 191 | 170 |

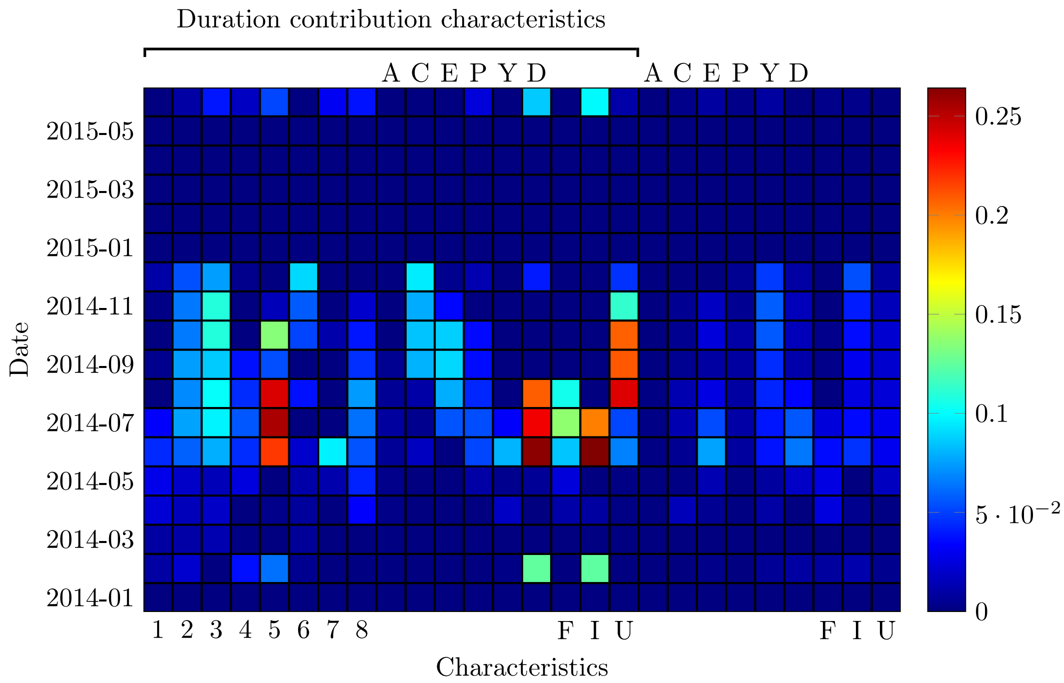
*Panel B: Overall results*

| Portfolio | $R^a$ | $\sigma^R$ | Dur. | Yield | OAS | TE | ATO* |
|---|---|---|---|---|---|---|---|
| Benchmark | -0.00 | 3.40 | 5.44 | 2.41 | 119 | – | 55.04 |
| Sample | -0.00 | 3.48 | 5.57 | 2.52 | 122 | 41 | 34.70 |

* The measure ATO, the adjusted total turnover, is calculated by summing up the total absolute differences in bond weights per month before and after rebalancing. The figure is then annualised by multiplying by twelve and dividing by the total observed months.

2015 (except for January 2015, to be elaborated below). Hence, for the strictness of our problem we were able to demand monthly turnovers of less than 5%, while satisfying the position restrictions. The question is at this point if we were also able to match the characteristics adequately. First of all, the resulting returns of the selected sample and the benchmark are displayed in figure 10. From the figure it seems that the sampled portfolio moves in similar direction as the benchmark and seems to match it quite accurately. In more detail, the performance figures are displayed in Table 7. We can now see that, while seemingly the returns are matched nicely, the tracking error is 41 basis points. If lower tracking error is desired, one either needs to increase the investable amount (and corresponding bounds on $M$), or examine the possible improvements

**Figure 11:** A display of the deviations of the sample characteristics with respect to the benchmark per year. The colour bar on the right shows the percentage deviation that a colour reflects. Each column corresponds to one of the 26 characteristics that were aimed to match in the optimisation, as introduced in Section 6.I. The first eight columns correspond to the duration contributions per maturity bucket, then to the duration contribution per sector and currency, and finally the pure weight exposure per sector and currency. To distinguish the labels of different strata (or characteristic groups), they are positioned alternately below and above the graph. For sectors the F corresponds to financials, I to industrials, and U to utilities. For currencies, A stands for Australian Dollar, C for Canadian Dollar, E for euros, P for Great British Pounds, Y for Japanese Yen, D for US Dollar, and . To clarify, in July 2014 there are big deviations in the matching on duration contribution for maturity bucket 5 (maturities between 7 and 10 years), and for Japanese Yen and Australian Dollar. January till May 2015 all characteristics were matched.




of the full optimisation from Problem 6.2 (or one of the other suggested improvements in the last section).

We moreover see from Table 7 that the absolute difference in bond weights, $\Delta^t$, sum up to the maximally allowed monthly turnover of 5% in most months. Also, we see that in January 2015 a higher turnover was needed in order to attain the constraints (due to a highly changed set of bonds due to the new year's update of constituents). However, afterwards lower turnover is observed while characteristics are fully matched in these months. This indeed raises the idea that the initial bond selection, and the one after one year in our simulation, determine the flexibility in the matching procedure considerably. In fact, we find that the set of selected bonds in 2015 was closer to the benchmark characteristics each month then in 2014, prior to the transactions. It is interesting to examine whether this can be predicted, so as to prevent the high peaks in characteristic deviations. The largest mismatch was in July 2014, where the sum of relative deviations from the benchmark characteristics was 174%.

Figure 11 shows in more detail the relative deviation per benchmark characteristic over time. The vertical axis corresponds to the different regarded months, each column contains the deviations for the duration contribution per maturity, sector, and currency, and the weight exposure per sector and currency. The bar right of the graph shows the deviation percentages represented per colour. For example, the US Dollar and Utilities duration contributions were mismatched by about 12% in February 2014, and around 25% in July. Three things are observed mainly. First, the higher penalty on weight exposure deviations leads to a better matching as all squares are tinted blue. It seems that certain characteristics are hard to match throughout, Yen and Pound duration contribution, while others, financials for example, are matched consistently. This seems to be as there are less bonds within the first group, making it harder to attain those attributes.

Panel B of Table 7 shows the overall results of the simulation. The annualised return and the annual volatility of the returns are close to each other, as is expected since the returns themselves were close to each other. The panel also shows that OAS and yield figures are close to each other, even though we only matched on duration. It does not seem necessary to constrain the optimisation more than we did now, but rather find ways to select bonds more efficiently (such that the characteristics can be met monthly with less trades).

Ultimately, we conclude that the fundamental indexation returns are possible to mimic even by portfolios with much less bonds. The tracking error of 42 basis points is already an acceptable figure but possibly can be reduced further by making the optimisation problem more precise. With regards to the previous chapter, we have found that the RAFI methodology outperforms a debt-weighted index and have shown that one of the index properties seems to be met. The next section investigates whether improvements can be found to the RAFI procedure itself.

# 7. Towards enhancing the RAFI method

We have seen that the RAFI methodology as suggested by Shepherd (2014) yields promising results, as it outperforms the debt-weighted index from a risk-return point of view and satisfies at least one of the index properties. It seems acceptable to say that regardless the screening, rebalancing, and weight assigning, purely the alternative weight determination itself improves the index performance. The motivation for the weighing factors, however, is from an economic perspective. This section examines whether additional motivation is found upon a quantitative analysis. Also, we slightly address the other two index requirements and suggest an alternative fundamental weighing method.

Our analysis is based on the observation that an alternative index (a better market proxy) is an implicit claim of better fundamental value determination. Arnott et al. (2010) show that, if observed credit spreads are noisy estimates of the true underlying default risks, a debt-weighted index overweights bonds that are priced too expensive (relative to their "true" default risk). Following the reasoning of Hsu (2004), this would lead to a positive correlation between the pricing error and price, as a consequence of market sentiment, and moreover a positive correlation with total face value per company as a consequence of the bums problem. A debt-weighing index systematically overweights these bonds and suffers from a performance drag.

Ultimately, one would therefore want to regard a regression of fundamentals on the pricing error, to see which fundamentals correct best for the market mis-pricing. We are aware that this is an ambitious aim, but feel that it can provide additional insight in what drives the outperformance of fundamental weighing. Moreover, it might provide a first step towards more advanced models in future research.

As stated in Section 2.V the other risk factors besides credit risk are: interest rate risk, option risk, and currency risk. Our analysis will consist of two steps: (1) to model the pricing error taking into account the bond risk factors, and (2) subsequently examine which company fundamentals explain the pricing error best.

## 7.I Modelling the pricing error

There exists extensive literature on modelling the returns of bonds. Where the CAPM is widely viewed as the basis model for pricing equities, the model of Chen et al. (1986) provides the "consensus" risk factors for corporate bond markets. Among others, they suggest a default premium and a term structure premium:

1. $DEF$: the general compensation investors require for the fact that a company issuing bonds can default, and their promised payments will not be met. The factor is constructed as the additional average returns on 7-10 year US corporate bonds over 7-10 year US treasuries.

2. $TERM$: the compensation demanded on the additional risk borne when promised payments reach further to the future, *i.e.* when a bond's maturity increases. Over time multiple characterisations are introduced, but a generally accepted one is the average return on 7-10 year US treasuries minus the one-month T-bill.

Subsequently, Fama and French (1993) argue that any pricing model should do well in explaining returns throughout asset classes, so not just for stocks nor just bonds. They therefore suggest taking the well-known equity factors, market return $MKT$, size $SMB$, and value $HML$, in addition to the default and term structure premia.[16] For any pricing model, testing should be conducted on a single, well-diversified, portfolio in order to minimise the random movements of individual assets. By doing so, they nicely show that about 70% of the variation in bond returns is explained by the combination of equity and bond factors, and about 95% for stocks. Fama and French (1993) do not include the momentum factor $MOM$ in their research, while this factor is now commonly included in the Carhart (1997)-model. $MOM$ is the average return on the stocks with highest previous 12-month returns minus the average return on stocks with lowest previous 12-month returns.[17]

Over time, different enhancements to the model have been suggested in the literature, that can be distinguished into two groups. First, models have been introduced to better explain part of the total returns (not necessarily taking them in excess of the 1-month T-bill). For example, Nederhof (2012) creates a model to forecast European corporate credit default swap spreads that proxy for the credit risk returns in bond total returns. As the second group, factors are presented that are left unexplained by the Fama and French (1993)-model in order to add risk-proxying portfolios that are based on those factors. For example, Houweling and Van Zundert (2014) and Israel et al. (2015) suggest swapping the equity factors by their bond market counterparts[18] and find that they hold significant explanatory powers. In fact, they increase the R-squared to an impressive 95%.

While promising, the approaches of Houweling and Van Zundert (2014) and Israel et al. (2015) are criticised as they construct decile portfolios in their factor definitions. By doing so, they select mostly bonds that are difficult to trade (small issues or ones close to maturity) and hence their results are argued to be dragged largely by a liquidity premium (that is not included separately in the models). A qualitative description of liquidity is at what cost, with what market impact, and in what time frame a certain amount of securities can be traded. In general, investors require a compensation if a bond is more illiquid (it bares more risk as, in the case one wants to sell, this is harder and more costly to do). An extensive research on available proxies for liquidity in corporate bonds can be found, for instance, in Bakker (2015), but no widely accepted measure exists.

We therefore stick to the basis model of Fama and French (1993) in order to obtain a reasonable cross-section comparison of stocks.[19] The found residuals will arguably contain a part of unexplained premia by including too few factors, but at least 70% of the portfolio return variation is explained, according to the results of Fama and French (1993). Their results span 1963 till 1991, which has no overlap with our observations. We therefore start off by checking whether their pricing model performs similarly on the more recent data of portfolio returns. Table 8 shows our results. Note, that in the last regression we replaced the market factor $MKT$ by the orthogonalised one, $RMO$, as suggested by Fama and French (1993). This factor is found by regressing the market return on the stock and bond factors, and summing up the constant term

[16]The $MKT$ factor is the value-weighted return on all NYSE, AMEX, and NASDAQ stocks in excess of the US 3-month T-bill rate. For the $SMB$ and $HML$ factors, all the stocks are sorted by their book-to-market value (book equity divided by market equity) and the lowest 30% is considered value, the highest 30% growth, and neutral in between. The stocks are also sorted just on size, and the bottom half (measured from the median) is small, the upper half is big. $SMB$ is the the average return of small value, neutral, and growth stocks minus the average on big value, neutral, and growth stocks. $HML$ is the average return on small and big value stocks minus the average return on small and big growth stocks.

[17]See `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/Data_Library/det_mom_factor.html`.

[18]Meaning that bond portfolios are constructed instead of the equity portfolios $SMB$, $HML$, and $MOM$. Moreover, the underlying size, value, and momentum factors are defined using bond characteristics.

[19]All stock factors are downloaded from the Kenneth French website `http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`.

**Table 8:** A reporting of the results of the regressions of Fama and French (1993), over a different time horizon. Our data spans January 1997- July 2015, versus 1963 - 1991. The first column shows for each row which table of Fama and French (1993) is replicated. All time series regressions are of the following form: $R_t^e = \alpha + \beta F_t + \varepsilon_t$. For each model the G0BC-index returns in excess of the 1-month US T-bill rate are the dependent variable. The factors $F_t$ differ per row, in accordance to the table of Fama and French (1993), and coefficients are shown only for the included regressors. Their t-statistics are shown below between parentheses. The last column shows the $R^2$ adjusted for the degrees of freedom.

| Table | $\alpha$ | $MKT$ | $SMB$ | $HML$ | $TERM$ | $DEF$ | $R^2$ |
|---|---|---|---|---|---|---|---|
| 3 | -0.03<br>(-0.48) | | | | 0.77<br>(18.58) | 0.90<br>(18.49) | 68.72 |
| 4 | 0.17<br>(1.43) | 0.13<br>(5.08) | | | | | 10.45 |
| 5 | 0.22<br>(1.83) | | 0.02<br>(0.65) | 0.04<br>(1.06) | | | -0.35 |
| 6 | 0.15<br>(1.28) | 0.14<br>(5.52) | -0.01<br>(-0.41) | 0.07<br>(1.91) | | | 11.55 |
| 8* | -0.08<br>(-1.17) | 0.07<br>(3.74) | 0.01<br>(0.58) | 0.03<br>(1.39) | 0.77<br>(19.06) | 0.90<br>(18.72) | 70.46 |

* In this regression the market return factor $MKT$ is replaced by the factor $RMO$, similar to Fama and French (1993). This is the orthogonalised market factor, found by regressing the excess market returns $MKT$ on the equity and bond premia, and then summing up the constant term and the residuals.

and regression residuals. In this way, the common variation of the risk factors is included in the $RMO$ factor, after all, Fama and French (1993) describe the market factor as a "hodgepodge" of pricing factors. The advantage of taking $RMO$ is that the same variation is explained, but the coefficients describe the influence per factor more adequately.

The fact that table 8 reports similar betas, t-statistics, and R-squared figures adds confidence that the risk premia are consistent through time. Hence, the model seems reasonable to take as a starting point in modelling the bond returns individually. As only addition we will examine the equity factor momentum $MOM$ in our research, and download the data as well from the Kenneth French website. In order to enable cross-section comparison we follow the procedure of Gebhardt et al. (2005), who take rating, duration, and bond type as control variables. Moreover, we will control for different currencies as we include bonds from multiple countries in the sample. In a similar fashion, we will also require bonds to have at least three years left to maturity, pay a coupon, and be non-convertible and investment grade. The first condition can be argued to deal with differences in liquidity among bonds, as mentioned above. Finally, we exclude bonds for which we have less than 13 observations and end up with 27505 bonds out of the 39760 to be selected. Note, that the panel is still (highly) unbalanced as some bonds have just 13 observations and maximally 222. Ultimately, we want to fit the following relation:

$$\begin{aligned} r_{i,t}^e &= \lambda X_{i,t} + \gamma Z_{i,t} + \varepsilon_{i,t}, \qquad \forall i = 1, \ldots, I, \text{ and } \forall t = 1, \ldots, T, \text{ where} \\ X &= [1, RMO, SMB, HML, MOM, TERM, DEF], \text{ and} \\ Z &= [RAT, DUR, TYPE, CUR]. \end{aligned} \tag{2}$$

This is the panel regression form of the excess bond returns for $I$ bonds on $T$ times, indexed by $i$ and $t$. The number of regressors in $X$ is denoted by $N$. Possible approaches are to regard a fixed effects model or a random effects model depending on the restrictions one wants to impose on the error term $\varepsilon$. However, as argued by Gebhardt et al. (2005), more reliable standard errors in the regression procedure can be found with a Fama and MacBeth (1973) approach. Hence, we will estimate equation (2) by first obtaining the estimated betas (coefficients for the risk factors) per firm, *i.e.* a time series regression per company, and then a cross-sectional regression per month. This results in the following stepwise models:

$$\underset{(T\times 1)}{r_i^e} = \underset{(T\times N)\times(N\times 1)}{X_i \quad b_i} + \underset{(T\times 1)}{e_i}, \qquad \forall i = 1,\ldots,I, \text{and then:} \tag{3}$$

$$\underset{(I\times 1)}{r_t^e} = \underset{(I\times N)\times(N\times 1)}{b_t \quad \lambda_t} + \underset{(I\times 1)}{\varepsilon_t}, \qquad \forall t = 1,\ldots,T. \tag{4}$$

The models are estimated by OLS-regressions with assumptions that error terms have mean zero. We follow the approach of Gebhardt et al. (2005) to correct for auto-correlation and heteroskedasticity in the time series regression by using Hansen-Hodrick-Newey-West corrected standard errors with 30 lags. Finally, Fama and MacBeth (1973) state that the averages of cross-sectional regression estimates then give the desired coefficients of equation (2):

$$\widehat{\lambda} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\lambda}_t \text{ , and } \widehat{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T}\widehat{\varepsilon}_{i,t}. \tag{5}$$

The sampling errors are obtained by taking the standard deviations of the estimates in equation (5), where we use the Hansen-Hodrick-Newey-West weighing matrix as stated above. We do note, however, that the Fama and MacBeth (1973) properties of (5) require the data to be balanced, which is not the case as we stated earlier. This problem is acknowledged by Petersen (2009), who shows that the Fama and MacBeth (1973) approach weights observations proportional to the number of companies in that month. Because the number of companies grows over time, recent observations become under-weighted. To correct for this, he suggest taking weighted averages (again proportional to the number of companies per month) to cancel out this effect and comparing the estimates and standard errors. We show the results for both approaches in Table 9.

As the first observation, we see that the effect of weighing observations differently over time is small. The coefficients are of similar values, both in the absolute terms as well as relative to the other factor premia coefficients. This adds evidence to time independent risk factors. Strikingly, we see that the orthogonal market factor has no significance in explaining the bond returns, neither has the stock momentum factor. The value and size premia, and the bond premia are significant and explain 43.63 percent of the variation after correcting for the rating, currency, type, and duration of the bonds. This is a significant decrease with respect to the 70.46 percent we found for a portfolio of bonds in Table 8. We distinguish the following reasons possible for this as: (1) either the randomness on individual bond level causes return movement that are not captured by the systematic pricing factors, or (2) we did not control for enough bond characteristics taking away explanatory power of the factors, or (3) we did but the idiosyncratic bond risk is getting priced by investors and we left out important bond specific pricing factors, or (4) a combination of the previous points. In any case, the residual captures the pricing error but includes different magnitudes of noise. Regardless, factors with explanatory power on the residual yield possible improvements on prices for the purpose of constructing an index.

Cochrane (2009) stresses the importance of checking whether the fitted pricing errors (regression residuals) are jointly zero, in order to see whether the Fama and MacBeth (1973) approach made sense. To do so, he suggests a sampling test on the pricing errors requiring estimation

**Table 9:** Here, the results of the Fama-Macbeth approach on the panel regression of equation (2) are reported, meaning that the regression is split up into 3 and 4. For each regressor the coefficient is shown along with its t-statistic. The standard errors are corrected for auto-correlation and heteroskedasticity, by using a Hansen-Hodrick-Newey-West weighing matrix based on 30 lags. We show the results for the regular Fama-Macbeth approach that overweighs observations at times of relatively few observations. In the next two columns, results are shown for the case where averages are weighted by their relative number of observations to correct for this. The last column shows the $R^2$ adjusted for the degrees of freedom.

| | | Fama-Macbeth | | Weighted Fama-Macbeth | | |
|---|---|---|---|---|---|---|
| Regressor | | Coefficient | t-statistic | Coefficient | t-statistic | $R^2$ |
| $X$ | constant | 0.25 | 4.62 | 0.21 | 3.72 | 43.63 |
| | $RMO$ | 0.03 | 0.63 | 0.02 | 0.35 | |
| | $SMB$ | -0.16 | -4.03 | -0.11 | -3.06 | |
| | $HML$ | 0.12 | 4.08 | 0.13 | 4.90 | |
| | $MOM$ | -0.03 | -0.34 | -0.08 | -0.85 | |
| | $TERM$ | 0.04 | 2.93 | 0.04 | 1.63 | |
| | $DEF$ | 0.03 | 2.67 | 0.02 | 3.47 | |
| $Z$ | $RAT$ | 0.02 | 336.04 | 0.02 | 343.62 | |
| | $CUR$ | -0.01 | -39.29 | -0.01 | -26.35 | |
| | $TYPE$ | 0.01 | 40.56 | 0.00 | 26.44 | |
| | $DUR$ | 0.03 | 246.63 | 0.03 | 267.06 | |

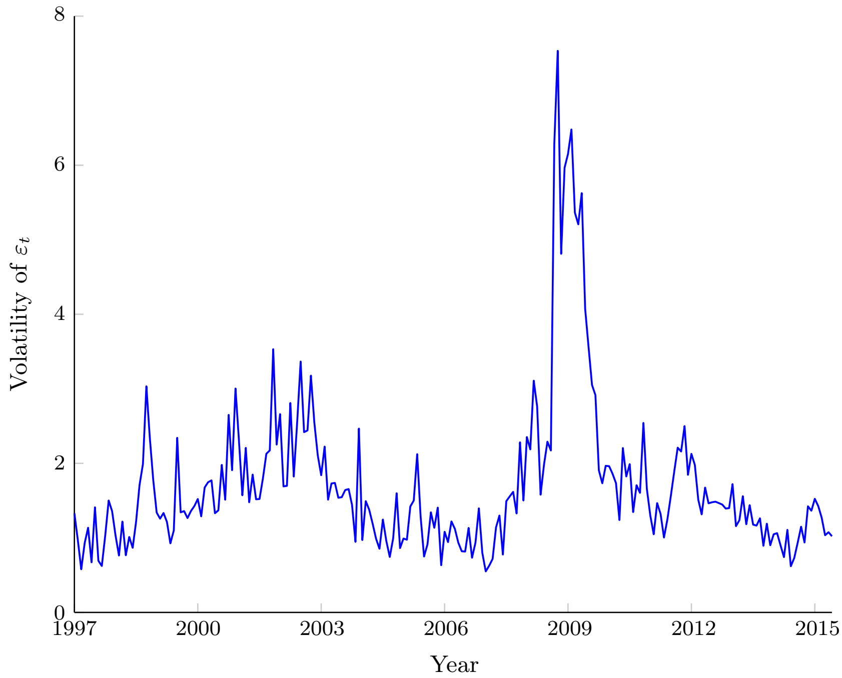
of a large covariance matrix and its inverse, making it unfeasible for our sample. We therefore follow the same procedure but take the identity matrix as covariance matrix, meaning that all correlations are assumed to be zero. By doing so, we find that the vector of pricing residuals cannot be rejected to be different from zero at all as we find 3,132 as test-statistic and 27,119 as critical value at 5% level. A better test should take into account the correlations, but with our simplifying assumption the Fama and MacBeth (1973) approach seems to be appropriate. As a final remark, we note that our model of the pricing error implicitly assumes an even split between over-pricing and under-pricing. Possibly, improvements can be found by allowing the average pricing error to fluctuate as well.

Summarising, the model explains about 44% of the variation in bond excess returns, while for a diversified portfolio of bonds 70% was explained. Still, the found residuals contain information on market sentiment, mis-pricings, that a fundamental index opts to correct for. Next section will examine the pricing residuals in order to see which company fundamentals yield the highest explanatory power.

## 7.II Explaining the pricing error

In the previous section we derived regression residuals of a bond pricing model that we regard as proxy for the pricing errors. It remains disputable whether they proxy for market sentiment well, or that too many relevant factors are left out in our pricing model. We will assume this is not the case, but leave it open to future research to examine more precise models. To back up our claim, we note that any model will face inexplainable noise due to the pure randomness in market sentiment. In this line of thought, we can only correct for structural mis-pricing that exists in the market. This leads to the idea of checking in our data whether we find the correlations that are suggested in the literature. As stated above, Hsu (2004) has shown why a positive relationship should be found between the price and pricing error as soon as mis-pricing

**Figure 12:** This graph shows the volatility of the residual $\varepsilon_t$ from the cross-section regression of equation (4), for each month in our sample of January 1997 to July 2015.




exist. Moreover, the bums-problem suggest a positive correlation between issue size and error. We find that both correlations are positive, though really small. The latter might add evidence that the structural part in mis-pricing is only small, or that our residual still contains important pricing factor variations (in which case our assumption is not adequate and the search for better models is interesting).

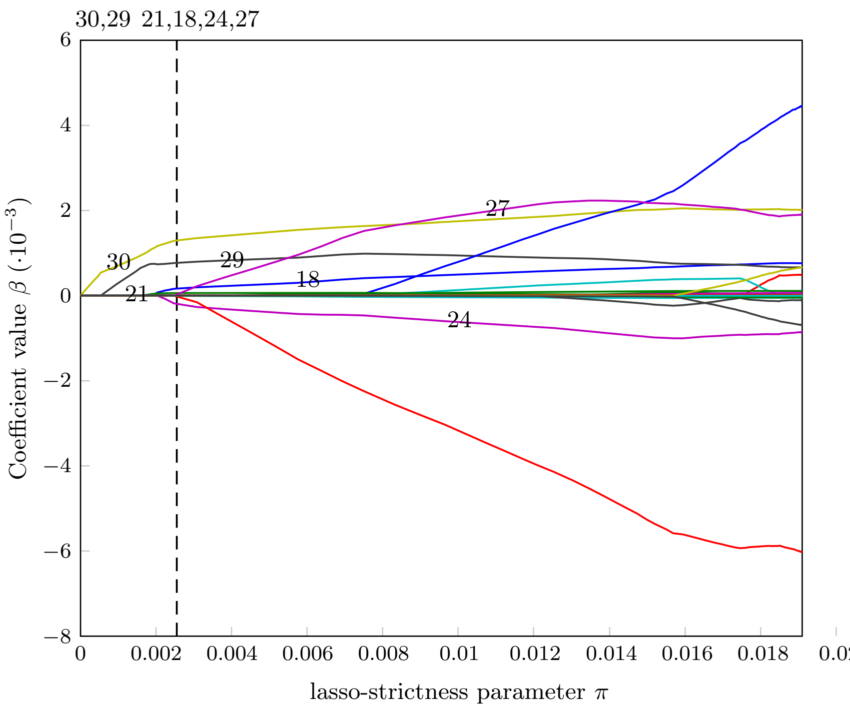
Now, the next step in our analysis is to find which fundamentals best explain the pricing error. Appendix E contains a description of the 40 different fundamentals that we include, along with our motivation. We will use these factors in the following regression and denote them by $F$:

$$\varepsilon_{i,t} = \beta F_{i,t} + \kappa C_{i,t} + \eta_{i,t}, \qquad \forall i = 1, \ldots, I, \ \text{ and } \ \forall t = 1, \ldots, T. \tag{6}$$

Here, the $\varepsilon$ comes from the regression 2 in the previous section. The variable $C$ contains the control variables `year` and `company` to concentrate on the fixed effects of the panel, and to deal with omitted variable bias. We do the regression twice, one time with the most recent fundamental data in $F$, and one with 5-year averages. Both give similar results in signs and significance for all of the regressors, rest of this section shows the results of the most recent fundamental data.

For the regression in equation (6) we suspect the fundamentals in $F$ to be subject to multi-collinearity, as we, for instance, use multiple measures of the company earnings (EBIT, EBITA, EBITDA, and net earnings). We still find that the matrix is invertible and we obtain coefficients for our regressors; the results are in Table 14 in Appendix E. However, the coefficients are highly biased but we will take care of this in the second step. For now, we see that the factors jointly explain 2% of the variance of the pricing error. A simple calculation indicates that in this way

**Figure 13:** Here, the coefficients resulting from the lasso-regression as described in this section are shown. The lasso-technique restricts the sum of absolute coefficients to be smaller than a parameter $\pi$. The graphs show per value of $\pi$ the coefficients that fit the data best. For the first six regressors that become significant the row numbers of the fundamentals in Table 14 in Appendix E are displayed. The dashed line shows for which value of $\pi$ the first six regressors obtained coefficients different than zero.




we have explained about 46% of the total excess return variation, which arguably leaves room for improvement. Still, we are only interested in correcting the valuation of the credit spread, so for this goal the results seem reasonable and we continue our investigation.

We find that 13 out of the 40 fundamentals are significant and that we only explain about 2 percent of the residual, see Appendix E. The small R-squared, does not have to be problematic as the pricing error possibly consists of much noise. However, we have two issues with the regression conducted on all fundamentals: first, we find that the number of significant regressors is too high, and second, we have that the standard errors are biased because of collinearity. Tibshirani (1996) provides a solution to both these issues with a shrinkage and selection technique called "lasso".[20] The results of the lasso-technique are displayed in figure 13. For each regressor the graphs indicate on the vertical axis the coefficient values, per level of strictness that is displayed on the horizontal axis. Hence, the lasso-technique becomes more lenient towards the right for higher values of $\pi$ (the notation of Tibshirani (1996) for the strictness parameter, but we will denote it with $\pi$). So, we see that gradually more regressors become significant, the precise coefficients per level of $\pi$ are in Table 14 in Appendix F. We see that the coefficients indeed go towards the coefficients of the unconstrained regression in (6) as $\pi$ increases. We discussed that those coefficient are presumably not reliable, so we want to regard a criteria to select an appropriate subset of regressors. To do so, we try two methods: (1) by regarding the marginal value of including one more variable in

[20]The term lasso stands for 'least absolute shrinkage and selection operator', introduced by Tibshirani (1996). In short, the lasso-technique constraints the sum of absolute regression coefficients to be smaller than some value $t$, while still minimising the sum of squared residuals over the coefficients. Resulting are a set of coefficients set to zero, and the remaining ones such that this subset explains the observed data best. This is a quadratic optimisation problem with linear constraints, for which efficient algorithms exist.

the order resulting from the lasso-technique, and (2) by performing F-test subsequently to see whether inclusion statistically adds value.

The marginal value of taking on an additional regressor towards the explained variation, as measured by the $R^2$, unfortunately does not give us a clear cut-off point. But, if we regard the outcomes of an F-test per step, inclusion of the seventh variable does not statistically add more explanation of the pricing error. The resulting factors are:

1. GROSS_MARGIN,
2. RETURN_ON_CAP,
3. NET_DEBT,
4. BS_LT_BORROW,
5. CF_REIMB_LT_BORROW,
6. RETURN_COM_EQY,

and their coefficients can be used to correct bond prices in building an index. This is what the next section seeks to confirm.

However, so far we only addressed robustness of the results by taking either the most recent figures for all fundamental values or the (up to) 5-year averages and found that the outcomes were similar. Ultimately, one would desire that the six variables above remain the best explaining set of fundamentals when the regarded period, sectors, currencies, or universes are changed. Possibly, the biggest challenge for the previous outcomes is to persist in different sub-periods, since markets have shown radical changes in behaviour over time. We therefore conduct the lasso-approach on the residuals of (6) for three different periods, and update the $\beta$ to be computed over the sub-periods only. We distinguish the pre-crisis, crisis, and post-crisis periods (1997–2008, 2008–2010, 2010–2015) and the results are in table 10.

*Note:* We do not update the residuals themselves, but just take the earlier found ones that are within the regarded sub-periods, company $\beta$'s are updated. The reason for this is that we argue that pricing factors should be constant over time and no substantial differences are to be expected by recalculating them per period. For the bond exposures towards the risk factors we think they might vary considerably over time and therefore do update those. For future research this is something to be checked.

Most remarkable is the observation that the last 5 years (post-crisis) give precisely the same six fundamentals in our selection procedure. In the other two periods only three of the six variables persist, CF_REIMB_LT_BORROW and GROSS_MARGIN in both periods, NET_DEBT pre-crisis, and BS_LT_BORROW post-crisis.[21] Furthermore, only bottom-half scores (rankings below 20) are found in the crisis period and for just two variables.

The fact that the exact same variables are found in the last period raises ambiguous doubt. At the one hand there was more data available over time, possibly making the results more accurate for the latter observations. In this case the data on the six variables prior to the last period was possible noisy, and therefore not yet adequately describing the errors. On the other hand, results might just be dragged mostly by the latter observations simply because there were more bonds in this period, that were moreover matched with fundamental data more often. Then, the

[21] We did not regard the different cut-off points, but merely took the number six as selection criterium. It might remain interesting to see how different criteria affect the number of variables to be selected.

**Table 10:** In panel A, the first column shows the six fundamentals that were found to best explain the pricing error over the time period (as described in this section), out the group of 40 fundamentals. The other columns shows how they are ranked if we regard the pre-crisis, crisis, and post-crisis periods (1997–2008, 2008–2010, 2010–2015). So, the lasso-technique is conducted over the residual of (6) for the months in these periods. Panel B shows the new 6 variables that hold the most explanatory power, the last row shows the R-squared adjusted for the degrees of freedom.

*Panel A: Rankings for the 6 best factors over the whole universe*

| Fundamental | Pre-crisis | Crisis | Post-crisis | Full period |
|---|---|---|---|---|
| GROSS_MARGIN | 2 | 5 | 1 | 1 |
| RETURN_ON_CAP | 9 | 28 | 2 | 2 |
| NET_DEBT | 4 | 12 | 3 | 3 |
| BS_LT_BORROW | 8 | 4 | 4 | 4 |
| CF_REIMB_LT_BORROW | 6 | 2 | 5 | 5 |
| RETURN_COM_EQY | 14 | 36 | 6 | 6 |

*Panel B: The best 6 factors over the different sub-periods*

| Rank | Pre-crisis | Crisis | Post-crisis |
|---|---|---|---|
| 1 | CF_INCR_LT_BORROW | PROF_MARGIN | GROSS_MARGIN |
| 2 | GROSS_MARGIN | CF_REIMB_LT_BORROW | RETURN_ON_CAP |
| 3 | BOOK_VAL_PER_SH | BS_CASH_NEAR_CASH_ITEM | NET_DEBT |
| 4 | NET_DEBT | BS_LT_BORROW | BS_LT_BORROW |
| 5 | SUSTAIN_GROWTH_RT | GROSS_MARGIN | CF_REIMB_LT_BORROW |
| 6 | CF_REIMB_LT_BORROW | SALES_REV_TURN | RETURN_COM_EQY |
| $R^2$ | 1.90% | 4.04% | 2.01% |

same results are produced regardless of being true. The latter reasoning obviously would greatly endanger our results. However, the fact that most factors remain highly ranked throughout, and the fact that half of them is selected in all sub-periods makes us lean towards the former argumentation.

We also see that the amount of the variation in the pricing residual that is explained (jointly) by all factors is of similar magnitude outside the crisis, but twice as high during the crisis period. As we saw earlier in figure 12, the volatility of the residual was also higher in this period. A possible explanation for this is that in this period the magnitude of the mis-pricing (*e.g.* due to panic among investors) was higher. This would increase the structural part of the pricing residual that can be explained by the company fundamentals, therefore leading to a R-squared that is twice as high. This add confidence that the pricing errors are adequately captured by our model.

So combined, we believe that residuals are reasonable proxies of the market mis-pricing, and that they are increasingly accurate over time due to more data being available. We therefore conclude that our analysis has led to a selection of fundamentals that might yield improvements with respect to the RAFI-factors (long term assets and cashflow). Hence, we continue our investigation in the next section to see how our results can be used for the construction of new fundamental approaches. Importantly, our results should be considered with care because of our specific assumptions. In stead of conducting more robustness checks at this point, we will do so for the resulting performance. Empirical evidence might add to the idea that directly correcting pricing errors is promising, this is what the next section examines.

### 7.III Our "fundamental indexation" approach

Before we build a portfolio based on the factors of the previous section, which we will dub "lasso-factors", we provide a reasoning why we *presume* that the approach of RAFI does not attain all of the index properties in Definition 2.1. While we have shown that the RAFI portfolio seems possible to be replicated, the information that is incorporated seems too few to enable equilibrium holding of the portfolio by all investors. An equal-weighted portfolio falls short of being an equilibrium portfolio because, if all investors hold it, all bonds (or other constituents) are priced equally despite different characteristics or risk attributes. Similarly, the RAFI portfolio would only address relative cashflow and long term assets as equilibrium determining factors. Arguably, many more factors should be taken into account. We will therefore regard two alternatives, one portfolio is the counterpart of the RAFI portfolio that takes constituent weights to be the average over the relative values of the lasso-factors (instead of the average over cashflow and long term assets).

Only few changes are needed to implement the RAFI method over the lasso-factors. We see that only net-debt has a negative coefficient meaning that lower values should get higher weights. We therefore subtract the largest net-debt from the net-debt per company, in this way all values are negative and the largest negative value gets the highest weight in the normalisation process. A company weight is obtained by taking the average of the relative lasso-factors. All other remains the same, hence the company weights are distributed over the bonds by relative face value, and the screening is done over the same ratios as before.

Second, we will therefore regard an alternative that uses price and the lasso-factors. As reasoning we argue that, even though price is subject to pricing errors, it (obviously) still reflects many relevant pricing factors.[22] As stated before, the fundamental indexes are designed to correct for the market mis-pricing, and we therefore will regard a method that tries to address this directly. The approach is as follows: (1) at the beginning of each month we use the previous month data to obtain a pricing residual using equation (2), (2) subsequently we regress the residuals on the six fundamentals of the previous section (that survived the lasso-approach) to obtain their coefficients using equation (6), and finally (3) correct the observed prices with the coefficients and fundamental data per bond. If no data exists we take the observed price and if data is found on some of the six fundamentals we simply correct for those ones only.

Hence, in line with the definitions of Section 4 our methodology is precisely the same as the debt weighted index in the fundamental universe, except that our prices are adjusted:

**Definition 7.1.** In line with Definition 4.1 our index takes $\Omega = \{p, n, a, F^L, \hat{\beta}^L\}$ and assigns target weights to each of the constituents, $i \in \mathcal{U}_t$, in the following way:

$$\varphi_i^M(f^M(t,\Omega_t)) = \begin{cases} \dfrac{f_i^M}{\sum_{j=1}^N f_j^M} & , \forall i \in \mathcal{U}_{\underline{t}} \\ 0 & , \text{otherwise} \end{cases} \quad , \text{ where:}$$

$$\underline{t} = \max\left(\tau \in \mathcal{F} \;:\; \tau \leq t\right),$$

$$f_i^M(t,\Omega_t) = p_{\underline{t},i} \cdot n_{\underline{t},i} + a_{\underline{t},i} + F^L_{\underline{t},i} \cdot \hat{\beta}^L_{\underline{t}}.$$

Hence, our index takes price $p$, the number of issues $n$, accrued interest $a$, and the six selected fundamentals of Section 7.II as factors in $\Omega$. The $p$, $n$, and $a$ are $N \times 1$ vectors with on the $i$-th

[22]We stated before that a methodology using price and fundamentals is not a fundamental index (since price is also used).

index the data of bond $i$ as ordered in $\mathcal{U}_t$. For the factors in $F^L$ $(N \times 6)$ and $\hat{\beta}^L$ $(6 \times 1)$ the upper script $L$ is used to clarify the dependence on the lasso-technique of Section 7.II, for the coefficients values are found upon a previous 12-month Fama-French regression. The $\underline{t}$ is defined as the first moment prior to $t$ that was a rebalancing moment (and hence included in $\mathcal{F}$). A value is assigned to each bond by the function $f$, note that the resulting values are indeed non-negative. Finally, the index function $\varphi^D$ normalises them assuring that all weights are between 0 and 1, and sum up to one, so that it is indeed in $\mathbb{W}$. △

An advantage of our approach is that all attributes of a debt-weighted index are *more likely* kept intact, by addressing the pricing error directly. We do not deem this as active investment as it satisfies the requirements of Definition 2.1, and we are not "picking" certain bonds. Moreover, we do not construct decile portfolios, nor use investment "styles", nor do we conduct a screening. Instead, all bonds receive a positive weight just like in the debt weighted index (we restricted our weights to be minimally zero but this bound was never attained). As a side note, we find that it is more natural for our method to be conducted over the whole universe by just taking the observed price if no data is available for a correction. Still, we regard the fundamental-universe of Definition 2.3 at this point to enable better comparison with RAFI. In this setting we find that the maximum price corrections are around 40% and only occur in the crisis period of 2008-2009. By far the most corrections are below 1%, the 95% percentile is 0.97%, the 99% percentile is 3.07%.

While these numbers seem to add confidence, a possible "look-ahead-bias" could appear as we found the six fundamentals by a regression analysis over the whole universe. We do not deem this problematic as this regression was never used in order to predict residuals, but merely to find fundamentals with the most explanatory power. Table 11 shows our results and figure 14 shows the development over time for the different indexes.

We see that the improvements of both approaches are impressive, even though we could explain only a small part of the pricing error. The rest of this section will elaborate on the performance of our method (that corrects prices), rather than the RAFI-method modified with our lasso-factors. We do this because we deem it more likely that our method attains all index properties, as argued above.

We note, however, that the results raise suspicion whether the performance is due to certain exposure shifts as we analysed in Section 5.II, or if the used fundamental factors are subject to a look-ahead bias after all (even though we initially argued this should not matter much). First, to address the possible look-ahead bias, we conduct our method once more, but with the cashflow and long term asset factors of Shepherd (2014) as described in Section 2. As those factors, arguably, provide a more clear economic interpretation, the bias is surpassed and we add the results to Table 11 and call it "Our method with RAFI factors". The results remain similar and add confidence that fundamentals have explanatory power on the pricing error, such that they can be used in correcting prices of a large group. With this I mean that the model was never constructed to try to predict individual bond mis-pricing, but only to explain the errors on aggregate level. I therefore do not see it as active strategy, but absence of a consensus definition makes this arguable.

Second, we check for the concentration and exposure shifts. Table 12 shows the exposures towards the bond risk premia in similar fashion as we did in Section 5. We find that our method reduces the exposure towards the default premium even further. In line of the argumentation of Arnott et al. (2005) the interpretation of this is ambiguous, as the alpha could be a result of

**Figure 14:** The solid line displays how a initial money holding in the BofA Merrill Lynch Global Corporate Index has evolved in the period January 1997 - June 2015. The dashed line shows how the same initial holding would have evolved by only debt-weighing the bonds for which fundamental data was found. The dotted line shows the performance of the fundamental indexation approach of Shepherd (2014), our approach is shown by the dashed-and-dotted line.

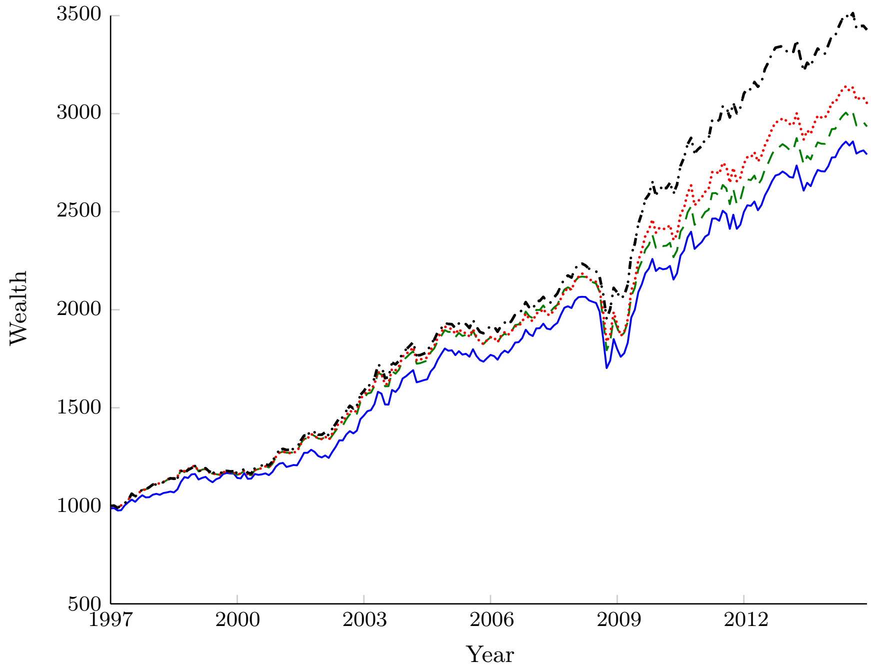



**Table 11:** Display of the performance of the debt-weighted index in the full universe (A), which is the G0BC-index. For the fundamental universe (F) the performance of the debt-weighted index, equal-weighted portfolio, fundamental method of RAFI, and our method are displayed. The fundamental index takes weights based on relative cashflow and long term assets, we also show the performance when our lasso-factors are taken as listed in Section 7.II. Conversely, our method corrects prices over the lasso-factors, but is also done for the RAFI factors. $R^a$ denotes average annualised return, $\sigma(R)$ its volatility, and SR the Sharpe ratio. Then yield, duration, option adjusted spread, and tracking error figures are displayed. All numbers are percentages, except for the Sharpe ratio that is a real number and OAS that is displayed in basis points. Tracking error is measured as the standard deviation of the difference in returns between the benchmark and stratified sample. TO is the turnover constructed by taking the lesser of monthly purchases and sells divided by the net asset value averaged over time and then annualised.

| Method | Univ. | $R^a$ | $\sigma(R)$ | SR† | yield | Dur. | OAS | TE | TO |
|---|---|---|---|---|---|---|---|---|---|
| G0BC | A | 5.87 | 6.34 | 0.57 | 4.59 | 5.38 | 134 | - | 27.54 |
| Debt-weighted | F | 6.16 | 6.18 | 0.64 | 4.93 | 5.50 | 139 | ref. | 20.79 |
| Equal-weighted | F | 6.51 | 5.87 | 0.72 | 4.98 | 5.54 | 144 | 1.31 | 36.36 |
| RAFI method... | F | 6.42 | 6.04 | 0.69 | 4.86 | 5.34 | 144 | 0.78 | 32.51 |
| ...on our factors | F | 6.86 | 5.28 | 0.86** | 5.21 | 6.05 | 150 | 2.26 | 28.89 |
| Our method... | F | 7.08 | 5.54 | 0.86*** | 4.92 | 5.62 | 139 | 1.72 | 24.61 |
| ...on RAFI factors | F | 6.92 | 5.59 | 0.83** | 4.92 | 5.62 | 139 | 1.52 | 24.65 |

† Statistical significance of Sharpe ratio difference with the G0BC index are displayed as *,**,*** for 90%, 95%, and 99% respectively. Results from a Jobson and Korkie-test.

**Table 12:** A display of the results of three linear regression models: $R_i^e = \alpha + \beta_1\text{TERM} + \beta_2\text{DEF} + \varepsilon$. TERM denotes a factor of average returns on 7-10 year duration Treasuries in excess of the one-month US T-bill. DEF is the default premium factor constructed as the returns on 7-10 year US corporate bonds minus 7-10 year US treasuries. The regression is done on monthly basis, $R_i^e$ are the regressed excess returns; per row is displayed which returns are taken.

| Dependent variable | Intercept | t-statistic | TERM | DEF | $R^2$ |
|---|---|---|---|---|---|
| Debt-weighted | 0.028% | 0.48 | 0.80*** | 0.92*** | 0.78 |
| RAFI | 0.055% | 0.88 | 0.78*** | 0.86*** | 0.74 |
| Our method | 0.113%*** | 1.94 | 0.76*** | 0.78*** | 0.79 |

Statistical significance levels are displayed as *,**,*** for 90%, 95%, and 99% respectively.

the regarded factors or the factors are driven by the found “market alpha”. We do note, however, that the intercept on our portfolio is the only one statistically different from zero. Figure 16 in Appendix G finally shows the sector allocations over time and a precise summation of the largest overweights and underweights are in Appendix H. We see that the sector allocations do not show any radical changes. Moreover, we see that the largest shift in weights on bond level are large issues that receive less weight in our portfolio. This possibly identifies the bums problem.

All in all, no obvious objections can be made and at the same time performance is improved. Moreover, the arguments of Shepherd (2014) that defend the RAFI index performance are met in all cases (and even improved as well). However, the found performance seems to make it a pyrrhic victory, as the returns appear to be too high to adequately describe the development of the underlying market, which is the global investment grade corporate debt market. So, eventually, our results seem to stress the need for more formal test on portfolios in order to say when the properties of Definition 2.1 are satisfied, *and* when one index is better than another. The absence of consensus on this makes it impossible to draw clear conclusions at this point.

## 8. Conclusions and recommendations

This thesis examined the fundamental indexation approach of Shepherd (2014), by adjusting the definitions of the weighing factors slightly, and by choosing an alternative universe including private companies with fundamental data. The five main contributions of this thesis are:

- improvements in the data matching approach (Section 3),
- formalisation of the indexation methods (Section 4),
- confirmation of the outperformance of the RAFI indexation method (Section 5),
- providing evidence that the RAFI index is replicable (Section 6),
- and that direct modelling of the pricing error may enhance the RAFI method (Section 7).

### Summary of our findings

We find that the debt-weighted index in the full bond universe, and the debt-weighted index over the bonds with fundamental data, are both outperformed by the fundamental indexation approach of Shepherd (2014). The outperformance cannot be solely addressed to one of the changes in the index construction, although performance diminishes when the rebalancing frequency is decreased from annually to monthly (which corresponds to the frequency of the regarded debt-weighted indexes). Still, the combination of changes yields results that are much in line with Shepherd (2014) adding evidence to its robustness. Generally, the outperformance is larger in the broader universes indicating that the index adds value particularly in allocating debt more efficiently than the debt-weighted indexes. This adds value to the claim that markets are not efficient, and as a result, that the debt-weighted indexes fall short of adequately proxying for their underlying markets because of this mis-pricing.

Subsequently, we check whether the fundamental approach also satisfies the other index requirements as described by Siegel (2003). Summarising, an index should be able to be held in equilibrium, should be replicable, and provide a useful benchmark for investors. We first regard whether the fundamental approach can be replicated. We find that investable portfolios can be constructed by a mixed integer linear programming approach that lead to acceptable replication on the basis of 18 months data, January 2014 – July 2015. The returns are matched with a tracking error of 41 basis points if we are allowed to invest with an amount of 200 million, and select between 150 and 300 bonds. This result is found by matching the portfolio characteristics with the ones of the fundamental index (per month), while simultaneously constraining and minimising transaction costs. This procedure is known as stratified sampling and possible improvements can be found by taking different strata. Moreover, the optimisation is time-intensive, and we suggest the search for approximation methods. Finally, improvements are possible by taking more adequate mathematical representations in the objective function of the tradeoffs between transaction costs and deviations in portfolio characteristics.

Then, we address the critique that the Shepherd (2014)-approach (and the fundamental indexation approach for stocks) do not fulfil the equilibrium property of an index. We agree that company's cashflows and long term assets incorporate for too few information to likely be able to do so. We therefore suggest a hybrid methodology that uses price and fundamentals, such that the fundamentals are only used to correct for the pricing error. We model the pricing error using a Fama and MacBeth (1973) approach in combination with the Fama and French (1993) model for corporate bond pricing. In the final step we use a lasso-method (least absolute shrinkage and selection operator) to reduce the number of selected fundamentals as well as taking care of the collinearity bias of our initial regression on 40 different fundamentals. We explain only a small part of the pricing residual, and argue that this is because of incompleteness of our model on the one hand, and because of the largest part of pricing error being purely random. Possibly, improvement can be found by conducting non-linear models, or by adding relevant factors.

Still, we decide to see whether the currently found fundamentals can be used for correcting prices. We do so by, at each month, using the previous 12 month data in order to find coefficients for the six factors. While prices are only adjusted marginally in this way, the resulting portfolio seems much more stable as less volatility is observed. Moreover, we do not find any suspicious shifts in allocations or risk exposures and use the argumentation of Shepherd (2014) that the portfolio therefore is likely to satisfy the index properties. In fact, if one believes that market mis-pricing exists our methodology possibly provides interesting results as the portfolio (1) provides a better proxy of the underlying market than the debt-weighted index by (partly) bypassing market mis-pricing, and is likely to fulfil the index requirements as it (2) incorporates all information that is captured by market prices enabling it for equilibrium holding by many investors, and (3) the weights are stable through time making replication of the portfolio likely.

#### Answering the research questions and hypothesis

*The original hypothesis: My hypothesis is that the RAFI methodology adds value as it bypasses price in the weight determination and thereby is not affected by market mis-pricing. However, I am anxious that the RAFI method incorporates for too few information and therefore expect undesirable concentration shifts or radical changes in certain exposures. For this reason I do not expect it to be possible to replicate the RAFI portfolio returns, and think that improvements can be found.*

We can now conclude that the RAFI-methodology indeed seems to outperform (partly) on the basis of avoiding market mis-pricing. However, the fact that too few information is taken into account does not seem to imply that investors cannot replicate its returns, but it does allow for possible improvements. We believe that either the RAFI-method with our factors, or our regarded hybrid-methods yield additional value. But, also for our method some doubt persists as the returns seem too high to reflect the underlying market, and therefore falling short of one of our index requirements.

#### Recommendations

Based on our findings we recommend for practitioners to decide which index properties are desired and for what purpose the benchmark is used:

- If you aim to replicate an index that bypasses price, and do not care about equilibrium properties, the RAFI method provides good results.
- In case equilibrium properties are desired, and further improvements in risk and returns, then a fundamental method on our factors or our hybrid methods (as described in Section 7) are preferable.

Suggestions for future research to validate our assumptions have been made in Sections 6 and 7 and are not repeated here. By aggregating the findings of this thesis we further recommend the following:

- Researchers must carefully define the index properties in order to avoid the semantic discussions that have appeared in current articles. Definitions are required before introducing new methods.

- For the definition of Siegel (2003) we suggest two additional checks to see if an underlying market is reflected. First, the long-term index returns should be close to the long-term trend of the market price. Second, relative performance of active fund managers should result in a zero-sum game with respect to the index.
- Finally, our index methodologies need to be tested in different datasets (*e.g.* other asset classes, periods, or universes) in order to address the robustness of our results.

## References


Altman, E. I. and Saunders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking & Finance*, 21(11):1721–1742.

Arnott, D., Hsu, J., and Moore, P. (2005). Fundamental indexation. *Research Affiliates White Paper*.

Arnott, R. D., Hsu, J. C., Li, F., and Shepherd, S. D. (2010). Valuation-indifferent weighting for bonds. *Journal of Portfolio Management*, 36(3):117–130.

Arnott, R. D. and Markowitz, H. M. (2008). Fundamentally flawed indexing: Comments. *Financial Analysts Journal*, 64(2):12–14.

Asness, C. (2006). The value of fundamental investing. `http://www.institutionalinvestor.com/Article/1082031/The-Value-of-Fundamental-Indexing.html`.

Bakker, M. (2015). Liquidity in the european corporate bond market. *Unpublished*.

Basu, A. K. and Forbes, B. (2014). Does fundamental indexation lead to better risk-adjusted returns? new evidence from australian securities exchange. *Accounting & Finance*, 54(3):699–728.

Best, M. and Grauer, R. (1991). On the sensitivity of mean-variance-efficient portfolios to changes in asset means: some analytical and computational results. *Review of Financial Studies*, 4(2):315–342.

Black, F. and Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5):28–43.

Blitz, D. and Swinkels, L. (2008). Fundamental indexation: an active value strategy in disguise. *Journal of Asset Management*, 9(4):264–269.

Blitz, D., Van der Grient, B., and Van Vliet, P. (2010). Fundamental indexation: Rebalancing assumptions and performance. *Journal of Index Investing*, 1(2):82–88.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Chen, C., Chen, R., and Bassett, G. W. (2007). Fundamental indexation via smoothed cap weights. *Journal of Banking & Finance*, 31(11):3486–3502.

Chen, N.-F., Roll, R., and Ross, S. A. (1986). Economic forces and the stock market. *Journal of business*, pages 383–403.

Cochrane, J. H. (2009). *Asset Pricing:(Revised Edition)*. Princeton university press.

Da Silva, A., Lee, W., and Pornrojnangkool, B. (2009). The Black-Litterman model for active portfolio management. *The Journal of Portfolio Management*, 35(2):61–70.

De Jong, M. and Wu, H. (2014). Fundamental indexation for bond markets. *Amundi Working Paper*.

Dijkstra, T. (2013). On perold's fundamentally flawed indexing. `https://www.researchgate.net/publication/238240589`.

Estrada, J. (2006). Fundamental indexation and international diversification. *Available at SSRN 949162*.

Fama, E. F. (1965). The behavior of stock-market prices. *Journal of business*, pages 34–105.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of financial economics*, 49(3):283–306.

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

Fama, E. F. and MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy*, pages 607–636.

Fuller, R. J., Giovinazzo, R., and Tung, Y. (2014). The stable roe portfolio: An alternative equity index strategy based on common sense security analysis. *The Journal of Portfolio Management*, 40(5):135–145.

Gebhardt, W. R., Hvidkjaer, S., and Swaminathan, B. (2005). The cross-section of expected corporate bond returns: Betas or characteristics? *Journal of Financial Economics*, 75(1):85–114.

Houweling, P. and Van Zundert, J. (2014). Factor investing in the corporate bond market. *Social Science Research Network*.

Hsu, J. C. (2004). Cap-weighted portfolios are sub-optimal portfolios. *Social Science Research Network*.

Israel, R., Kang, J., and Richardson, S. (2015). Investing with style in corporate bonds. *Social Science Research Network*.

Jensen, M. C., Black, F., and Scholes, M. S. (1972). The capital asset pricing model: Some empirical tests.

Jobson, J. and Korkie, M. (1981). Performance hypothesis testing with the sharpe and treynor measures. *Journal of Finance*, pages 889–908.

Kritzman, M., Page, S., and Turkington, D. (2010). In defense of optimization: the fallacy of 1/n. *Financial Analysts Journal*, 66(2):31–39.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The review of economics and statistics*, pages 13–37.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *Journal of economic perspectives*, pages 59–82.

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

Markowitz, H. (1959). Portfolio selection: Efficient diversification of investments. *New York: John Wiley & Sons*.

Markowitz, H. M. (2005). Market efficiency: A theoretical distinction and so what? *Financial Analysts Journal*, 61(5):17–30.

Nederhof, P. (2012). Forecasting european corporate cds spreads. *Unpublished*.

Perold, A. (2007). Fundamentally flawed indexing. *Financial Analysts Journal*, 63(6):31–37.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of financial studies*, 22(1):435–480.

Roll, R. (1977). A critique of the asset pricing theory's tests part i: On past and potential testability of the theory. *Journal of financial economics*, 4(2):129–176.

Russo, A. (2013). Low risk equity investments. *Amundi Working Paper*.

Sharpe, W. (1965). Risk-aversion in the stock market: Some empirical evidence. *Journal of Finance*, 20(3):416–422.

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

Shepherd, S. (2014). Smart beta investing in corporate bonds: Conceptual and empirical grounds. *Research Affiliates White Paper*.

Shleifer, A. (2000). *Inefficient markets: An introduction to behavioral finance*. Oxford university press.

Shleifer, A. and Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52(1):35–55.

Siegel, J. (2006). The noisy market hypothesis. *Wall Street Journal*, 14:A14.

Siegel, L. B. (2003). Benchmarks and investment management.

Simon, H. A. (1965). *Administrative behavior*, volume 4. Cambridge Univ Press.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288.

Treynor, J. (1961). *Toward a theory of market value of risky assets*. Unpublished.

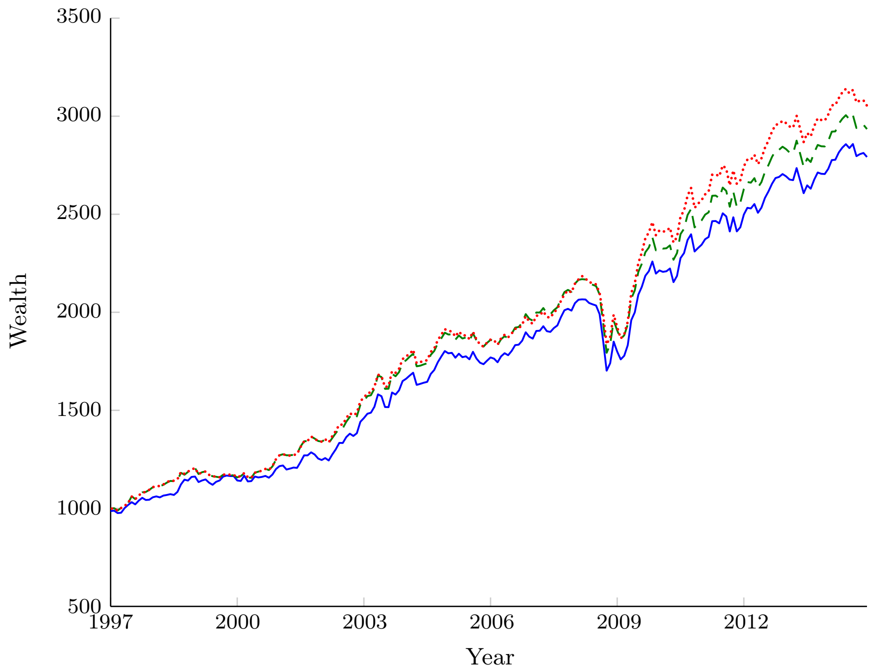
Treynor, J. (2005). Why market-valuation-indifferent indexing works. *Financial Analysts Journal*, 61(5):65–69.

Treynor, J. (2008). Fundamentally flawed indexing: Comments. *Financial Analysts Journal*, 64(2):14–14.

# Appendices

## A Index performance detail

**Figure 15:** The solid line displays how a initial money holding in the BofA Merrill Lynch Global Corporate Index has evolved in the period January 1997 - June 2015. The dashed line shows how the same initial holding would have evolved by only debt-weightng the bonds for which fundamental data was found. The dotted line shows the performance of the fundamental indexation approach of Shepherd (2014).




## B Example of variable replacement

*Proof.* We will show per case that the restrictions $iv$, $v$, $vi$, $vii$ from problem 6.2 imply the relation: $\Delta^t \equiv |w_i - \hat{w}_i|$.

<u>*Case a*</u>: $w_i - \hat{w}_i > 0$

Then $z'$ and $z''$ become:

$$\begin{cases} -z_i + w_i - \hat{w}_i \leq 0 \implies z_i = 1 \\ z_i - (w_i - \hat{w}_i) \leq 1 - w^L \implies z_i \in \{0, 1\} \end{cases}, \quad \implies z_i = 1.$$

Now the constraints become:

$$iv: \; -(w_i - \hat{w}_i) - \Delta_i^t \leq 0 \implies \Delta_i^t \in [0, 1]$$
$$v: \; w_i - \hat{w}_i + \Delta_i^t - 2 \leq 0 \implies \Delta_i^t \in [0, 1]$$
$$v: \; -(w_i - \hat{w}_i) + \Delta_i^t \leq 0 \implies \Delta_i^t \in [0, w_i - \hat{w}_i]$$

$$vi: \; -(w_i - \hat{w}_i) - \Delta_i^t \leq 0 \implies \Delta_i^t \in [w_i - \hat{w}_i, 1].$$

Hence:

$$\Delta_i^t \equiv w_i - \hat{w}_i.$$

<u>*Case b*:</u> $w_i - \hat{w}_i \leq 0$

Then $z'$ and $z''$ become:

$$\begin{cases} -z_i + w_i - \hat{w}_i \leq 0 \implies z_i \in \{0,1\} \\ z_i - (w_i - \hat{w}_i) \leq 1 - w^L \implies z_i = 0 \end{cases}, \quad \implies z_i = 0.$$

Now the constraints become:

$$\begin{aligned} iv: \; & -(w_i - \hat{w}_i) - \Delta_i^t \leq 0 \implies \Delta_i^t \in [-(w - \hat{w}_i), 1] \\ v: \; & w_i - \hat{w}_i + \Delta_i^t \leq 0 \implies \Delta_i^t \in [0, -(w - \hat{w}_i)] \\ vi: \; & -(w_i - \hat{w}_i) + \Delta_i^t - 2 \leq 0 \implies \Delta_i^t \in [0, 1] \\ vii: \; & -(w_i - \hat{w}_i) - \Delta_i^t \leq 0 \implies \Delta_i^t \in [0, 1]. \end{aligned}$$

Hence, now:

$$\Delta_i^t \equiv -(w_i - \hat{w}_i).$$

So indeed the non-redundant constraints adequately make sure the auxiliary variable equals the absolute difference in rebalanced and pre-rebalanced weights. □

## C Proof of proposition 6.1

*Proof.* It is sufficient to show per case that the most strict bound of $i$, $ii$, or $iii$ coincides with equation 1, as the negative counterpart make sure they are then attained:

<u>*Case a*:</u> $w_i = 0$ & $w_i \leq \overline{w}_i^{BM}$

Then c' and c" become:

$$\begin{cases} -c_i \leq 0 \\ c_i \leq 1 - w^L \end{cases}, \quad \implies c_i = 0, \text{ since } c_i \in \{0,1\}.$$

Also, $w_i - \overline{w}_i^{BM} \geq -1$ so the constraints become:

$$\begin{aligned} i: & \; \Delta_i^w \geq -(w_i - \overline{w}_i^{BM}) - 2 \geq -1 \\ ii: & \; \Delta_i^w \geq w_i - \overline{w}_i^{BM} \geq -1 \\ iii: & \; \Delta_i^w \geq 0 \equiv \Delta^w. \end{aligned}$$

<u>Case a'</u>: $w_i = 0$ & $w_i > \overline{w}_i^{BM}$

$$\implies \overline{w}_i^{BM} < w_i = 0, \text{ contradiction since } \overline{w}_i^{BM} \geq 0!$$

<u>Case b</u>: $w_i > 0$ & $w_i \leq \overline{w}_i^{BM}$

We now have that $w_i - \overline{w}_i^{BM} \leq 0$, and for the control variables that:

$$\begin{cases} w_i - c_i & \leq 0 \\ -w_i + c_i & \leq 1 - w^L \end{cases}, \quad \implies 1 = c_i \geq w_i > 0, \text{ since } c_i \in \{0, 1\}.$$

Now we can regard the constraints:

$$\begin{aligned} i &: \Delta_i^w \geq -(w_i - \overline{w}_i^{BM}) \equiv \Delta^w \geq 0 \\ ii &: \Delta_i^w \geq w_i - \overline{w}_i^{BM} \geq -1 \\ iii &: \Delta_i^w \geq -1. \end{aligned}$$

<u>Case c</u>: $w_i > 0$ & $w_i > \overline{w}_i^{BM}$

In this case we have $w_i - \overline{w}_i^{BM} > 0$ and:

$$\begin{cases} w_i - c_i & \leq 0 \\ -w_i + c_i & \leq 1 - w^L \end{cases}, \quad \implies 1 = c_i \geq w_i > 0, \text{ since } c_i \in \{0, 1\}.$$

Now the constraint give:

$$\begin{aligned} i &: \Delta_i^w \geq -(w_i - \overline{w}_i^{BM}) \geq -1 \\ ii &: \Delta_i^w \geq w_i - \overline{w}_i^{BM} \equiv \Delta^w \geq 0 \\ iii &: \Delta_i^w \geq -1. \end{aligned}$$

So indeed in all cases the highest lower bound (and in fact the only non-redundant one) is equal to equation 1 and the one that will be attained by imposing the opposite constraints $i^-$, $ii^-$, and $iii^-$. The auxiliary constraints and variables therefore admit replacing of the non-linear deviation constraint in problem 6.1. □

## D Matlab implementation of stratified sampling MILP

```
% Matlab implementation of: min_x p'*Dc + tau*Dt s.t.
%    A * x <= b
%    Aeq * x = beq
%    lb <= x <= ub
%    and x(int_idx) are integer
%    where x = {w,c,Dc,z,Dt}.
%
% Written by Ronald Smits, 13-12-2015.

rows = [kron(1:N,[1 1]),...          %c'
    kron(N+1:2*N,[1 1]),...          %c''
    (2*N+1)*ones(1,N),...            %max. selection
    (2*N+2)*ones(1,N),...            %min. selection
    kron(2*N+3:3*N+2,[1 1 1]),...    %iv
    kron(3*N+3:4*N+2,[1 1 1]),...    %iv-
    kron(4*N+3:5*N+2,[1 1]),...      %iv'
    kron(5*N+3:6*N+2,[1 1]),...      %iv-'
    kron(6*N+3:7*N+2,[1 1]),...      %s'
    kron(7*N+3:8*N+2,[1 1]),...      %s''
    (8*N+3)*ones(1,N),...            %max. turnover
    8*N+4:9*N+3,...                  %max. position
    kron(9*N+4:10*N+3,[1 1]),...     %min. position
    kron(10*N+4:11*N+3,[1 1]),...    %min. trade
    ];

cols = [reshape([1:N; N+1:2*N],1,[]),...
    reshape([1:N; N+1:2*N],1,[]),...
    N+1:2*N,...
    N+1:2*N,...
    reshape([1:N; 2*N+1:3*N; 3*N+1:4*N],1,[]),...
    reshape([1:N; 2*N+1:3*N; 3*N+1:4*N],1,[]),...
    reshape([1:N; 2*N+1:3*N],1,[]),...
    reshape([1:N; 2*N+1:3*N],1,[]),...
    reshape([1:N; 3*N+1:4*N],1,[]),...
    reshape([1:N; 3*N+1:4*N],1,[]),...
    2*N+1:3*N,...
    1:N,...
    reshape([1:N; N+1:2*N],1,[]),...,...
    reshape([2*N+1:3*N; 3*N+1:4*N],1,[]),...
    ];

vals = [kron(ones(1,N), [1 -1]),...
    kron(ones(1,N), [-1 1]),...
    ones(1,N),...
    -ones(1,N),...
    kron(ones(1,N), [1 1 -2]),...
    kron(ones(1,N), [-1 1 2]),...
    kron(ones(1,N), [-1 -1]),...
    kron(ones(1,N), [1 -1]),...
    kron(ones(1,N), [1 -1]),...
    kron(ones(1,N), [-1 1]),...
    ones(1,N),...
    I(t-S+1)./data(:,14)',...
    reshape([-I(t-S+1)*ones(1,N); alfa*data(:,14)'],1,[]),...
    reshape([-I(t-S+1)*ones(1,N); delta*ones(1,N)],1,[]),...
```

```
    ];

charConstr = [MaturityBuckets;...
    CurrencyDurationContribution;...
    SectorDurationContribution;...
    CurrencyExposure;...
    SectorExposure];

A = [sparse(rows,cols,vals,max(rows),max(cols)+n_chars);...
    charConstr; -charConstr];

b = [zeros(N,1);...                    %c'
    ones(N,1) - wL;...                 %c''
    M;...                              %max. selection
    -Mmin;...                          %min. selection
    wSS_BR;...                         %iv
    wSS_BR + 2*ones(N,1);...           %iv-
    -wSS_BR;...                        %iv'
    wSS_BR;...                         %iv-'
    wSS_BR;...                         %s'
    ones(N,1)- epsilon- wSS_BR;...     %s''
    epsilon;...                        %max. turnover
    beta*ones(N,1);...                 %max. position
    zeros(N,1);...                     %min. position
    zeros(N,1);...                     %min. trade
    BenchmarkTargets;...               %max. distance
    -BenchmarkTargets];                %min. distance

Aeq = [ones(1,N) zeros(1,3*N) zeros(1,n_chars)];

beq = 1;

int_idx = [N+1:2*N, 3*N+1:4*N];
lb = [zeros(3*N,1); zeros(N+n_chars,1)];
ub = [ones(4*N,1); Inf*ones(n_chars,1)];

f = [zeros(2*N,1); tau/N*ones(N,1); zeros(1*N,1); ones(n_chars-n_currs-n_secs,1);
    2*ones(n_currs+n_secs,1)];

[x,fval,exitflag,output] = intlinprog(f,int_idx,A,b,Aeq,beq,lb,ub,options);
```

## E Overview of fundamentals

**Table 13:** This table lists the fundamental variables used in the fixed effects panel regression of Section 7.II: $\varepsilon_{i,t} = \beta F_{i,t} + \eta_{i,t}$. Each row corresponds to one of the regressors in $F$. The first column displays the Bloomberg Field name, and a description. Then, for each variable the coefficient is displayed with Hansen-Hodrick-Newey-West corrected t-statistics between parentheses, the next column counts the significant results. Finally, the order of significance is displayed as result of the lasso-approach. The R-squared adjusted for the degrees of freedom is 2.04%.

| | Variable | Description | $\beta$ | Y/N | Order |
|---|---|---|---|---|---|
| 0 | `Constant` | The constant term in the regression | 39.43 (0.50) | | 21 |
| 1 | `NET_INCOME` | The profit after all expenses have been deducted, including all extraordinary gains and losses raising possibility that investors overreact | -7.8559 (-6.99) | 1 | 21 |
| 2 | `EARN_FOR_COMMON` | Net income minus total cash of preferred dividend and other adjustments. Possibly better proxy of 1 | 7.5974 (6.83) | 2 | 35 |
| 3 | `IS_EPS` | Earnings per share, possibly better proxy of 1 | 19.01 (0.66) | | 14 |
| 4 | `BEST_EBIT` | Analyst forecast of earnings possibly leading to overreaction of investors due overconfidence or herding | -0.02 (-0.65) | | 36 |
| 5 | `EBITA` | Possibly better proxy of 1. | 0.17 (1.82) | 3 | 8 |
| 6 | `EBITDA` | Often seen as proxy of generated cash flow and possibly overlooking underlying facets | -0.14 (-0.66) | | 28 |
| 7 | `SALES_REV_TURN` | Total of operating revenues | 0.01 (0.52) | | 32 |
| 8 | `GROSS_PROFIT` | Net sales minus cost of goods sold | 0.01 (0.22) | | 11 |
| 9 | `BS_CUR_ASSET_REPORT` | Total of reported assets incl. cash, short-term investments, inventories, and accrued income | 0.02 (1.75) | 4 | 25 |
| 10 | `NET_DEBT_TO_SHRHLDR_EQTY` | Proxy for leverage, net debt divided by total equity | 0.70 (1.89) | 5 | 10 |
| 11 | `BOOK_VAL_PER_SH` | Total common equity per share | 0.00 (0.24) | | 29 |
| 12 | `CASH_FLOW_TO_TOT_LIAB` | Cash from operations divided by total liabilities | -59.99 (-2.71) | 6 | 15 |
| 13 | `CASH_FLOW_TO_INT_EXPENSE` | Cash from operations divided by total incurred interest | 1.11 (1.30) | | 16 |
| 14 | `CF_CASH_FROM_OPER` | Net income plus depreciation and amortisation and non-cash adjustments | 0.13 (0.56) | | 33 |
| 15 | `EQY_SH_OUT` | Number of outstanding equity shares | 0.06 (2.74) | 7 | 7 |
| 16 | `NUM_OF_EMPLOYEES` | Number of total hired staff in FTE | -0.00 (-0.26) | | 17 |
| 17 | `CUR_MKT_CAP` | Current shares outstanding times the last market price of equity | 0.01 (0.64) | | 37 |
| 18 | `BS_LT_BORROW` | Sum of all interest-bearing obligations over one year | 0.01 (0.44) | | 4 |
| 19 | `BS_ST_BORROW` | Sum of all interest-bearing obligations within one year | 0.04 (1.92) | 8 | 26 |
| 20 | `TOT_DEBT_TO_TOT_EQY` | Total debt divided by common equity | -0.47 (-1.42) | | 31 |
| 21 | `NET_DEBT` | Liabilities and debts net from cash and cash equivalents and marketable securities | -0.04 (-2.73) | 9 | 3 |
| 22 | `DIVIDEND_YIELD` | Trailing 12-month dividend per share | -1.13 (-0.09) | | 30 |

Table continues on next page

| | Continuation of table on previous page | | | | |
|---|---|---|---|---|---|
| | Variable* | Description | $\beta$ | Y/N | Order |
| 23 | CF_DVD_PAID | Portion of paid dividends paid in cash to shareholders | -0.24 (-0.86) | | 39 |
| 24 | CF_REIMB_LT_BORROW | Reimbursement of long term borrowings | 0.06 (1.30) | | 5 |
| 25 | CF_INCR_ST_BORROW | Increase in short term borrowings as of last reporting date | -0.47 (-1.89) | 10 | 12 |
| 26 | CF_INCR_LT_BORROW | Increase in long term borrowings as of last reporting date | 0.07 (1.36) | | 34 |
| 27 | RETURN_COM_EQY | Return on common equity | 0.12 (0.07) | | 6 |
| 28 | RETURN_ON_ASSET | Return on common assets | 46.72 (1.93) | 11 | 23 |
| 29 | RETURN_ON_CAP | Return on capital | 6.69 (0.82) | | 2 |
| 30 | GROSS_MARGIN | Net sales minus cost of goods sold divided by net sales | 20.00 (5.13) | 12 | 1 |
| 31 | OPER_MARGIN | Operating income divided by total revenue | -8.14 (-1.17) | | 13 |
| 32 | PRETAX_MARGIN | Pretax income divided by revenue | -0.83 (-0.06) | | 24 |
| 33 | PROF_MARGIN | Net income divided by revenue | 6.11 (0.39) | | 40 |
| 34 | DVD_PAYOUT_RATIO | Dividends paid as portion of net income | 0.05 (0.49) | | 27 |
| 35 | EFF_TAX_RATE | Total tax paid divided by the pretax income | -0.04 (-0.12) | | 41 |
| 36 | SUSTAIN_GROWTH_RT | Percentage of the return on common equity that is not paid as dividend | 0.26 (0.24) | | 38 |
| 37 | TOT_COMMON_EQY | Total shareholder invested amount | -0.00 (-0.18) | | 42 |
| 38 | BEST_SALES | Analyst estimates of revenue | 0.00 (0.29) | | 19 |
| 39 | BS_CASH_NEAR_CASH_ITEM | Cash in vaults and deposits in banks excluding restricted cash, including investments with maturities of less than 90 days | -0.14 (-3.91) | 13 | 20 |
| 40 | Working capital | Self-constructed variable by taking ... | 0.04 (0.68) | | 18 |
| 41 | Company ID | Control variable on the parent company | -0.06 (-1.17) | | 22 |
| 42 | Year | Control variable on the regarded year | 7.64 (2.52) | 14 | 9 |

* All variables are from Bloomberg.

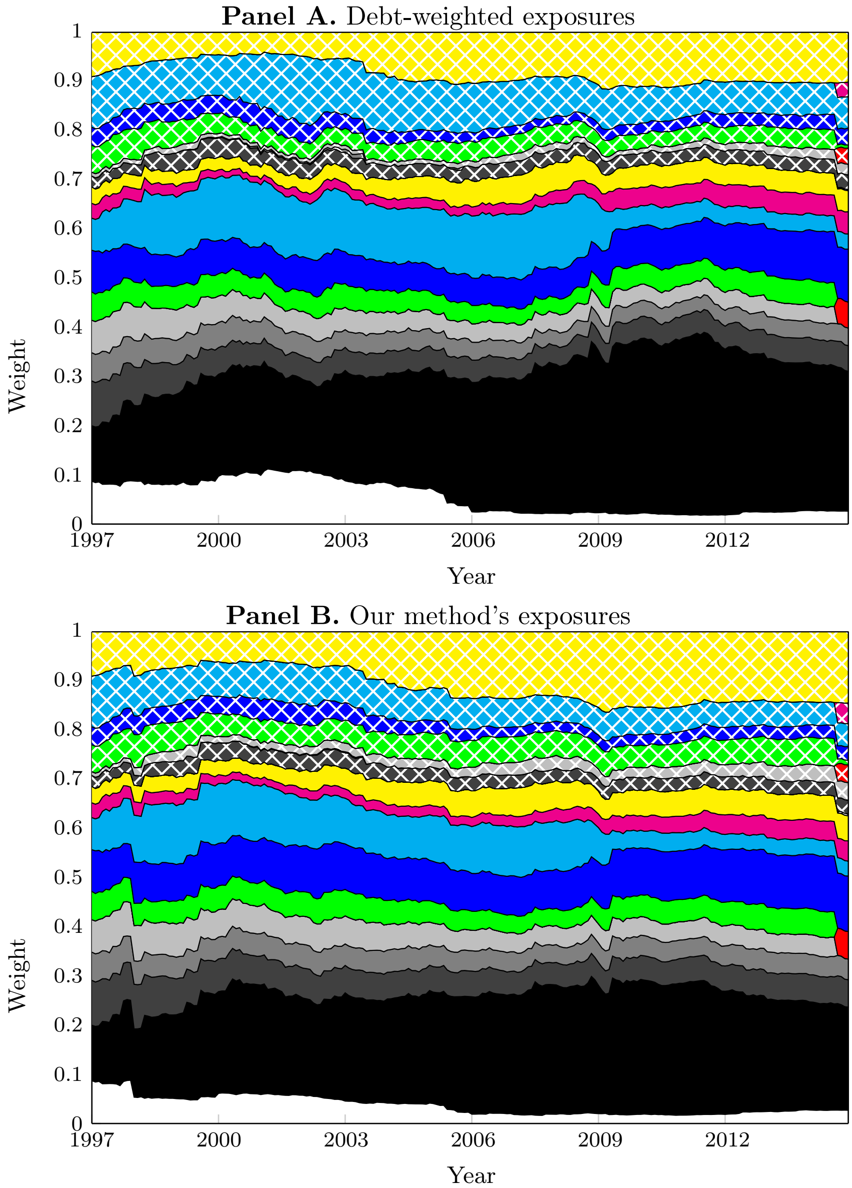
# F Lasso results

**Table 14:** Detail of the Tibshirani (1996) lasso-technique results. Per strictness (given by a value for $t$, following the notation of Tibshirani (1996)) the coefficients are given, all numbers are magnified with $10^5$. Row numbers are regressors of Table 14. Omitted values are zeros, omitted rows are regressors with zeros for all $t$.

| regressor | **Value for Tibshirani (1996)'s $t$** | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 0.03 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.09 | 0.10 | 0.13 | 0.15 | 0.18 | 0.22 | 0.27 | 0.32 | 0.39 | 0.46 | 0.56 | 0.67 | 0.81 | 0.98 | 1.18 |
| 1 | | | | | | | | | | | -0.01 | -0.03 | -0.04 | -0.07 | -0.12 | -0.14 | -0.17 | -0.18 | -2.13 | -5.39 | -6.01 |
| 2 | | | | | | | | | | | | | | | | | | 0.00 | 1.94 | 5.17 | 5.77 |
| 3 | | | | | | | | 2.23 | 4.33 | 9.62 | 11.63 | 15.21 | 16.51 | 18.38 | 21.90 | 22.33 | 21.79 | 21.63 | 20.84 | 18.95 | 18.60 |
| 4 | | | | | | | | | | | | | | | | | | | -0.01 | -0.02 | -0.02 |
| 5 | | | | | 0.00 | 0.02 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.07 | 0.08 | 0.13 | 0.14 | 0.15 | 0.15 | 0.15 | 0.16 | 0.16 |
| 6 | | | | | | | | | | | | | | | 0.00 | 0.03 | 0.05 | 0.05 | 0.03 | 0.00 | -0.00 |
| 7 | | | | | | | | | | | | | | | | | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 |
| 8 | | | | | | 0.00 | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.02 | 0.03 | 0.03 | 0.04 | 0.03 | 0.03 | 0.03 | 0.02 | 0.01 | 0.01 |
| 9 | | | | | | | | | | | | | 0.00 | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.02 |
| 10 | | | | | | 0.02 | 0.02 | 0.04 | 0.04 | 0.06 | 0.07 | 0.08 | 0.09 | 0.10 | 0.12 | 0.16 | 0.33 | 0.40 | 0.51 | 0.63 | 0.66 |
| 11 | | | | | | | | | | | | | | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 12 | | | | | | | | -1.55 | -5.27 | -15.01 | -17.72 | -22.45 | -25.60 | -30.30 | -41.38 | -47.16 | -52.66 | -54.96 | -58.00 | -58.79 | -58.72 |
| 13 | | | | | | | | | 0.07 | 0.32 | 0.39 | 0.52 | 0.57 | 0.67 | 0.88 | 0.95 | 1.04 | 1.07 | 1.10 | 1.12 | 1.11 |
| 14 | | | | | | | | | | | | | | | | | 0.00 | 0.01 | 0.04 | 0.03 | 0.03 |
| 15 | | | | 0.00 | 0.01 | 0.01 | 0.01 | 0.02 | 0.02 | 0.03 | 0.04 | 0.04 | 0.05 | 0.05 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 | 0.07 |
| 16 | | | | | | | | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | -0.00 |
| 17 | | | | | | | | | | | | | | | | | | | 0.00 | 0.00 | 0.00 |
| 18 | | | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.01 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | -0.00 | 0.00 | 0.00 |
| 19 | | | | | | | | | | | | | | 0.00 | 0.01 | 0.02 | 0.02 | 0.03 | 0.03 | 0.04 | 0.05 |
| 20 | | | | | | | | | | | | | | | | -0.03 | -0.19 | -0.25 | -0.34 | -0.43 | -0.45 |
| 21 | | -0.00 | -0.00 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.01 | -0.02 | -0.02 | -0.03 | -0.03 | -0.03 | -0.04 | -0.04 | -0.04 | -0.04 | -0.04 | -0.04 | -0.04 |
| 22 | | | | | | | | | | | | | | | | -1.11 | -1.92 | -2.27 | -1.43 | -1.24 | -1.32 |
| 23 | | | | | | | | | | | | | | | | | | | | -0.11 | -0.14 |
| 24 | | | 0.00 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.03 | 0.06 | 0.06 |
| 25 | | | | | | | -0.02 | -0.09 | -0.11 | -0.17 | -0.22 | -0.29 | -0.32 | -0.37 | -0.47 | -0.51 | -0.54 | -0.54 | -0.52 | -0.48 | -0.48 |
| 26 | | | | | | | | | | | | | | | | | | 0.01 | 0.03 | 0.06 | 0.07 |
| 27 | | | | 0.36 | 0.47 | 0.57 | 0.57 | 0.62 | 0.63 | 0.65 | 0.66 | 0.66 | 0.61 | 0.53 | 0.34 | 0.33 | 0.62 | 0.72 | 0.59 | 0.47 | 0.39 |
| 28 | | | | | | | | | | | | 0.45 | 2.72 | 6.91 | 15.69 | 19.26 | 22.09 | 24.00 | 31.19 | 40.07 | 41.79 |
| 29 | | 2.74 | 5.69 | 7.46 | 7.48 | 7.46 | 7.46 | 7.98 | 8.27 | 8.94 | 9.33 | 9.87 | 9.76 | 9.53 | 8.87 | 8.58 | 8.02 | 7.68 | 7.38 | 6.81 | 6.75 |
| 30 | 1.29 | 6.90 | 8.88 | 10.84 | 11.29 | 12.21 | 12.21 | 13.52 | 14.07 | 15.46 | 15.80 | 16.36 | 16.69 | 17.32 | 18.72 | 19.61 | 20.18 | 20.33 | 20.28 | 20.29 | 20.34 |
| 31 | | | | | | | -0.70 | -2.67 | -3.17 | -4.31 | -4.47 | -4.65 | -5.17 | -5.99 | -7.59 | -8.74 | -9.66 | -9.93 | -9.51 | -9.03 | -9.05 |
| 32 | | | | | | | | | | | | | 0.49 | 1.14 | 2.57 | 3.10 | 3.62 | 3.78 | 3.96 | 1.16 | 0.15 |
| 34 | | | | | | | | | | | | | | -0.00 | -0.02 | -0.02 | -0.03 | -0.03 | -0.02 | 0.00 | 0.01 |
| 35 | | | | | | | | | | | | | | | | | | | | | -0.03 |
| 36 | | | | | | | | | | | | | | | | | | | | 0.07 | 0.14 |
| 38 | | | | | | | | | | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 39 | | | | | | | | | | | -0.01 | -0.02 | -0.02 | -0.04 | -0.07 | -0.08 | -0.09 | -0.09 | -0.11 | -0.13 | -0.13 |
| 40 | | | | | | | | | | 0.01 | 0.01 | 0.02 | 0.02 | 0.03 | 0.04 | 0.04 | 0.05 | 0.05 | 0.05 | 0.05 | 0.05 |
| 41 | | | | | | | | | | | | -0.01 | -0.01 | -0.02 | -0.04 | -0.04 | -0.04 | -0.04 | -0.05 | -0.06 | -0.06 |
| 42 | | | | | 0.34 | 1.18 | 1.18 | 1.97 | 2.26 | 3.02 | 3.42 | 4.11 | 4.38 | 4.83 | 5.84 | 6.25 | 6.56 | 6.74 | 7.15 | 7.56 | 7.64 |

## G Detail of sector exposures

**Figure 16:** Panel A of this figure displays the sector exposures of the debt-weighted index in the fundamental universe, panel B for our method's portfolio. The sectors are from the third level of Bloomberg's classification. The white area corresponds to automotive, then banking, basic industry, capital goods, consumer goods, energy, financial services, healthcare, insurance, leisure, media, non-pfandbriefe covered, real estate, retail, services, technology and electronics, telecommunications, transportation, and finally utility.



## H Detail of the under- and over-weights of our method

**Table 15:** Detail of the largest absolute under- and overweights of our index (as described in Section 7.III) with respect to the debt-index, both in the fundamental universe, at the date of December 2014 and after rebalancing. The deviations are regarded per date and the largest 20 are displayed.

| | | Company weights | | |
|---|---|---|---|---|
| Company name | Ticker | Debt-weighted | Our index | Difference |
| Verizon Inc. | VZ US | 1.30% | 0.49% | 0.81% |
| Goldman Sachs Inc. | GS US | 1.37% | 0.58% | 0.80% |
| J.P. Morgan Chase & Co. | JPM US | 1.32% | 0.59% | 0.74% |
| Morgan Stanley | MS US | 1.10% | 0.46% | 0.64% |
| General Electric Corp. | GELK US | 1.67% | 1.04% | 0.63% |
| Rabobank Group | RABO NA | 1.03% | 0.59% | 0.43% |
| Wells Fargo & Co. | WFC US | 0.85% | 0.45% | 0.39% |
| Petroleo Brasileiro | PETR4 BZ | 0.64% | 0.25% | 0.39% |
| Electricite de France | EDF FP | 0.69% | 0.32% | 0.37% |
| Petroleos Mexicanos | 1232Z MM | 0.56% | 0.26% | 0.30% |
| HSBC Holdings | HSBA LN | 0.56% | 0.26% | 0.29% |
| Wal-Mart Stores Inc. | WMT US | 0.61% | 0.35% | 0.26% |
| Telefonica SA | TEF SM | 0.57% | 0.31% | 0.26% |
| Kyushu Electric Power Inc. | 9508 JP | 0.13% | 0.39% | −0.26% |
| Kansai Electric Power Inc. | 9503 JP | 0.14% | 0.40% | −0.25% |
| Bank of America Corp. | BAC US | 0.59% | 0.34% | 0.25% |
| Oracle Corp. | ORCL US | 0.42% | 0.18% | 0.24% |
| East Japan Railway | 9020 JP | 0.11% | 0.34% | −0.23% |
| Intesa Sanpaolo | ISP IM | 0.52% | 0.29% | 0.22% |
| ING Groep | INGA NA | 0.46% | 0.24% | 0.22% |