Benchmarking Global Fixed Income Portfolios*

Jules H. van Binsbergen† Ralph S.J. Koijen‡

November 5, 2018

Abstract

We study the tradeoffs involved in designing a benchmark for global fixed-income portfolios. Our starting point is a benchmark portfolio based on market-value weights across countries and fixed income segments and we discuss potential reasons for deviations from these market-implied weights. We also study the performance of various other potential benchmarks, including the use of each country's Gross Domestic Product (GDP) as the basis for a benchmark's portfolio weights. Next, we propose an idiosyncratic tail risk measure that is informative on the type of adjustment factors that are required to mitigate concentration risk for any particular region or market segment. Lastly, due to secular trends in fixed income markets as well as the risk profile of corporate bonds, which inevitably involves a Peso problem, we argue that there exists an important role for stress testing.

*November 2018. This report is prepared for the Norwegian Ministry of Finance.

†The Wharton School of the University of Pennsylvania, National Bureau of Economic Research, and Center for Economic and Policy Research

‡University of Chicago, Booth School of Business, National Bureau of Economic Research, and Center for Economic and Policy Research

1 Executive summary

We would like to thank the Norwegian Ministry of Finance for appointing us to this expert group and evaluate the benchmark of the Global Fixed Income Portfolio for the Government Pension Fund Global (GPFG). We are aware of the importance that this Norwegian Government Pension Fund has for the Norwegian people and for the economic future of the country. We believe that the construction of the benchmark portfolio, through its impact on investment decisions, can have significant implications for the potential value that this extraordinary endowment can create going forward.

Our mandate contains several guiding principles for the investment strategy of the fund. These principles, which derive from the overall objective of the Fund to obtain the highest possible long-term financial return within an acceptable level of risk, are as follows:

1. Broad diversification (within listed markets).
2. A moderate degree of active management.
3. Harvesting of risk premiums over time.
4. Responsible management.
5. Cost efficiency.
6. Clear division of roles and responsibilities.

The purpose of the Fund's fixed-income investments is to (i) reduce the volatility of overall Fund returns, (ii) provide liquidity, and (iii) provide exposure to fixed-income risk factors, such as interest rate risk and credit risk.

A thorough investigation and a careful tradeoff of the arguments involved in constructing a fixed-income benchmark is therefore warranted. This report summarizes these various arguments, both from a theoretical and an empirical perspective.

Based on our analysis, we reached the following conclusions and recommendations. We recommend a benchmark that is broadly diversified across geographies and across fixed income segments to minimize risk and maximize investment capacity. In particular in fixed income markets, volatility and correlations can quickly change and, most importantly, in ways that are hard to predict based on historical data alone. Hence, in the spirit of the fund's objective to be broadly diversified and to be exposed to the main risk factors in financial markets via the benchmark, we argue that a portfolio that covers a wide set of countries and the main fixed income segments (treasuries, inflation-linked bonds, corporate bonds, and covered bonds) not only diversifies risks during normal times, but also across changes in market regimes.

We summarize the evidence on the term premium and the credit risk premium. Historically, the term premium has been positive in most, if not all, countries, implying that investors generally require additional compensation for holding long-term bonds relative to short-term bonds. One candidate source of risk is inflation risk as long-term bonds are more affected by changes in inflation. Historically, inflation has varied quite substantially in the U.S, but been remarkably stable over the past years. Whether this is a temporary phenomenon or a permanent feature of U.S. bond markets remains to be seen. Also, there are other countries in the Fund's investment universe that have a different level of inflation volatility. We recommend that the duration of the portfolio is chosen in accordance with the duration that follows from market weights. We view the decision to increase or decrease exposure to the term premium as an active investment management decision.

Recent research provides support for the existence of a credit risk premium after correctly adjusting for term risk. While corporate bonds can be reasonably replicated by a portfolio of risk-free bonds and stocks for long periods of time, this correlation breaks down during certain periods. If one were to restrict the benchmark to nominal government bonds in developed markets only (and perhaps even a subset of currencies within this smaller set), one would omit approximately $23 trillion US dollars in market value compared to approximately $25 trillion US dollars in government bonds in developed markets. Hence, the investment capacity would be reduced and concentration risk would rise as a result.

In summary, a natural starting point for the benchmark is a portfolio that mimics the market portfolio. We discuss potential reasons to deviate from market weights, such as market segmentation, but given that the research on these questions is relatively recent, we currently view such deviations as an active investment management decision.

We have documented the high concentration in fixed income markets in the different segments. It is worth noting that different countries play a prominent role in different fixed income segments. This high degree of concentration may expose the fund to idiosyncratic movements in a country or fixed income segment. Historically, the fund has addressed this concentration concern by using GDP weights in the government bond portfolio. While this measure has perhaps been effective given the past and current distribution of GDP, this may change in the future. For instance, according to the International Monetary Fund (IMF) in 2017, in the U.S., nominal GDP is $19 trillion, it is $12 trillion in China, and $5 trillion in Japan. Bloomberg announced in March of 2018 that China will be included in the Bloomberg Barclays Global Aggregate Index as of next year, which would make it the second-largest position in the benchmark if no further adjustments are made.

Risk measurement in fixed income markets is challenging as backward-looking risk measures may not reveal the potential risks going forward. While prices and / or quantities in fixed-income markets may look unusual compared to historical data, such episodes are often also characterized by arguments in favor or against why "this time is different." For instance, the convergence in European treasury yields following the introduction of the Euro provides a useful example. From a risk measurement and management perspective, this is a major challenge and perhaps more challenging for fixed income markets as we have seen a secular decline for almost four decades in interest rates and the path going forward is uncertain.

To address the last two items, we propose an idiosyncratic tail risk exposure measure for each country that depends on three inputs: (1) the country's market portfolio weight, (2) the country's government bond duration and (3) adverse yield change scenarios. We have used the available historical data across countries to deduce a set of extreme scenarios that is the same for all countries. This

approach recognizes that it is very hard to predict in which country or region the next extreme interest rate changes will occur, particularly over longer horizons, and it is therefore useful to apply the most extreme yield changes in the whole panel to generate the scenarios. We note that it is of course possible to model the scenarios as a function of country or region characteristics such as fiscal strength if one has a particularly strong view on where the next crisis is most likely to occur. We also note that that these idiosyncratic tail risk measures could be used to construct adjustment factors. We believe such explicit views belong in the active management discussion of the fund rather than in the discussions regarding the benchmark.

In addition to using historical data, we recommend additional scenario analysis and stress testing as a potential tool to highlight the potential risks in the current portfolio. Stress testing is being used more extensively in the context of bank and insurance company supervision, and we think that the insights from these analyses are helpful to complement traditional risk measures, both in terms of understanding the benchmark risks and the active risks.

2 Preface

We would like to thank the Norwegian Ministry of Finance for appointing us to this expert group and evaluate the benchmark of the Global Fixed Income Portfolio for the Government Pension Fund Global (GPFG). We are aware of the importance that this Norwegian Government Pension Fund has for the Norwegian people and for the economic future of the country. We believe that the construction of the benchmark portfolio, through its impact on investment decisions, can have significant implications for the potential value that this extraordinary endowment can create going forward. A thorough investigation and a careful tradeoff of the arguments involved in constructing such a benchmark is therefore warranted. This report summarizes these various arguments, both from a theoretical and an empirical perspective.

3 Mandate

3.1 Guiding principles

Before presenting the report, we review the mandate (enclosed in the Appendix) we were given in writing this report. The mandate is to design and/or evaluate possible fixed-income benchmarks that the fund could use for its investments, including the benchmark that is currently in use. In this light, the mandate contains several guiding principles for the investment strategy of the fund. These guiding principles, which derive from the overall objective of the Fund to obtain the highest possible long-term financial return within an acceptable level of risk, are as follows:

1. Broad diversification (within listed markets).
2. A moderate degree of active management.
3. Harvesting of risk premiums over time.
4. Responsible management.
5. Cost efficiency.
6. Clear division of roles and responsibilities.

Diversification benefits can be achieved not only by investing broadly across asset classes, but also by investing broadly within asset classes. In the context of the fixed income category, this implies diversification across countries, currencies, and fixed-income segments. Another important consideration is that the fund has a moderate (to low) level of active management. This implies that currently the riskiness of the Fund's portfolio is to a very large extent driven by the choice of benchmark portfolio, and much less so by the relatively small active deviations from the benchmark. A question is whether these relatively small active deviations will continue to be small forward. If they do, then changes in the choice of benchmark portfolio will translate almost one-for-one to changes in the investment of the fund.

The guiding principles above broadly apply to the investments within and across asset classes, not just fixed income. In addition, there are several considerations specific to fixed income that need to be taken into account. In particular, the purpose of the Fund's fixed-income investments is to (i) reduce the volatility of overall Fund returns, (ii) provide liquidity, and (iii) provide exposure to fixed-income risk factors, such as interest rate risk and credit risk.

In this report we will discuss the various available benchmark choices in the context of the objective of the fund, as well as the supporting guiding principles that help achieve this objective where relevant.

3.2 Current benchmark strategy, the Bank's mandate, and risk limits

The current fixed-income benchmark index was adopted in 2012 on the basis of a strategic equity share of 60 percent, and reflects the stated purposes. The benchmark is based on index products provided by Bloomberg L.P., and comprises a government bond portion (70 percent) and a corporate bond portion (30 percent). The 70/30 shares are fixed and the portfolio is fully rebalanced to these weights at a monthly frequency. The fixed-income benchmark is exclusively comprised of investment-grade securities. Bonds from Norwegian issuers and bonds denominated in Norwegian kroner are excluded from the benchmark index.

In terms of fixed-income segments, Figure 1 summarizes the way the index provider categorizes the various segments in the Bloomberg Barclays fixed-income flagship index. We explore in this report whether segments should be included or omitted from the Fund's benchmark.

Government bond portfolio The government bond portfolio consists of nominal government bonds, inflation-linked government bonds, and bonds issued by supranationals that are included in the above-mentioned indices at a given point in time (but excluding Norwegian issuers and bonds denominated in NOK).[1] The weights of each country within the government bond portion of the

[1] A further exception has been made for Indonesian Rupiah, which was included in the underlying Bloomberg Barclays Global Treasury GDP Weighted by Country Index as of June 1, 2018. An adjustment factor of 0 was set to Indonesia in the benchmark for the GPFG, pending the decision on a new fixed-income benchmark to be presented to Parliament

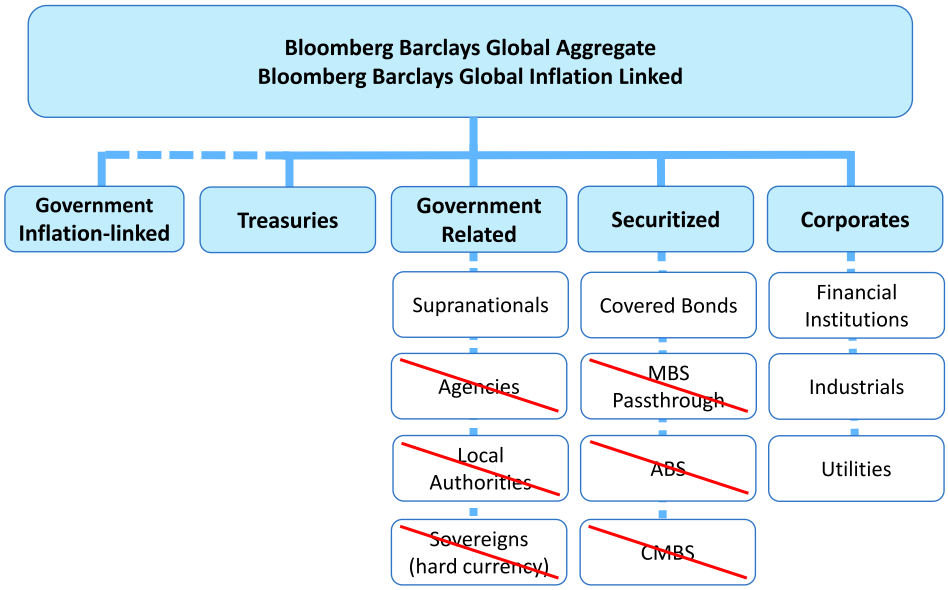

Figure 1: Segment classification fixed income within the flagship index of Bloomberg Barclays Global Aggregate index (+ Global Inflation Linked). The crossed out segments (in red) are the segments excluded from the fixed income benchmark as reported in Meld. St. 17 (2011-2012) Report to the Storting (white paper)

portfolio is calculated annually on the basis of the size of the economy of each country, according to the methodology of Bloomberg Barclays Global Treasury GDP Weighted by Country index.[2] These country weights are rebalanced back to the original weights on a monthly basis. Within each country, sub-segments and individual bonds are weighted by market weights. Certain country weights in the government bond portion are supplemented by adjustment factors motivated by investability considerations and to limit concentration risk.[3]

in the spring of 2019.

[2]However, there are certain adjustments within the Euro area due to bonds issued by Supranationals in Euro. There are separate country weights for such bonds and concomitant reductions in the country weights within the Euro area.

[3]An adjustment factor of 0.25 is applied to the country weights for Chile, Hong Kong and Russia. An adjustment factor of 0 has been applied to Indonesia.

Corporate bond portfolio The corporate bond portion comprises corporate bonds and covered bonds issued in seven approved currencies.[4] Within the corporate bond sub-index, the bonds are weighted using the methodology for the Bloomberg Barclays Global Aggregate Bond Index.

The Bank's mandate and risk constraints In its management mandate, the Ministry has defined an investment universe for fixed-income instruments, which is broader than the benchmark index itself. The Bank shall seek to keep any deviations from the benchmark index in the composition of the actual portfolio within the mandated limit for expected tracking error of 1.25 percentage point. This limit is given for equities and bonds combined.

In addition, the Bank is required to have limits for the minimum overlap between the actual fixed-income portfolio and the fixed-income benchmark, separate limits for tail risk, and credit risk limits at the portfolio level as well as for individual investments, and take differences in fiscal strength across countries into account in the management of government bonds. In order to ensure that Norges Bank is not forced to immediately divest bonds that are dropped from the benchmark index as the result of a credit rating downgrade below investment grade, it is permitted to hold up to 5 percent of the fixed-income portfolio in high-yield bonds.

3.3 Questions to be answered

Based on our mandate and the above principles and guidelines, we set out to answer the following questions:[5]

1. What are the considerations of various portfolio construction rules and which currencies should be included in the index?

2. Which segments should be included in the benchmark?

[4] US Dollar (USD), the Euro (EUR), the British Pound (GBP), the Swedisch Krona (SEK), the Danish Krone (DKK), the Swiss Franc (CHF), and the Canadian Dollar (CAD).

[5] See the document titled "Terms of reference for an expert group appointed to review the fixed-income investment framework for the Government Pension Fund Global" which is attached to this report.

3. How should the duration (that is, the sensitivity to interest rate risk) of the bond portfolio be chosen?

4. Should inflation-linked bonds be included in the index?

5. Are there risk premiums which should not be harvested through the benchmark index?

Ad 1: We are requested to analyze how various country and currency compositions in the fixed-income benchmark may contribute to meeting the principle of broad diversification. The analysis shall be conducted on the basis of various weighting principles, including market weights (with or without adjustment factors) and GDP weights. Advantages and disadvantages of the various weighting principles shall be discussed. The significance of emerging markets for diversification shall be addressed. The group shall also assess whether the long investment horizon of the Fund suggests that the index rules should pay special heed to capturing any changes in the fiscal strength of government bond issuers over time.

GDP weights may lead to excessive turnover due to rebalancing, as Norges Bank has pointed out in its advice to the Ministry. Alternative weighting schemes could take into account the market capitalization of the country.

Ad 2: Our group is requested to assess whether other segments than nominal government bonds from developed economies should be included in the benchmark index, including corporate bonds, covered bonds, bonds issued by supranationals, and nominal government bonds from emerging economies issued in local currency. The group is requested to assess the expected credit premium in this context, and how one should ensure exposure to such premium.

Norges Bank has pointed out in its advice to the Ministry that emerging market countries do drop in and out of the index occasionally and some markets may not be investable for the fund due to a high GDP compared to the market size of local currency government bonds. The bank has also pointed out that smaller corporate bond markets (e.g. Switzerland) are challenging to invest in.

Ad 3: Fixed income securities have a fixed maturity. The average time it takes an investor to recover

the cash flows of an investment is called the duration. It is well-known that the duration of a fixed income investment measures its sensitivity to interest rate changes. It is therefore a measure of interest rate risk. The duration choice is important because by investing in longer-term securities the so-called term premium (i.e., the difference between long-term and short-term bond returns) can be earned. In this context, we are asked to assess what the expected term premium is, and how exposure to that premium is ensured. The ministry has further asked whether there is a reason to assume segmentation in the bond market and whether such segmentation would influence the index rules.

Ad 4: Our group is requested to assess whether inflation-linked government bonds should be included in the fixed-income benchmark.

Ad 5: If there are bond market risk premiums to which the GPFG should be exposed, in addition to term and credit premiums, but which are not suited for inclusion in the benchmark index, an assessment is requested as to how the Ministry should stipulate risk limits in the mandate for any exposure to such premiums.

A recent practitioner and academic literature shows that expected returns and (co)variances vary with characteristics, also called factors (see Ang (2014) for an introduction). In some cases, as we will discuss below, these factors capture term and credit risk premia, but new factors have been discovered in recent years that may improve risk-return opportunities. Depending on the factor, a certain degree of dynamic trading is required to implement factor-based investing, which may make them more appropriate for the active part of the portfolio as opposed to the benchmark.

4 Data

4.1 Data from Bloomberg

We use data provided to us by Bloomberg, which cover most segments of the global fixed income market, with the exception of securitized assets. We summarize the coverage in this section by segment.

For each country and segment, we have data on the yield-to-maturity (YTM), the coupon return, the price return, the total return, the duration, amount outstanding (in terms of face value), the market value, the credit rating, and the number of issues used to compute returns. While not all fields are always available, in most cases they are.

Nominal government bonds We have data from on Treasury bonds from Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Hong Kong, Ireland, Italy, Japan, Latvia, Lithuania, Luxembourg, Netherlands, New Zealand, Norway, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, the U.K., and the U.S.[6]

For emerging markets, we have data from Argentina, Brazil, Chile, China, Colombia, Croatia, Czech Republic, Egypt, India, Indonesia, Israel, Malaysia, Mexico, Nigeria, Offshore China, Peru, Philippines, Poland, Romania, Russia, South Africa, South Korea, Taiwan, Thailand, and Turkey.

Table A1 summarizes the sample available for different countries where at least one issue is available as provided to us for developed markets. Table A2 summarizes the same information for emerging markets.

Inflation-linked government bonds We have data on inflation-linked bonds from Australia, Canada, Denmark, France, Germany, Greece, Italy, Japan, New Zealand, Spain, Sweden, the U.K., and the U.S.

Table A3 summarizes the sample available for different countries where at least one issue is available as provided to us.

Government-related fixed income securities We have data on bonds issued by agencies, local authorities, sovereigns, and supra-nationals, where the definitions follow the Bloomberg Global Aggregate Index.

[6]We include Norway for completeness, but excluding it has a minimal impact on the results.

Table A4 summarizes the sample available where at least one issue is available as provided to us.

Corporate bonds We have data on investment-grade corporate bonds from Australia, Austria, Belgium, Canada, France, Germany, Italy, Japan, Mexico, the Netherlands, South Korea, Spain, Sweden, Switzerland, the U.K., and the U.S.[7]

Table A5 summarizes the sample available for different countries where at least one issue is available as provided to us.

Covered bonds We have data on covered bonds from Austria, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, the U.K., and the U.S.

Table A6 summarizes the sample available for different countries where at least one issue is available as provided to us.

4.2 GDP data

We use nominal GDP from the World Bank, which is revised annually, in USD. We use the December values of a given year to construct weights for the subsequent year.[8]

4.3 Factor returns

Data on momentum and value factors, using country-level returns, is from AQR's data library following the construction of Asness, Moskowitz, and Pedersen (2013). For equity (fixed income) factors, the sample is from February 1980 (January 1983) until December 2017. The return on carry

[7]The corporate sub-index in the GPFGs fixed-income benchmark is based on the currency denomination of the bond, not the issuance country. The corporate sub-index currently includes 7 approved currencies, issued by corporations spread across in excess of 50 countries.

[8]The Global Aggregate Index uses weights based on average GDP during the last three years with declining weights. Given the persistence in GDP levels, such adjustments are likely to have a minor impact on the results.

strategies are based on the carry1-12 strategy of Koijen, Moskowitz, Pedersen, and Vrugt (2018) using country-level returns. The fixed income factor starts in November 1983. The series both end in December 2017. The low-beta factors are taken from Frazzini and Pedersen (2014). The low-beta equity factor starts in February 1980 and the fixed income factor starts in July 1989. Both factors end in March 2012. All factors are long-short strategies that are on average close to market neutral and the returns are currency hedged to US dollars. We refer to the original papers for further details on how the factors are constructed.

5 Institutional background

We describe in this section the institutional environment of the fund, including the performance of the fund across fixed income and equity, the degree of activeness of the portfolio, as well as the constraints that are imposed in the investment process.

5.1 Current overall portfolio

The Fund is globally diversifie-d across various asset classes, including equity, fixed, income, and unlisted real estate. In Figure 2 and 3, we plot the returns on the fixed income portion of the Fund's portfolio and compare them with the benchmark. Figure 2 plots simple returns and Figure 3 plots those returns cumulatively. While the graph illustrates that the fund has outperformed the benchmark, it also shows that the returns have closely tracked the benchmark's return. One marked period of underperformance was the financial crisis in 2008 and 2009.

A very similar picture emerges when studying the equity portion of the portfolio. In Figure 4 and 5, we plot the simple and cumulative returns on the fund as well as on the benchmark for this part of the portfolio. Once again, over this 20 years period the fund has outperformed the benchmark, while closing following its risk profile.

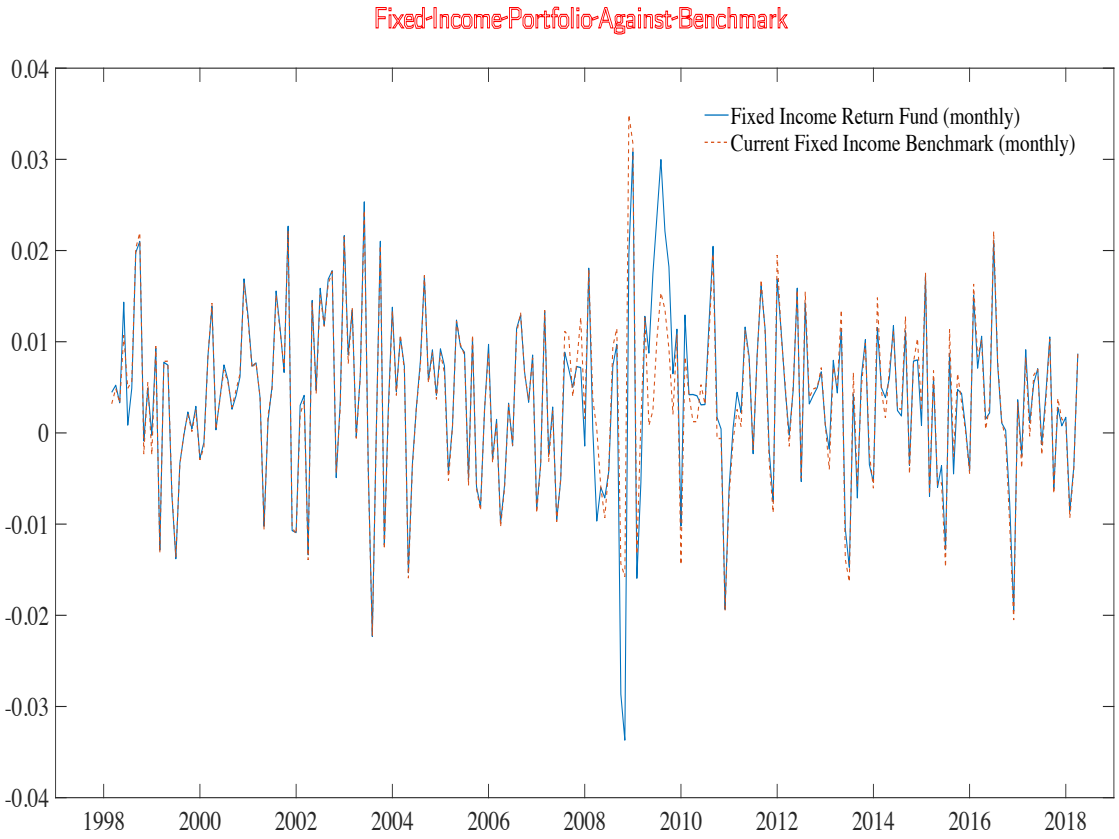

Figure 2: Monthly returns on the fixed income portion of the Fund's investments plotted against the fixed income benchmark using data between February 1998 and March 2018.

To formalize the distance between the Fund's return and the benchmark, we compute in Table 1 the fraction of the variance of the fund's fixed income and equity returns that can be attributed to the benchmark return, where the remainder is due to the active positions by which the fund deviates from that benchmark. The results show that in both case the vast majority of the variability is attributable to the benchmark, though less so for fixed income than for equity. For the equity returns, a mere 0.25% of the variability (or variance) of returns is attributable to active positions, whereas for fixed income, this fraction is higher at 9.58%. As argued before, this close relation between the fund's return and the benchmark return, highlights the importance of the choice of benchmark, though even more so for equity than for fixed income historically.

One may wonder to what extent the activeness of the portfolio has varied over time. That is,

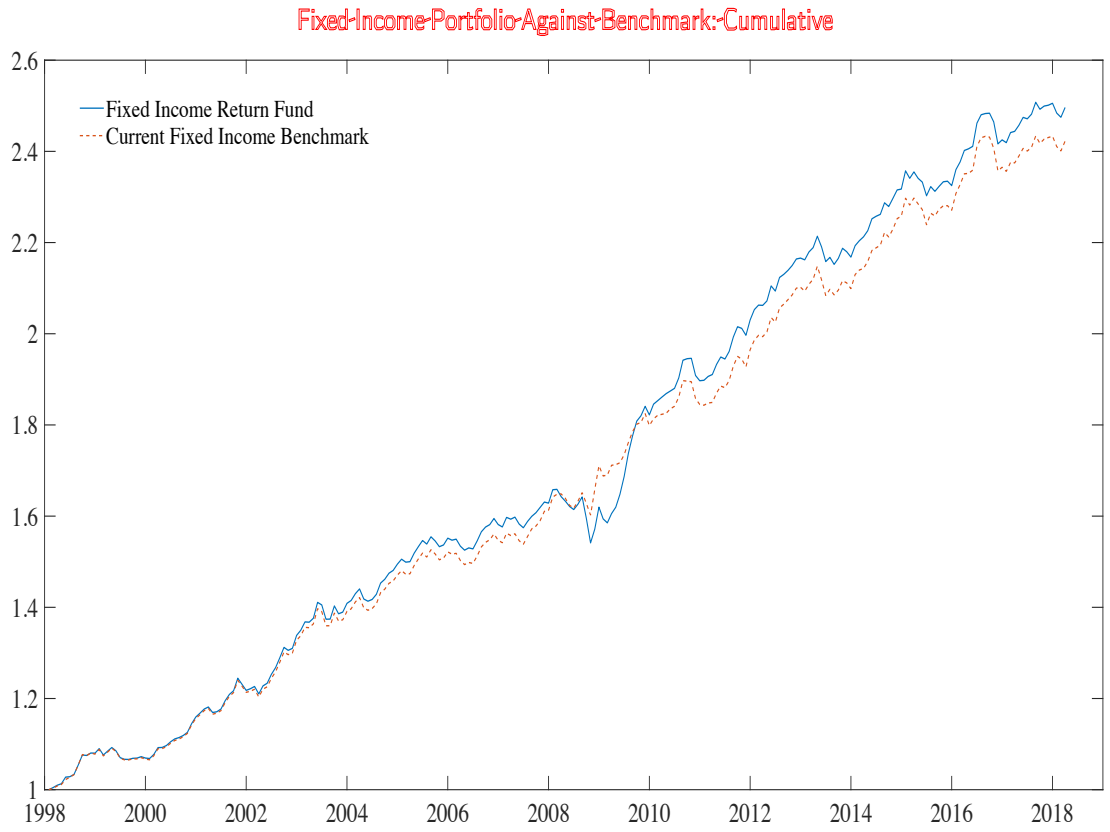

Figure 3: Cumulative returns on the fixed income portion of the Fund's investments plotted against the fixed income benchmark using data between February 1998 and March 2018.

the numbers in Table 1 represent an average level of activeness, but this could potentially vary over time. In Figure 6 we compute the same statistics as in Table 1, but now on a 2-year rolling basis. For example, the numbers plotted in December 2017 are based on data between January 2015 and December 2017.

The picture shows a marked difference between the equity and fixed income portfolio. Whereas for equities, the fraction of return explained by the benchmark has stayed consistently high, even during the recent financial crisis, this is not true for the fixed income part of the portfolio. The graph shows that the fixed income portfolio return has also stayed very close to the benchmark for most years. Before 2008, the numbers between fixed income and equity are comparable. The difference between the two is almost entirely driven by the financial crisis period. We can conclude that for both equity

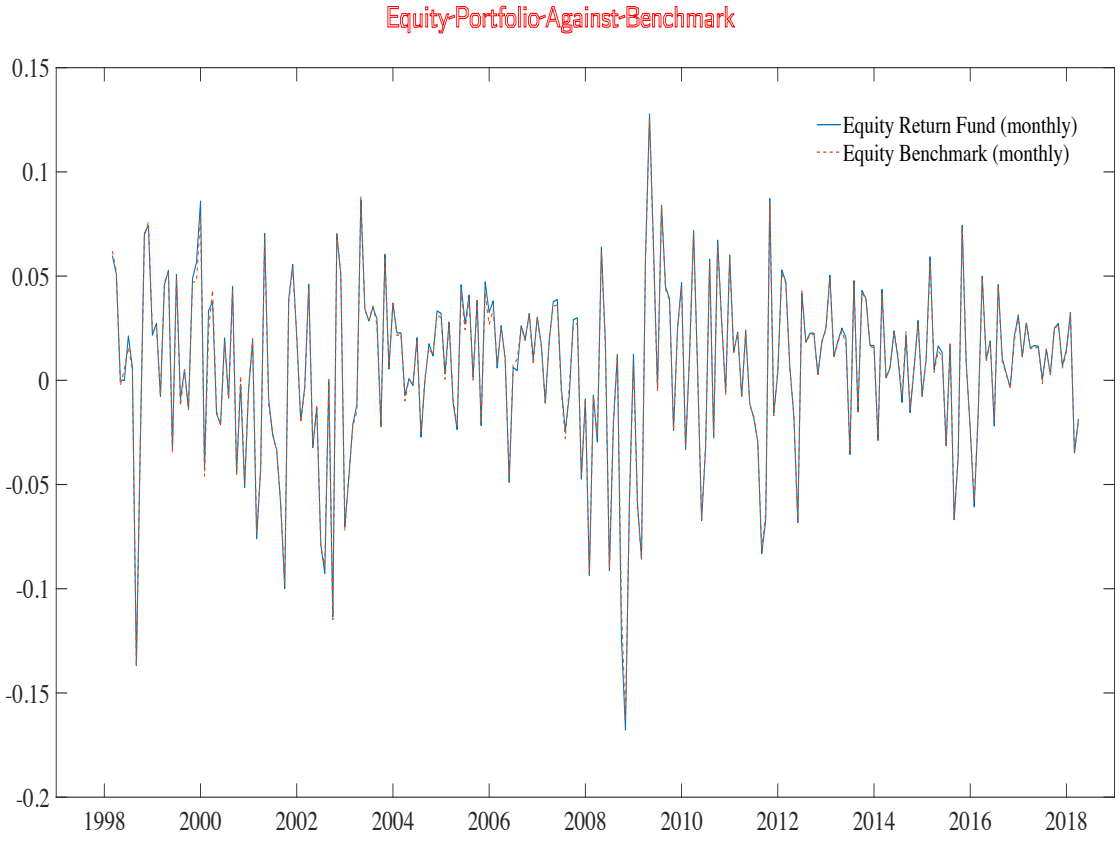

Figure 4: Monthly returns on the equity portion of the Fund's investments plotted against the equity benchmark using data between February 1998 and March 2018.

and fixed income, the fund's return has closely tracked that of the index, particularly outside of the financial crisis episode.

6 Designing benchmarks and risk limits: Theory

In this section, we briefly summarize several theoretical considerations in designing fixed-income benchmarks and risk limits.

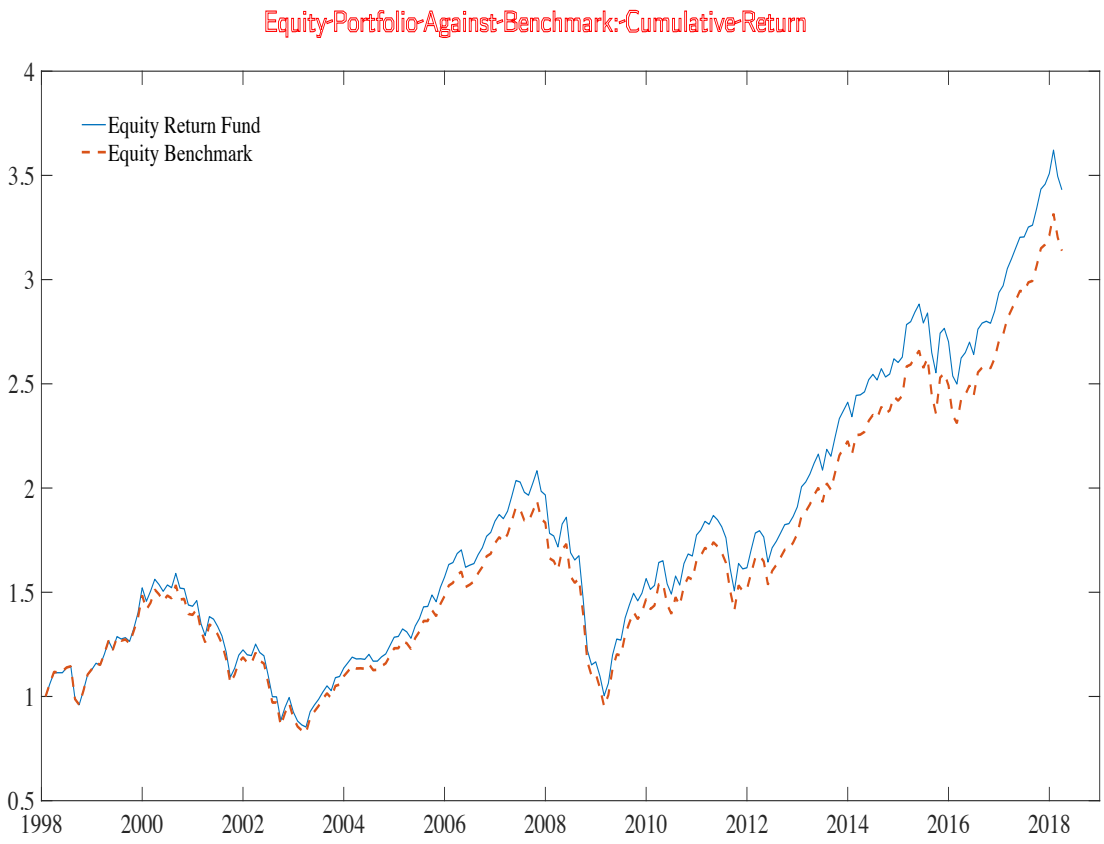

Figure 5: Cumulative returns on the equity portion of the Fund's investments plotted against the equity benchmark using data between February 1998 and March 2018.

6.1 The market portfolio and reasons to deviate

The common starting point in designing benchmarks is the market portfolio, following the insights from the capital asset pricing model (CAPM) (Sharpe, 1964). In the context of fixed income, it is useful to first define the market portfolio. Starting from a single country, government bonds are in zero-net supply and cancel from a representative agent's market portfolio. Fixed income instruments issued by non-financial firms or banks (e.g., covered bonds), however, do not and are part of the market portfolio. If investors are identical (up to risk aversion), they will hold the same risky market portfolio.

If investors face different risks outside of their portfolios, for instance coming from liabilities in case of insurance companies and pension funds, then this creates additional demand for certain types of

Table 1: Variance Decomposition of Returns: 1998-2018.

	Fixed Income	Equity
Benchmark	90.42%	99.75%
Active	9.58%	0.25%
Total	100%	100%

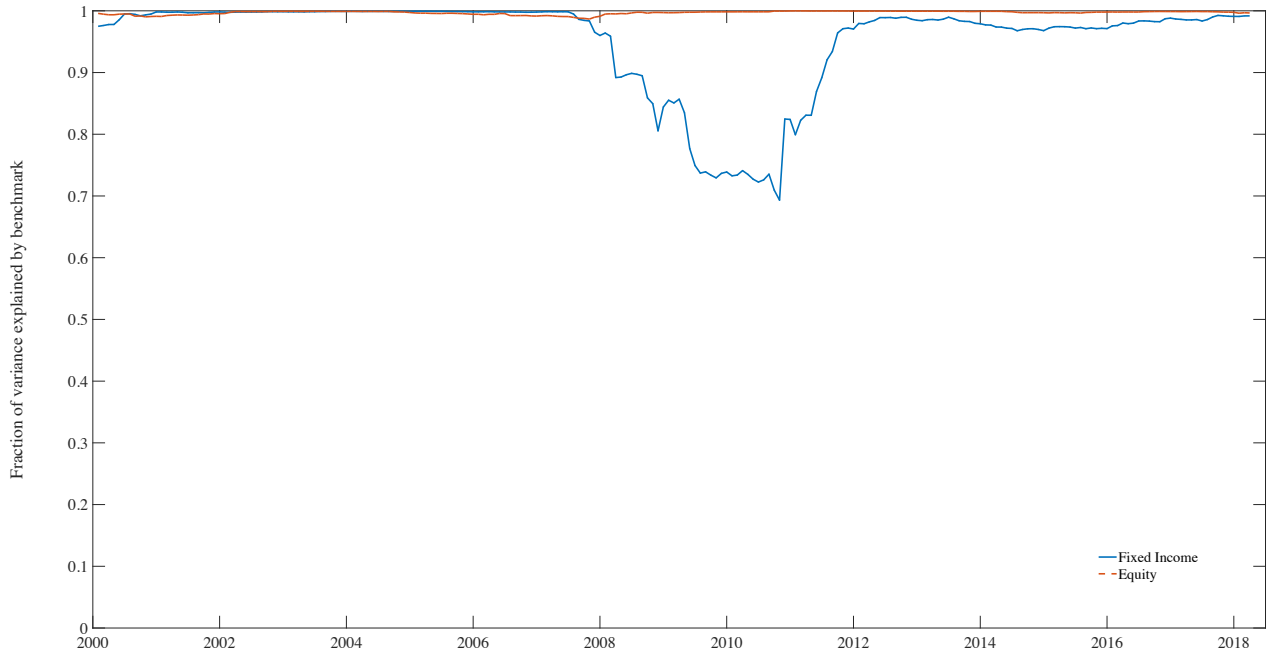

Figure 6: Fraction of return variance explained by the benchmark using a two-year rolling window using data between 1998-2018.

bonds. Depending on investors' willingness to substitute across bonds with different maturities, credit quality, and bonds issued in different currencies, such hedging demands can have significant price effects.[9] Vayanos and Vila (2009) develop a simple preferred-habitat asset pricing model in which one group of investors has inelastic demand for certain maturities and illustrate the impact on equilibrium asset prices.

A growing literature illustrates the empirical relevance of institutional and regulatory constraints,

[9]The willingness of investors to substitute across bonds with different characteristics is central in the debate about the impact of quantitative easing programs, see for instance Krishnamurthy and Vissing-Jorgensen (2011), Maggio, Kermani, and Palmer (2016), and Koijen, Koulischer, Nguyen, and Yogo (2018).

where some of the best evidence comes from bond markets, see for instance Ellul, Jotikasthira, and Lundblad (2011) and Greenwood and Vissing-Jorgensen (2018). Maggiori, Neiman, and Schreger (2018) furthermore show that investors have a strong currency bias in fixed income portfolios, beyond the traditional home bias that has been well documented. Krishnamurthy and Vissing-Jorgensen (2007) and Koijen, Koulischer, Nguyen, and Yogo (2018) directly estimate the demand system for government bonds by investor type and document significant heterogeneity in investor preferences.

The insights from this literature could be relevant in designing benchmarks. If the demand of certain investors is tilted towards bonds with certain characteristics, and if this group is sufficiently large to significantly affect prices, then an unconstrained investor can take advantage of attractive risk-return opportunities that arise as a result. This may provide a reason to deviate from strict market weights. Thus in answer to Section 3.3 question 3 (and particularly the discussion in Ad. 3), we conclude that market segmentation could potentially play a role in the investment strategy of the fund as well as the design of the benchmark, and a deviation from strict market weights could be desirable. However, while the previous literature documents an impact of regional regulatory frameworks on asset prices, it remains challenging to measure the quantitative impact on risk premia across countries and fixed-income segments. Taking advantage of such local demand effects, which may gradually dissipate over time, fits in our view currently better in the active mandate than in the benchmark.

Currently, the fund does in fact deviate from market weights in the government bond portfolio by using GDP weights. However, while this may avoid concentrated allocations to countries with large amounts of sovereign debt, it is not obvious that these particular tilts improve the risk-return tradeoff. For tilts to improve the risk-return tradeoff, they need to relate positively to future (expected) returns, or negatively to future levels of risk. The question is whether this is true for the tilts induced by using GDP weights. We will revisit this question in the empirical analysis below.

6.2 Potential drawbacks of using market weights

A potential concern with market weights in fixed income markets is that some countries have a large amount of debt (e.g., Japan), which may expose the Fund to idiosyncratic shocks. In response, the Fund currently uses GDP weights and applies adjustment factors to these weights in the case of investability concerns. However, any alternative that does not directly confront the issue of concentration risk may be problematic at some point in the future. For instance, as we discuss below, GDP weights may work well given the current composition of the index, but China would receive a very large weight (compared to its market weight) if it becomes part of the index due to its large GDP share. Similar challenges may arise with other ad-hoc weighting schemes, such as for instance using the equity market capitalization of a country as a weighting scheme. While they may look appealing in today's market environment, such rule may require adjustments if the distribution of market equity changes, which seems undesirable to us.

We therefore propose in Section 10 a way to compute adjustment factors to strict market weights that are directly motivated to mitigate exposures to idiosyncratic factors. This approach is consistent with the objective of diversification while harvesting factor premia.

Lastly, one concern sometimes expressed with strict market weights is that the portfolio is tilted towards countries that are fiscally irresponsible or companies that borrow excessively. However, it is important to note that prices (and maturity) adjust accordingly. Deviations away from market weights can therefore be viewed as active positions (for instance, that the market underestimates the fiscal risk or inflation risk of a country) and we think those are therefore more naturally taken in the active portfolio instead of the benchmark.

6.3 Challenges in measuring risk in fixed income markets

The combined equity and fixed income active portfolio of the Fund can deviate from the benchmark up to a tracking error limit of 1.25 percentage points. In fixed income markets, short-term and

backward-looking risk constraints may not accurately reflect the risk over the residual period of the bond. A bond is an investment in a risk-free bond and a short position in a put option. The latter represents the credit risk of the bond. For high-quality bonds this put option is well out-of-the-money (OTM). This implies that corporate bonds share some of the risk measurement challenges that writing OTM puts have; the returns are attractive and stable for potentially long periods of time, but can lead to large losses in bad times.

As a simple illustration, we plot the yield on 10-year government bonds from Germany and Italy from January 1990 until June 2018 using data from Eurostat in Figure 7. Yields converge rapidly leading up to the introduction of the Euro, seemingly marking a new period of sovereign stability in the Euro area. For close to a decade, German and Italian yields move in lock step, resulting in traditional risk measures, for instance volatility and yield spreads, to imply that both countries are equally risky. However, these measures would have underestimated the potential risks going forward following the financial crisis, but primarily following the European sovereign debt crisis in 2011 and 2012.

As part of the first question of the mandate, one may wonder whether measures of fiscal strength could be useful in constructing benchmark weights. However, as this episode illustrates, this may not be an easy task. A fast-growing literature tries to forecast financial crisis based on measures of credit growth (Schularick and Taylor (2012), Mian, Sufi, and Verner (2017). One paper that is particularly relevant in this context is Krishnamurthy and Muir (2016), who show that the joint dynamics of credit growth and changes in spreads of corporate bonds is informative about the risk of financial crises. If credit growth is demand driven, the spread and quantity go up simultaneously. If credit growth is supply driven, however, credit growth and changes in spreads are negatively correlated. These periods are more likely followed by financial crises, as well as longer and deeper recessions.

Another metric that may be useful to complement traditional risk measures is the composition of credit issued. Greenwood and Hanson (2013) show for instance that the face-value weighted average rating changes over the credit cycle. At the end of the expansion, low-quality firms are able to issue,

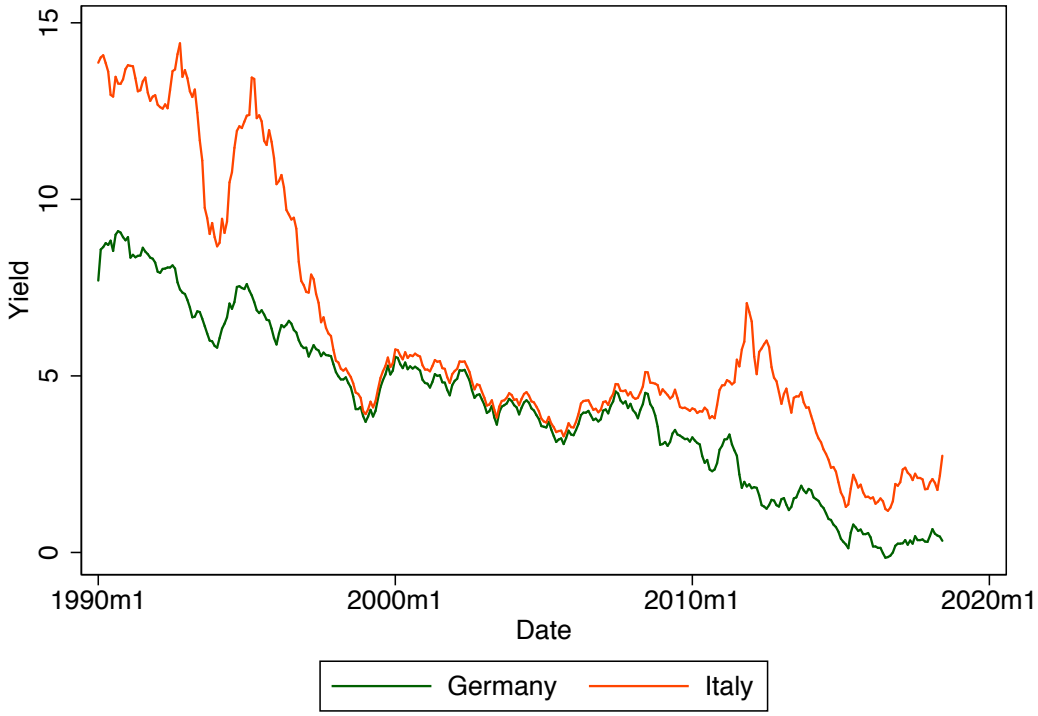

Figure 7: 10-year yields in Germany and Italy from January 1990 until June 2018.

while during economic downturns, mostly high-quality firms issue debt.

These facts are interesting in particular in light of the recent growth in the BBB segment of the market, which now amounts to a larger fraction of the investment-grade corporate bond market in the U.S., as measured by face value compared to 2010. BBB is the cutoff rating for a bond to qualify as investment grade. In addition, there has been a significant growth in the presence of mutual funds and exchange-traded funds (ETFs) in the U.S. market, which often have fairly strict mandates in terms of the fraction of their portfolios that need to be allocated to investment-grade bonds. Putting all this together could point to a potential fragility in the corporate bond market.

In order to complement traditional risk measures, we propose to use stress tests later on in the report. We show how to implement these tests by using stress scenarios from other countries and how to augment it with scenarios one may be worried about going forward (but that have not happened in the past). By testing how the benchmark and active portfolio behaves during those hypothetical stress scenarios, one will be able to get a more complete picture of the potential risks.

7 The risk-return tradeoff in fixed income markets

The literature on factor investing aims to summarize the important dimensions of risk and returns in financial markets. In this section, we summarize parts of this literature, and provide supportive empirical evidence of factors that have been explored in global fixed income markets.

We start with a discussion of the evidence on traditional risk premia in fixed income markets: the term premium and the credit risk premium. We then discuss additional factors that have been explored in recent years.

7.1 Evidence on term and credit premia

We first discuss the evidence on the term premium. Let us first define what the term premium is. One often-used definition of the term premium is the excess yield that investors require to commit to holding a long-term bond instead of a series of shorter-term bonds. Suppose that the current 10-year US government bond yield is 2.5% and that investors expect the average 1-year bond yield to be 2% over the next ten years. Then today's 10-minus-1 year term premium is equal to 2.5%-2% = 0.5%. When computing measures for the term premium, any long maturity bond portfolio can be compared with a short maturity bond portfolio.

Historically, the term premium has been positive in most, if not all, countries, implying that investors generally require additional compensation for holding long-term bonds relative to short-term bonds. One candidate source of risk that requires such risk compensation for longer term bonds is inflation risk as long-term bonds are more affected by changes in inflation. Thus, at least part of the term premium is explained by the so-called inflation risk premium. Historically, inflation has varied quite substantially in the U.S., with inflation rates in the double-digits in the late seventies and early eighties. That said, it has been remarkably stable over the past years. This would suggest that currently less compensation is required for that type of risk, leading to a lower term premium. Whether the stability in U.S. inflation rates is a temporary phenomenon or a permanent feature of

U.S. bond markets remains to be seen.[10] Also, there are other countries in the Fund's investment universe that have a different level of inflation volatility.

As part of a recent study, Brooks and Moskowitz (2017) examine the properties of government bond yields across a number of countries, including Australia, Belgium, Canada, Japan, Sweden, the UK and the US over the sample period 1972-2015. We report their results in Table 2. They find that over this period, the average excess return of 10-year government bonds over the short-term risk-free rate is positive in all 7 countries. The average spread across the 7 countries is 1.2%. Given the reported volatilities, the average over this sample period is statistically significant at the 5% level for all countries with the exception of the UK.

Table 2: Evidence on the term spread

	Excess Returns 10-year Bonds						
	AU	BD	CN	JP	SD	UK	US
Mean	1.31%	1.08%	1.26%	1.14%	1.59%	0.88%	1.07%
St. Dev	5.62%	4.63%	4.74%	3.93%	5.06%	6.14%	6.31%
t-stat	2.50	3.06	2.9	3.21	3.01	1.76	2.24

As we will discuss further later in this report, several important secular trends have affected fixed-income markets in the past decades. As a consequence, caution is in order and it should not be assumed that historical average returns are necessarily indicative of the returns that can be obtained in the future.

Next, we summarize the evidence on the so-called credit risk premium. The credit risk premium is the compensation investors receive for taking on credit risk. Consider the following example. Suppose a corporate bond with a face value of $100 expires in one year from today and it has no remaining coupon payments (coupons are annual). Further, the risk-free interest rate is 0% and there is a 97% chance that the firm will pay back the face value of the bond. If the bond currently trades for $97 (that is, the expected payoff of the bond), then the credit risk premium is zero. For any price below $97, there is a credit risk premium. Suppose that the current price is $96, then the credit risk premium

[10]There is no conclusive evidence that we are aware of that shows that the term premium, or the credit risk premium for that matter, has disappeared in global fixed income markets.

equals 1.04%, computed as:

$$\text{Credit Risk Premium} = \frac{97}{96} - 1 = 0.0104. \tag{1}$$

In recent work, Asvanunt and Richardson (2017) find strong evidence of the existence of a credit risk premium after correctly adjusting for term risk. They use data spanning 80 years in the U.S., and nearly 20 years in Europe and find that the average annual credit excess return on investment grade corporate bonds over the 1936-2014 period is 137 basis points with a Sharpe ratio of 0.37.

Over the more recent August 1988-December 2014 period they find that the average annual credit excess return for the aggregate high yield corporate bond index is 248 basis points with a Sharpe ratio of 0.26. They further find that this risk premium is not spanned by other commonly used risk factors (see also the next section).

In the example we presented above, the risk free interest rate was 0% making it relatively straightforward to measure the credit risk premium. However, when the risk-free interest rate is positive, then the discount of the bond price related to the time value of money needs to be appropriately separated from the discounting due to credit risk. For this separation, it is important to accurately match the duration of the corporate bond portfolio with that of the corresponding risk free bond portfolio (i.e., a matched government bond portfolio). Asvanunt and Richardson (2017) argue that if this matching is not done accurately, inadvertently the duration of the corporate bond portfolio could be lower than that of the government bond portfolio. In this case, the measurement of the credit risk premium is contaminated by the term premium, leading to an estimate of the credit risk premium close to 0 as in Fama and French (1993). Indeed, when using the same matching methodology as Fama and French (1993), Asvanunt and Richardson (2017) find a credit risk premium equal to 7 basis points.

7.2 Additional factors in government bond markets

For nominal government bonds, four main factors have been explored: betting-against-beta (BAB), carry, momentum, and value. An appealing feature of these factors is that the same factors also work in a wide variety of other asset classes, such as global equities, corporate bonds, commodities, currencies, and in some cases options. This implies that concerns about data mining, which have been discussed extensively in the recent asset pricing literature, see for instance Harvey, Liu, and Zhu (2016), are less of a concern for these factors.

We briefly discuss how the factors are constructed in turn.

BAB Frazzini and Pedersen (2014) define the BAB factor as a portfolio that is long low-beta and short high-beta countries. For government bonds, the betas are computed with respect to a GDP-weighted portfolio. The long and short ends of the portfolios are weighted such that the overall portfolio is market neutral.

Carry Koijen, Moskowitz, Pedersen, and Vrugt (2018) define carry as the return if market conditions do not change. They show in particular that carry, once applied to fixed income markets, equals the yield spread plus a roll-down component coming from the changing maturity of the bond. Empirically, most of the variation in the signal over time is due to the yield spread, which therefore mirrors the well known result that the yield spread forecasts future bond returns.

Momentum Asness, Moskowitz, and Pedersen (2013) use a standard definition of momentum across asset classes as the return during the last 12 months, while skipping the most recent month.

Value Of all factors, value is perhaps the most ambiguous signal to define as it generally corresponds to a measure of the price relative to fundamentals. Asness, Moskowitz, and Pedersen (2013) define value as the 5-year change in the yields of 10-year bonds. The use of this measure is motivated by the

equity literature, going back to De Bondt and Thaler (1985), who show that long-term returns are useful to identify growth and value firms.

Relation to alternative predictors in the bond market We already discussed the link between carry and the yield spread, a traditional predictor variable in bond markets. Brooks and Moskowitz (2017) also provide a further connection between the factors listed above and the Cochrane and Piazzesi (2005) and Cieslak and Povala (2015) predictor variables. Both are linear combinations of yields, potentially adjusted for long-term inflation expectations, and have been shown to predict U.S. nominal bond returns in the time series. Dahlquist and Hasseltoft (2013) extend this evidence for the Cochrane and Piazzesi factor to other countries. Given the evidence in Brooks and Moskowitz, we focus on these four factors that also summarize important dimensions of the risk-return tradeoff in other markets.

7.3 Additional factors in corporate bond markets

Similar to the earlier papers, Israel, Palhares, and Richardson (2018) show that carry, low-beta, momentum, and value are important drivers for the cross-section of corporate bond returns within the U.S., which is by far the largest corporate bond market globally as we show below. This suggests that the same factors also extend to corporate bond markets.

7.4 Currency returns and bond markets

The returns on the factors that we report are hedged against movements in the currency. It is however interesting to explore the interaction between signals from fixed income markets and exchange rate changes. The most natural connection that has been extensively studied is the currency carry trade, in which short-term interest rate differentials are related to currency returns. However, this connection weakens for longer-maturity bonds, see for instance Lustig, Stathopoulos, and Verdelhan (2018), implying that local term premia counter currency risk premia. While beyond the scope of this report, it would be of interest in future work to study the link between fixed income markets

and currency returns more broadly, in particular the interaction between currency risk premia, term premia, and credit risk premia.

7.5 Facts about fixed income factors and their relation to the benchmark

Data on all factors are available for nominal government bonds, and we focus on this segment as a result. Frazzini and Pedersen (2014) and Koijen, Moskowitz, Pedersen, and Vrugt (2018) provide additional results for corporate bond markets.

The sample for momentum and value starts in January 1983, for carry in November 1983, and for BAB in July 1989. For carry, momentum, and value, we have data until December 2017. For BAB, the sample ends earlier in March 2012. As all factors are in USD, we also compare the factors returns to the USD returns on the fund's actual and benchmark portfolio.

For the results below, we consider three potential samples: (i) the longest possible sample for each factor, (ii) the sample for which all series are available (January 1998 to March 2012), and (iii) the longest overlapping sample excluding BAB (January 1998 to December 2017).

Table 3: The table provides summary statistics of the fixed income factors and the fund's return and benchmark return.

		Factors				Fund	
	Sample	BAB	Carry	Momentum	Value	Actual portfolio	Benchmark
Mean		0.4%	2.8%	0.7%	0.6%	3.4%	3.2%
St.dev.	Longest	2.9%	6.4%	5.5%	4.8%	7.2%	6.9%
Sharpe Ratio		14.1%	44.1%	13.5%	12.6%	47.0%	46.5%
Mean		0.1%	2.8%	0.8%	0.1%	4.1%	3.8%
St.dev.	Overlapping	2.3%	4.4%	2.5%	3.0%	8.0%	7.6%
Sharpe Ratio		3.2%	62.7%	30.5%	1.7%	51.1%	50.1%
Mean			2.4%	0.5%	0.1%	3.4%	3.2%
St.dev.	Overlapping		4.0%	2.4%	2.7%	7.2%	6.9%
Sharpe Ratio	w/o BAB		58.6%	19.4%	4.0%	47.0%	46.5%

In Table 3, we report the annualized average returns, standard deviations, and Sharpe ratios of the four factors for the different sample periods. For comparison, we report the same moments of the fund's actual portfolio return and its benchmark. We convert both return series to excess returns using the U.S. 30-day T-bill rate as the factors are zero-cost (that is, long-short) portfolios as well.

As the factors are long-short portfolios, we can scale the standard deviation to an arbitrary target. The Sharpe ratio is therefore most meaningful in this comparison. For all factors and for all sample periods, we find that the Sharpe ratios are positive. However, there is significant variation across factors, with a strategy that sorts on carry (which is closely related to the yield spread) resulting in the highest Sharpe ratio.

In comparison to the fund's portfolio and benchmark returns, the Sharpe ratios may look modest, but this interpretation is misleading as the factors are designed to be neutral to the market. Therefore, combining factors and the benchmark may still improve the overall risk-return tradeoff. To test this directly, we regress the factors on the fund's actual portfolio and the benchmark returns.

Table 4: Factors and the fund's actual portfolio and benchmarking returns.

		Fund's actual portfolio return				Fund's benchmark return			
	Sample	BAB	Carry	Momentum	Value	BAB	Carry	Momentum	Value
Alpha		0.5%	2.4%	1.0%	-0.2%	0.5%	2.4%	1.0%	-0.2%
T-stat	Overlapping	0.82	2.00	1.54	-0.26	0.85	2.01	1.56	-0.28
Beta		-0.10	0.09	-0.06	0.06	-0.11	0.10	-0.07	0.07
IR		21.9%	54.6%	41.9%	-7.1%	22.7%	55.1%	42.3%	-7.6%
		BAB	Carry	Momentum	Value	BAB	Carry	Momentum	Value
Alpha			2.0%	0.6%	-0.1%		2.0%	0.6%	-0.1%
T-stat	Overlapping		2.21	1.17	-0.09		2.22	1.20	-0.11
Beta	w/o BAB		0.10	-0.05	0.05		0.11	-0.06	0.05
IR			50.8%	26.9%	-2.1%		51.0%	27.4%	-2.5%

The results are presented in Table 4. The left columns compare the factors to the fund's actual portfolio returns, while the right columns compare the factors to the fund's benchmark returns. In each case, we regress the factor returns on the fund's excess return, either the actual return or the

benchmark return. We report the alpha, its t-statistic, the beta, and the information ratio, that is, the ratio of the alpha to the standard deviation of the residual. The alpha and the information ratio are annualized.

Given small tracking error that we highlighted before, it does not come as a surprise that the results are very similar for the actual fund returns and the benchmark returns. While the alphas for value turn slightly negative, the alphas for BAB, carry, and momentum are positive. In the shorter sample, the alpha on BAB is 0.5% per year, on carry it is 2.4% per year, and for momentum it is 1.0%. These numbers deteriorate for both carry and momentum if we add the period from April 2012 to December 2017, which has been an unusual period in fixed income markets with a binding zero lowerbound and large-scale asset purchases in many countries. We will return to this in the conclusions. The alphas are significant for carry in both sample periods, while the alphas for BAB and momentum are too small and noisy to be statistically significant.

If we look at the information ratio, then we notice that they are substantial. For the overlapping sample period, for instance, the information ratio of BAB is around 20%, for carry around 55%, and for momentum around 40%. While these numbers also deteriorate if we expand the sample to include the last five years of data, the information ratios remain substantial. These factors provide some guidance in terms of directions in which to tilt the portfolio to improve the overall risk-return tradeoff.

8 Are credits spanned by government bonds and equities?

8.1 Theoretical arguments

In this section, we explore to what extent corporate bond returns are spanned by returns on equity and returns on risk-free debt. One may argue that if this is the case, then corporate bonds can simply be replaced by the appropriate portfolio of stocks and risk-free bonds, without adjusting the risk profile of the Fund.

Before discussing spanning in more detail, we note that it has not been explicitly stated that the recent change in the stock-bond allocation to 70%-30% would go together with simultaneously de-risking the fixed-income portfolio. That is, it has not been stated that part of the increased allocation towards equity was to reduce the fraction invested in corporate bonds. The question addressed in this section is therefore if the 30% currently invested in corporate bonds (of the 30% of the overall portfolio invested in fixed income, and the 9% invested in corporate bonds of the overall portfolio) can be replaced by an appropriate combination of stocks and risk-free bonds.

There exist theoretical reasons for why this spanning of corporate bond returns could be achieved. In particular, Merton (1974) shows that equity can be viewed as a call option on the underlying value of the firm, while corporate debt can be viewed as a portfolio of risk-free debt plus writing a put option on the underlying value of the firm. In addition, Black and Scholes (1973) have shown that dynamic spanning arguments can be used to price call and put options. That is, one can construct dynamic portfolios of riskfree debt and the underlying asset to replicate the payoffs of put and call options. Combining these arguments then gives a motivation for why corporate debt could be dynamically spanned by some combination of equity and risk-free debt.

There are several reasons for why achieving this spanning could be challenging in practice. First, the arguments rely on dynamic spanning implying that the relative portfolio weights vary over time. When managing the Fund's portfolio, this dynamic trading can be costly depending on the level of transaction costs. Second, if government bond markets, equity markets and corporate bond markets are segmented, or otherwise incomplete, spanning will fail. For instance, a large demand for safe assets can drive up the price of government debt and thus down the rate of return on risk-free debt, creating a positive wedge between corporate bond returns observed in the market and the returns that dynamic spanning would imply. We refer to Krishnamurthy and Vissing-Jorgensen (2012) and Jiang, Krishnamurthy, and Lustig (2018) for evidence on this convenience yield and its broader implications.

Lastly, we note that even if corporate bonds are (close to) being spanned by equities and risk-free bonds, it may be attractive to keep corporate bonds in the benchmark as it adds capacity and it may

lower trading costs as a result, which may be a relevant consideration given the Fund's size.

8.2 Empirical results

As an exploration of how well spanning works empirically, we take the U.S. as an example, which is the largest corporate bond market in our sample. We use the total U.S. aggregate corporate bond returns between January 2001 and August 2018 and regress these monthly returns on the return on the U.S. aggregate stock market (using S&P500 returns leads to very similar results), as well as on returns on risk free debt. For the latter, we use the zero coupon yield data provided by Jonathan Wright, which allows us to match the maturity of the risk free debt to the duration of the corporate bond returns. The time series of this duration is plotted in Figure 8.

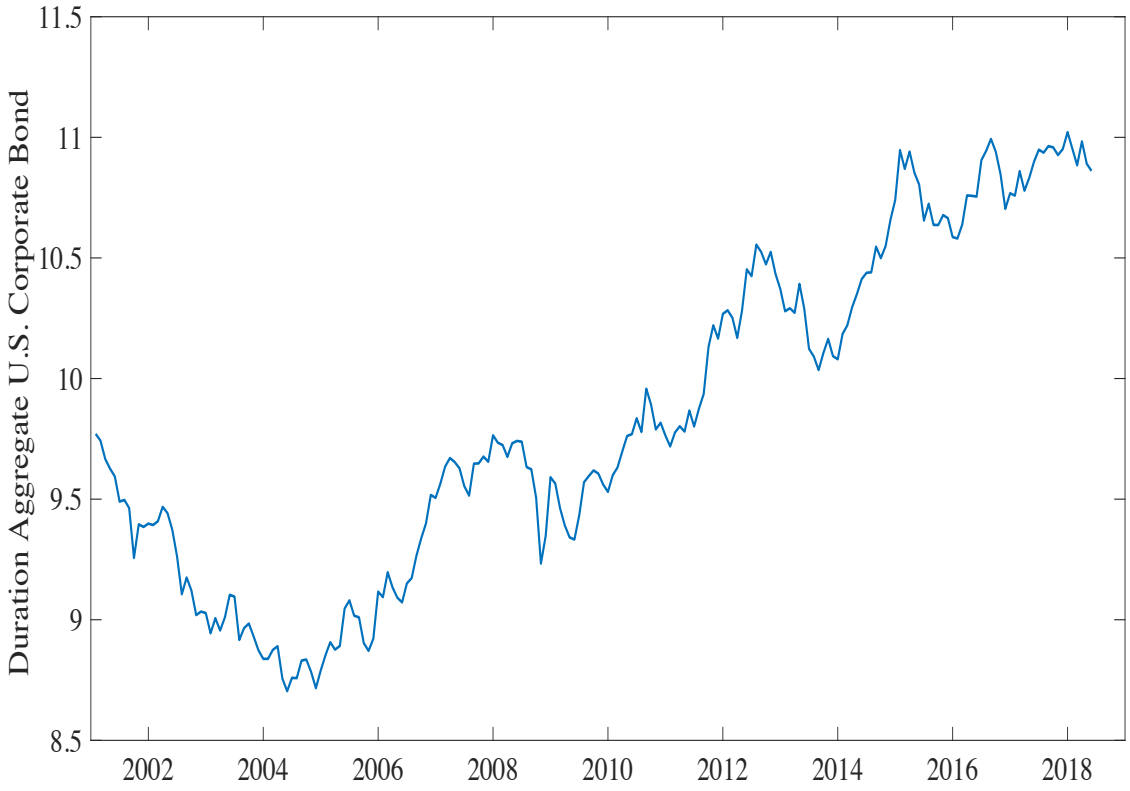

Figure 8: Duration Aggregate U.S. Corporate Bond Index.

To allow for dynamic adjustments in the portfolio weights of the spanning portfolio, we estimate the regression on an annual rolling basis. That is, each regression only uses 12 monthly returns. The results are summarized in Figure 9. The metric to assess spanning is the R-squared. After all, an asset is spanned by another portfolio of assets if the payoffs are identical, not just on average, but state-by-state. While it is impossible to test spanning for states that have not realized yet, which is the Peso problem we have highlighted before, we can test whether a replicating portfolio existed for the states that did realize historically. By allowing the weights to change over time, we are giving the replicating portfolio of equity and risk-free bonds the best chance to replicate the corporate bond index.

The figure shows that the adjusted R-squared of the spanning regression varies substantially over time. There are periods where near perfect spanning is achieved. There are other periods, however, where the explained variation is low. The average of this rolling adjusted R-squared over the sample period is 57% (this compares to an R-squared of 45% when we only run the regression once over the full sample period, thereby not allowing for time varying weights).

In line with earlier literature, our analysis suggest that corporate bond returns have quite a large amount of unspanned variation. Replacing corporate bonds by a replicating portfolio of government bonds and equity can thus leads to tracking error. Our results further indicate that the amount of unspanned variation varies over time. This implies that the return difference between a replicating portfolio and corporate bonds is time-varying and can at times potentially be large. Historically, as can be seen from Figure 9, such periods have partly coincided with periods of financial market turmoil.

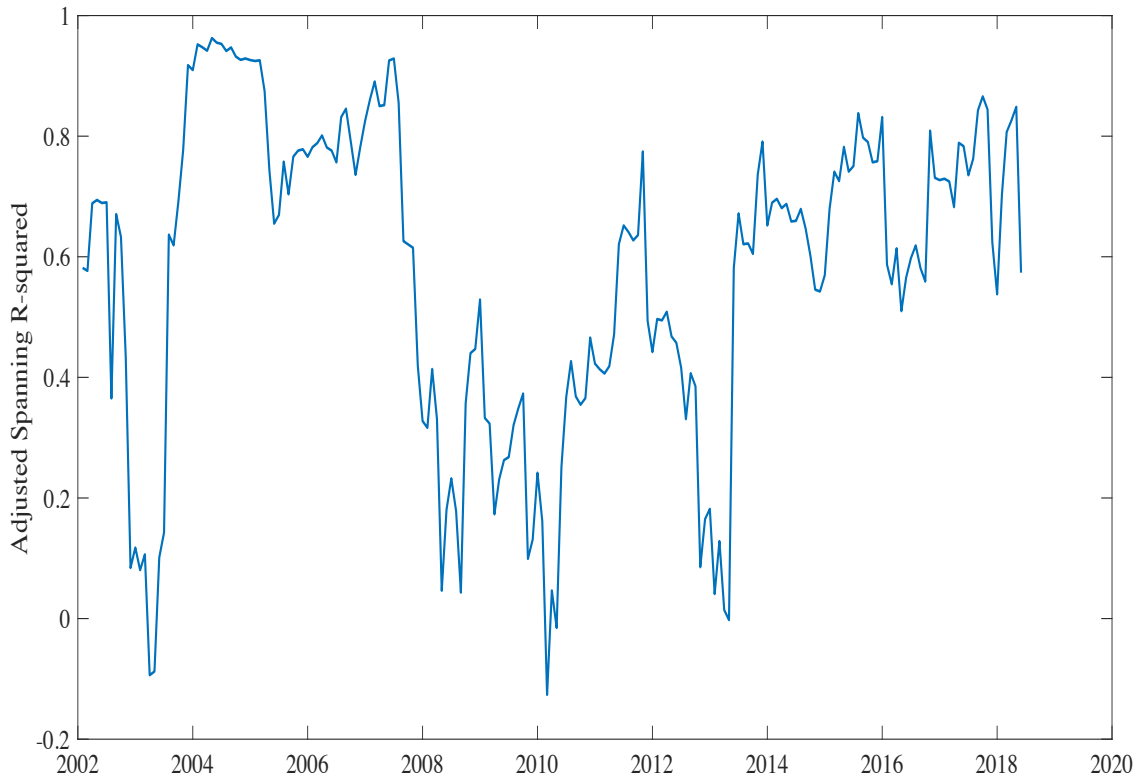

Figure 9: Adjusted R-squared of the spanning test.

9 Benchmarking fixed-income benchmarks

9.1 The size and composition of fixed income markets

Figure 10 summarizes the size of different fixed income segments from January 2001 to the end of our sample as covered by the Bloomberg data.

The nominal government bond market in developed markets grows to $25.3 trillion by December 2017 and to $4.4 trillion in emerging markets. The inflation-linked bond market grows to $2.9 trillion, other government securities to $6.0 trillion, the corporate bond market to $9.0 trillion, and the covered bond market to $0.9 trillion in terms of market values.

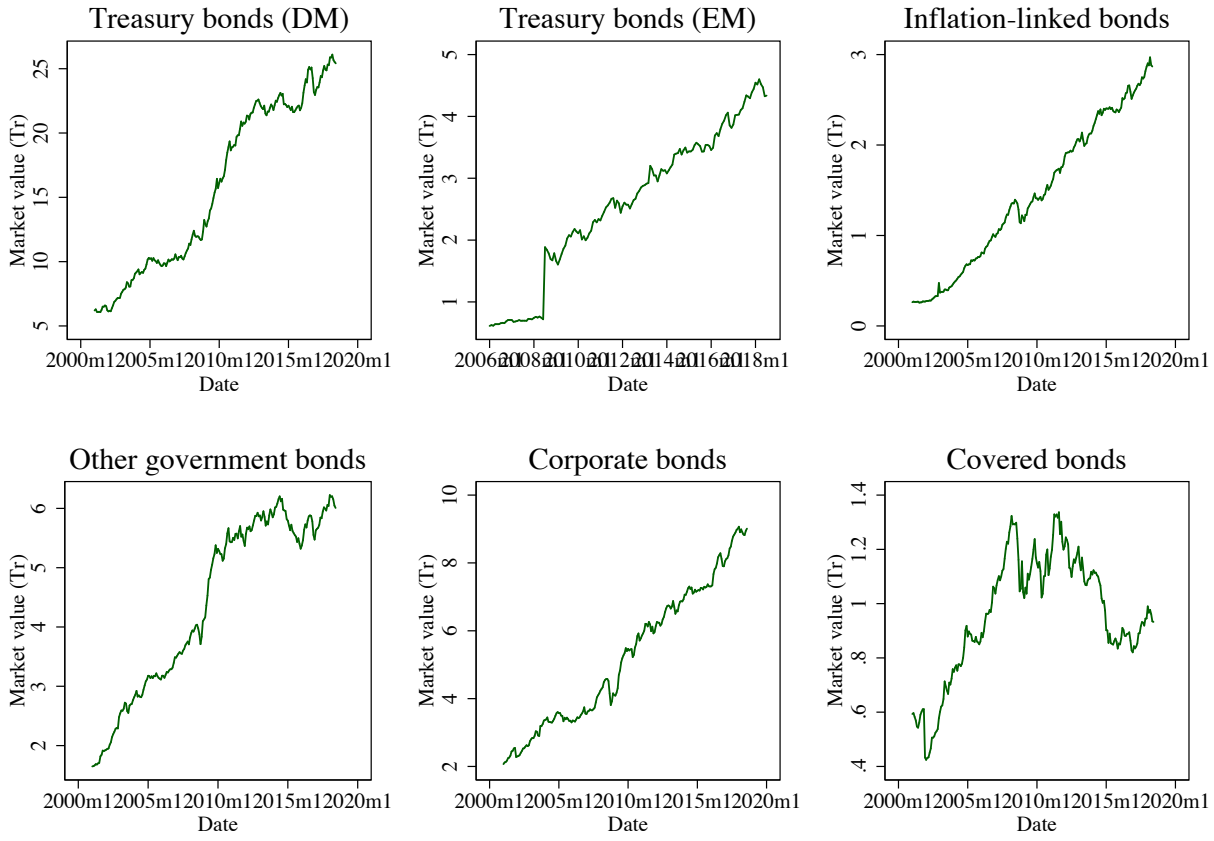

Figure 10: Size of fixed income segments in market values

To separate changes in market values into issuances and price effects (that is, yield changes), we repeat the same figures in Figure 11 but now in terms of face values. The broad patterns are similar with the expansion of the nominal government bond segment in Figure 10, in particular during the financial crisis and the stagnation of the covered bond market.

The decline in the covered bond market may be related to banks reducing the size of their balance sheets or banks may be able to access other sources of funding more cheaply. In Table 5 we report the log growth rate of the covered bond market across 17 countries in our sample from June 2011 to June 2018. While some of the countries severely affected by the financial crisis and the sovereign debt crisis experience the largest contractions, like Ireland, Spain, and Portugal, the Italian covered bond market is an exception as it grows by 7% during the same period. This episode illustrates though that due to changes in issuance behavior across segments, the risks to which a market-weighted portfolio is

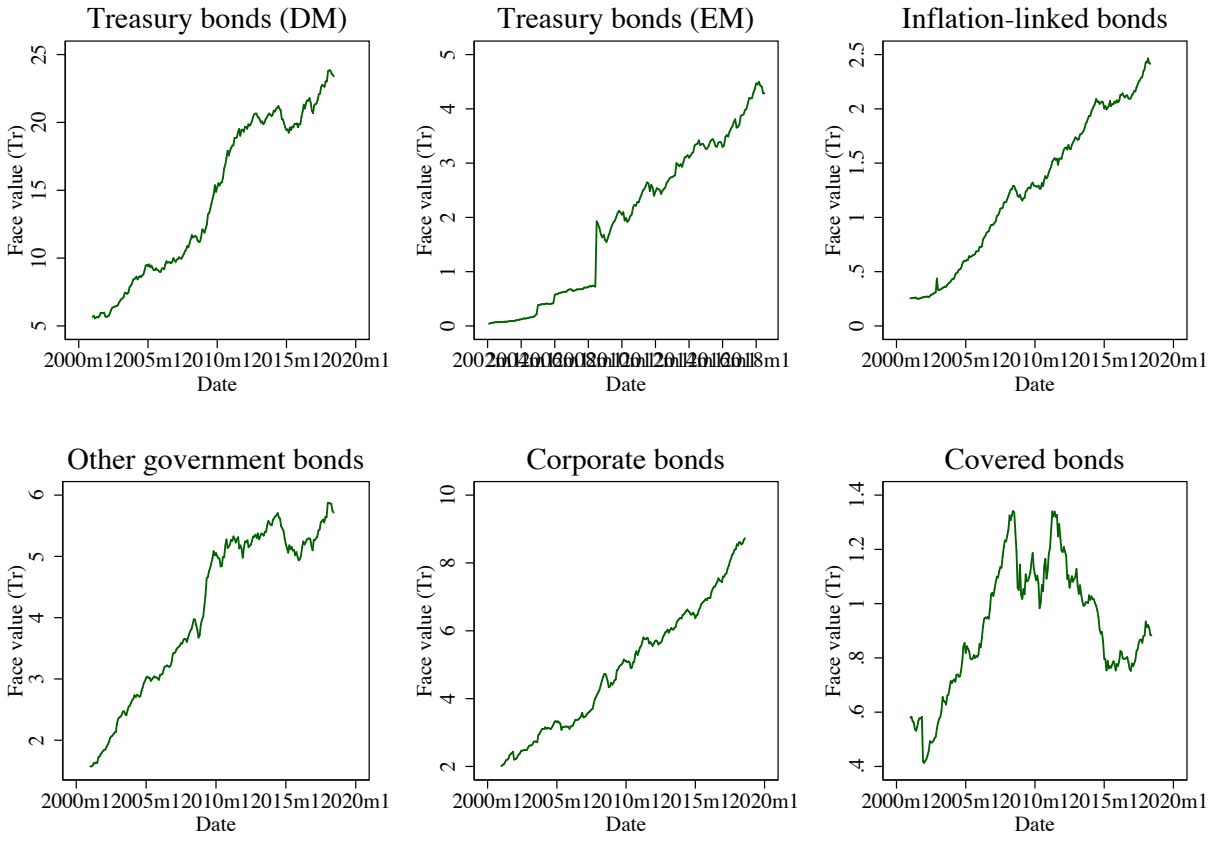

Figure 11: Size of fixed income segments in face values

exposed may changes significantly over time.

Next, we turn to the issue of concentration, which is often a reason to deviate from strict market weights in the context of fixed income markets. We report the square root of the Herfindahl as our measure of concentration, $HHI_t = \sqrt{\sum_i (MC_{it}/MC_t)^2}$, where MC_{it} is the market cap of country i and MC_t is the total market cap of the segment. We report the results in Figure 12.

We should preface these results once again by the observation that these results are based on the bonds covered by the Bloomberg Barclays index. Government bonds are the least concentrated segment, followed by covered bonds by the end of the sample. Corporate bonds and inflation-linked bonds are both highly concentrated.

Table 5: Log growth in the covered bond market between June 2011 and June 2018.

Country	Growth	Country	Growth
Ireland	-132%	Sweden	-3%
Spain	-120%	Netherlands	7%
Luxembourg	-107%	Italy	7%
Portugal	-96%	Denmark	12%
UK	-70%	Austria	47%
Switzerland	-55%	Norway	49%
Germany	-55%	Finland	60%
France	-32%	Canada	213%

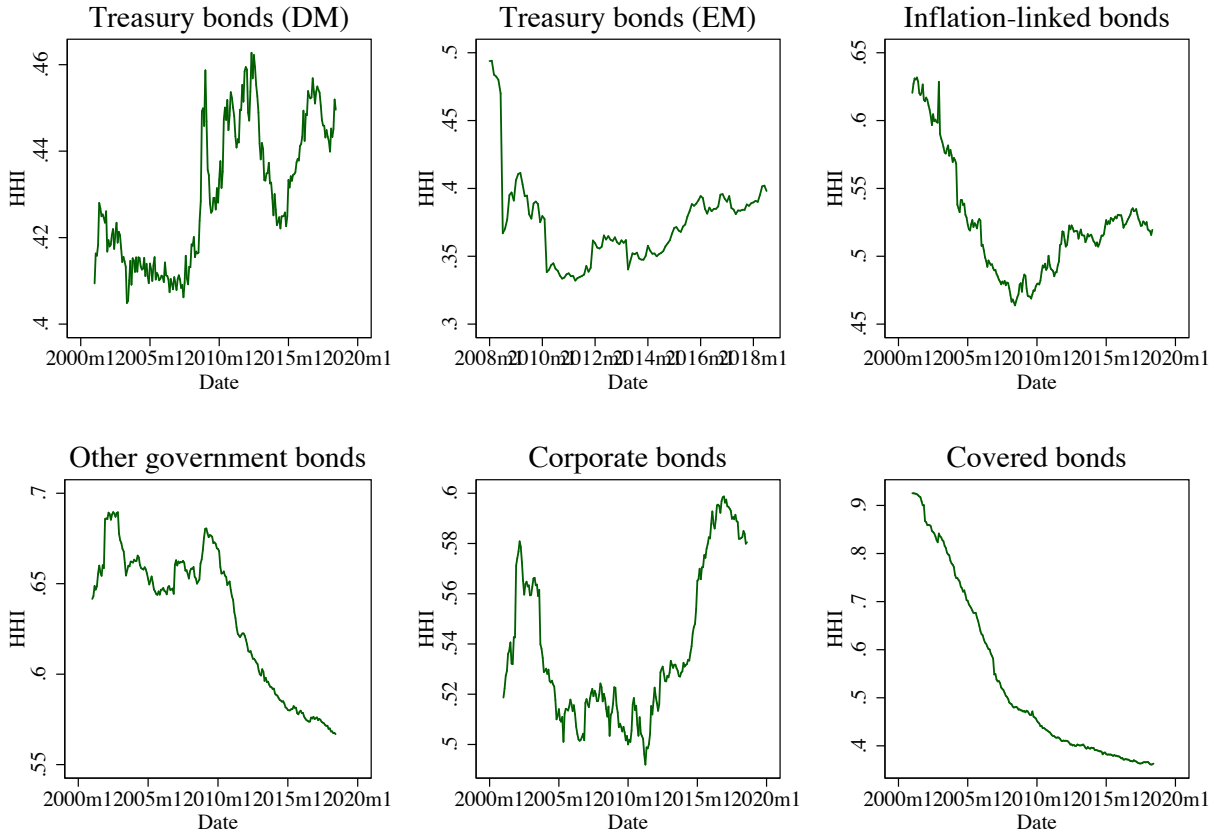

Figure 12: Concentration of fixed income segments.

9.2 Comparing GDP and market weights

Before computing the returns on various benchmarking strategies, we highlight the difference between market weights and GDP weights for government bonds. We use the final month of our sample,

May 2018. In Figure 14, we plot the market weights versus GDP weights in level (left panel) and logarithms (right panel).

If we look at the right panel first, then there is obviously a strong correlation between market and GDP weights. A simple regression of log GDP weights on log market weights in May 2018 provides a slope of 0.70 (t-stat = 11.0) and an R-squared of 83%.

However, the right panel understates the impact of using GDP weights if we look at the left panel, which displays the weights in levels. The weight that Japan receives in market weights, for instance, is much larger than the GDP weight. A regression in levels results in a slope of 0.87 (t-stat = 7.0) and an R-squared of 67%.

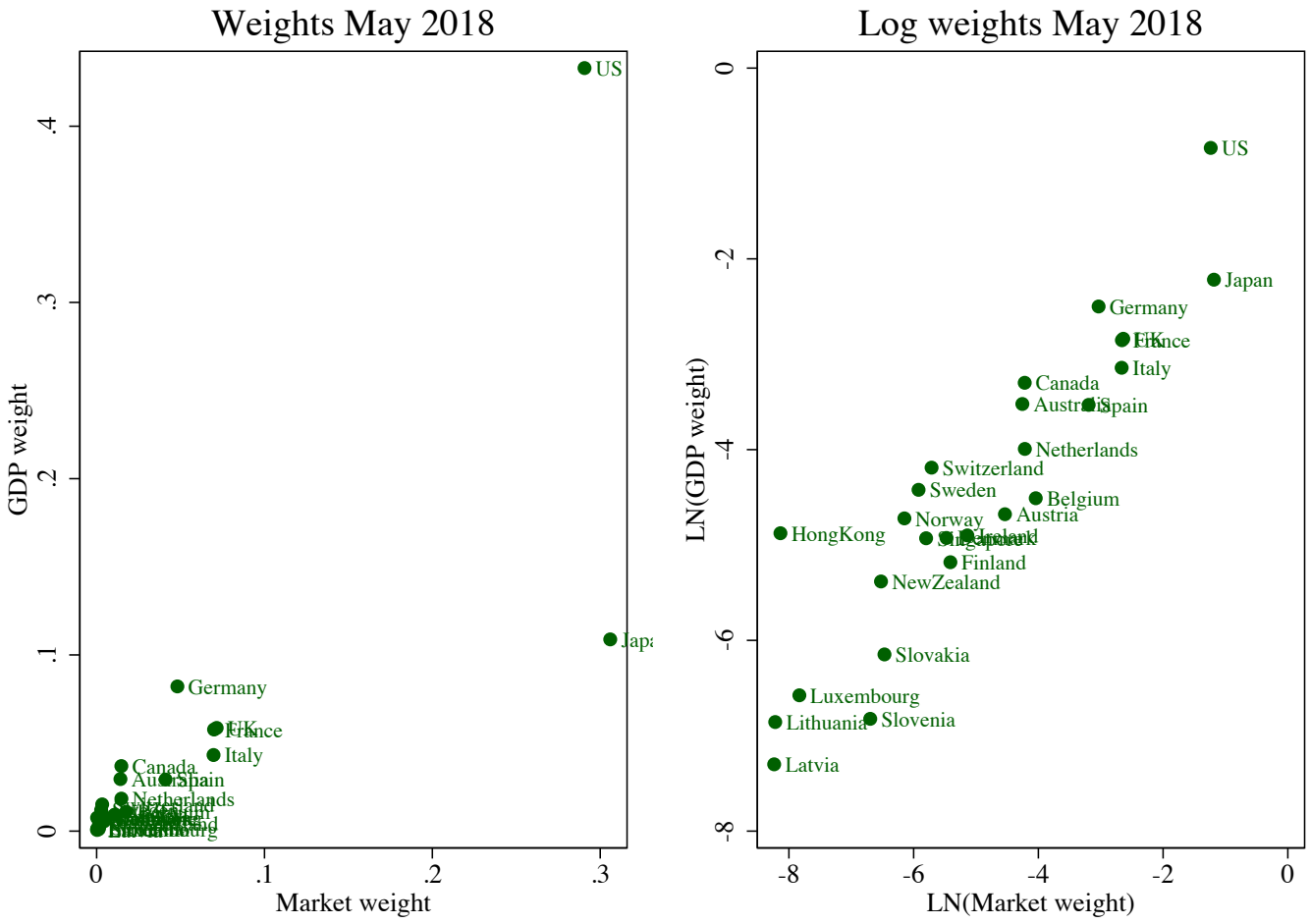

Figure 13: GDP versus market weights in May 2018 in levels (left panel) and logarithms (right panel).

9.3 Return properties of China and India

Using GDP weights also raises a new concern. While perhaps effective given the current distribution of GDP, there is no reason why this cannot change in the future. Bloomberg announced on March 23 2018 that it will add Chinese RMB-denominated government- and policy-bank securities to the index. These securities will be gradually added over a 20-month period starting April 2019. If one were to use GDP weights, China would be the second-largest holding, more than twice as large as the position in Japanese government bonds. However, in terms of market value, China would only rank sixth in December 2017 with a market value of debt that is more than five times smaller than Japan's. Another rapidly-developing economy, India, is ranking ninth in terms of market capitalization.

For completeness, we therefore document the return and yield properties of China and India, in comparison with the United States. We do note, however, that the sample only starts in July 2008 for both China and India.

In Figure 14, we plot the yield dynamics of all three countries, which shows that the yield of India is substantially higher than the one of China. This is consistent with the higher estimated inflation rate in India compared to China and the US. The World Bank estimates the inflation rates in China, India, and the US to be 2.0%, 4.9%, and 1.3% in 2016, respectively.

Another aspect that is worth noting is that the yields in both China and India responded more sharply to the "taper tantrum" in the summer of 2013, with yields in India increasing by more than 2%. The time-series correlation between the yields in China and India is 57%, between China and the US it is 10%, and between India and the US it is -29%. In terms of returns (which are closely related to yield changes), the same correlations change to 26%, 18%, and 10%. This suggest that there is a large amount of idiosyncratic yield variation in both China and India.

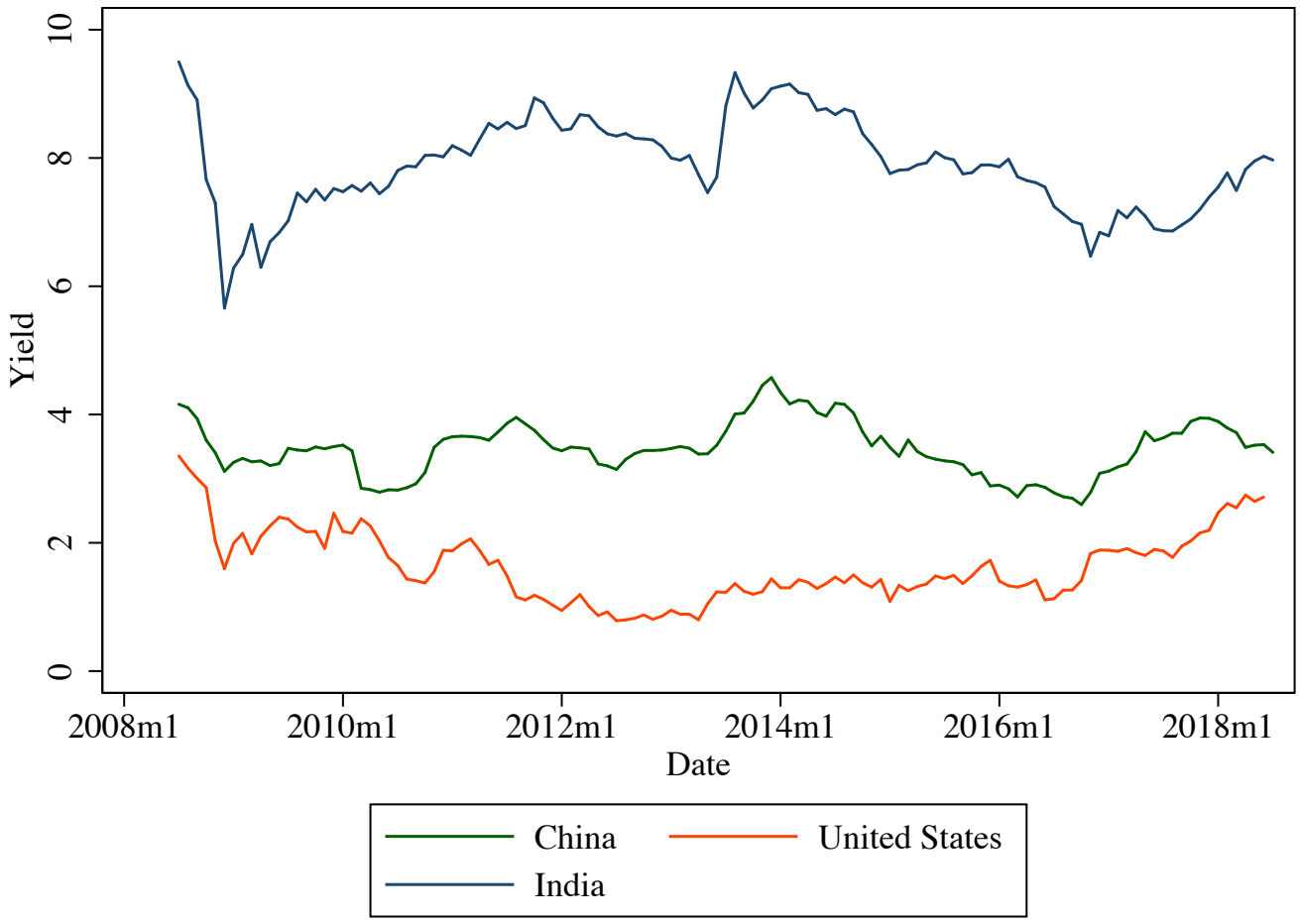

Figure 14: Dynamics of yields in China, India, and the United States.

9.4 Benchmarking benchmarks

In previous sections we have discussed a range of theoretical and practical issues and tradeoffs when choosing benchmarks. We have come to the conclusion that market weights should be the starting point for any benchmark design. We have also discussed situations and have presented arguments for why deviations from market weights could be warranted, see Section 6.1.

In this section, we compute several candidate benchmarks and compare their empirical properties. Recall that in the construction of the benchmark we consider the following categories and subcategories:

- Government

    1. Nominal government bonds of developed economies
    2. Nominal government bonds of emerging economies
    3. Inflation-linked government bonds
    4. Other government-related bonds

- Non-government
    1. Corporate bonds
    2. Covered bonds

Based on these categories, we consider the following four benchmarks:

1. A benchmark with a 70/30 split between government and non-government fixed income products and market-value weights within the government and non-government categories.
2. A benchmark with a 70/30 split between government and non-government fixed income products and market-value weights within the government and non-government categories excluding inflation-linked bonds.
3. A benchmark with a 70/30 split between government and non-government fixed income products and market-value weights within the government and non-government categories excluding emerging market government bonds.
4. A benchmark with a 70/30 split between government and non-government fixed income products and market-value weights within the non-government categories. Within the developed government bond, the emerging market government bond, as well as within the inflation-linked categories we apply GDP country weights. To determine the relative proportion of these three subcategories, we use market value weights.

Table 6: The table reports mean returns and standard deviations of 4 potential benchmarks using data between January 2001 and June 2018. The table reports the mean and standard deviation of monthly returns as well as annualized versions of those numbers (i.e. $(1+\mu)^{12}$ for the mean and multiplying by $\sqrt{(12)}$ for the standard deviation.)

Potential Benchmarks	Monthly mean (%)	Monthly st.dev. (%)	Monthly kurt.	Ann. mean (%)	Ann. st.dev. (%)
1: MW within	0.404	1.81	3.71	5.0	6.28
2: MW, no inflation linked	0.401	1.81	3.65	4.9	6.26
3: MW, no EM govt bonds	0.401	1.82	3.70	4.9	6.30
4: GDP weights	0.410	1.76	4.06	5.0	6.11

The summary statistics of these 4 benchmarks are reported in Table 6.

The table shows that the benchmark returns have highly similar characteristics in terms of mean returns and standard deviations. Over this sample period, the mean return of Portfolio 4 (using GDP weights) is somewhat higher than Portfolio 1 (70/30 market weighted portfolio), and the difference equals 8 basis points per year. The standard deviation of Portfolio 4 is somewhat lower (by 0.17% per year). On the other hand the kurtosis of Portfolio 4 is substantially higher. This can also be seen from the minimum monthly return that has occurred during this sample period, which equals -5.71% for Portfolio 4 and -4.44% for Portfolio 1. In the next section we will elaborate on stress testing scenarios in which worst case scenarios are discussed in more detail.

Next, we compute in Table 7 the correlations between the monthly returns on the four benchmarks. We find that the portfolios have extremely high correlations with each other. All correlations are equal to or above 0.988. Overall, we can conclude that the differences in the summary statistics of

Table 7: Return Correlation Matrix

Potential Benchmarks	Port 1	Port 2	Port 3	Port 4
Port 1: 70/30, Market weights within	1.000			
Port 2: 70/30, Market weights, no inflation linked	1.000	1.000		
Port 3: 70/30, Market weights, no emerging market	0.999	0.999	1.000	
Port 4: 70/30, GDP weights	0.991	0.989	0.988	1.000

the portfolios are small over this sample period. Whatever differences there are, should not receive an overly large weight in the choice of benchmark as they may be specific for this particular sample. We do note, however, adding the additional categories may increase capacity and hence lower trading

costs.

10 Concentration risk and adjustment factors

In this section, we provide a simple approach to measure concentration risk, defined as impact on the overall portfolio of idiosyncratic yield changes. One can compute the idiosyncratic yield changes using historical data, as we do in this section, and complement it with additional scenarios as we discuss in the next section.

We illustrate our approach using the government bond portfolio in developed markets, but readily extends to other fixed-income segments. The concentration risk measures can subsequently be used to construct adjustment factors in a systematic way, and we discuss a particular way of doing so that results in weights that are as close as possible (as we define precisely below) to the current benchmark based on GDP weights, which may be helpful in transitioning from the current GDP weights to a market-weighted index with concentration limits.

10.1 Measuring concentration risk

The composition of the government bond part of the fixed income benchmark for the GPFG is based on GDP weights. When the current benchmark was adopted in 2012, both the Ministry of Finance and Norges Bank emphasized that GDP weights may, as compared to market weights, to a larger degree reflect the ability of governments to repay debt.[11] For example, if the market capitalization of a particular country's government bond portfolio is large and, in addition, the duration of the country's bonds is large, then upward shifts in that country's yield curve may cause significant losses in the overall portfolio of the fund. While it should be noted that overexposure is only defined relative to some notion of normal exposure, the worry seems to stem mainly from situations where the weights

[11] The Bank has in its advice to the Ministry, as referred to in our mandate, stated that the benchmark should continue to use GDP weights for the government bond markets.

implied by the market capitalization differs substantially from the weights implied by GDP. Overall, we can conclude that substantial losses coming from one particular region or country is considered to be undesirable by the stakeholders in the Fund.

To address the concern described above, we propose a scenario analysis approach to measure idiosyncratic tail risk exposures across countries/regions. In combination with defined risk limits, these exposures can then be employed to create adjustment factors for the portfolio weights of individual countries, if so desired. That is, the exposure measures can be used to scale down the portfolio weights of countries with high exposures and to increase the portfolio weights of countries with low exposures.

To implement this scenario analysis, we need (1) the market weights of each country, (2) the duration of each country's bond portfolio, and (3) scenarios for idiosyncratic yield curve shifts.

The first two items are readily available. The third item is somewhat more challenging. While in principle we could use the time series of historical annual yield changes in each *individual* country to generate scenarios for that country, the limited available time-series information may not be sufficient to properly identify tail events in each country. To address this concern, we follow the literature on rare disasters (Nakamura, Steinsson, Barro, and Ursua, 2013) and use the whole historical panel of countries to better understand the importance of tail events. That is, we use cross-sectional information across countries to generate tail scenarios within each country.

Let $y_{t,n}$ denote the government bond yield for country $n \in \{1, \ldots, N\}$ at time $t \in \{1, \ldots, T\}$ measured in months. We then compute for each time t a cross-section of yield changes, taking out a time fixed effect which adjusts for systematic movements in yields across countries:

$$\varepsilon_{t,n} = y_{t,n} - y_{t-12,n} - \frac{1}{N} \sum_{n=1}^{N} \left(y_{t,n} - y_{t-12} \right). \tag{2}$$

We then stack all observations across time and across countries and compute the 95th and 99th percentiles of idiosyncratic yield changes. Figure 15 below plots the distribution of $\varepsilon_{t,n}$. The 95th percentile of the distribution is at 0.78% and the 99th percentile is 1.63%. We will use the latter

number in our computations below.

We note that it is also possible to construct idiosyncratic shocks per segment, or to condition on credit rating. The reason that we prefer not to do this is that we are trying to assess the risk over long horizons (that is, the horizon of the bond). While a country may have a high rating now, this may quickly change in times of crises, as illustrated by the Italian yield dynamics before. We therefore take a more agnostic approach and compute the idiosyncratic risk across all developed markets, without any conditioning information.

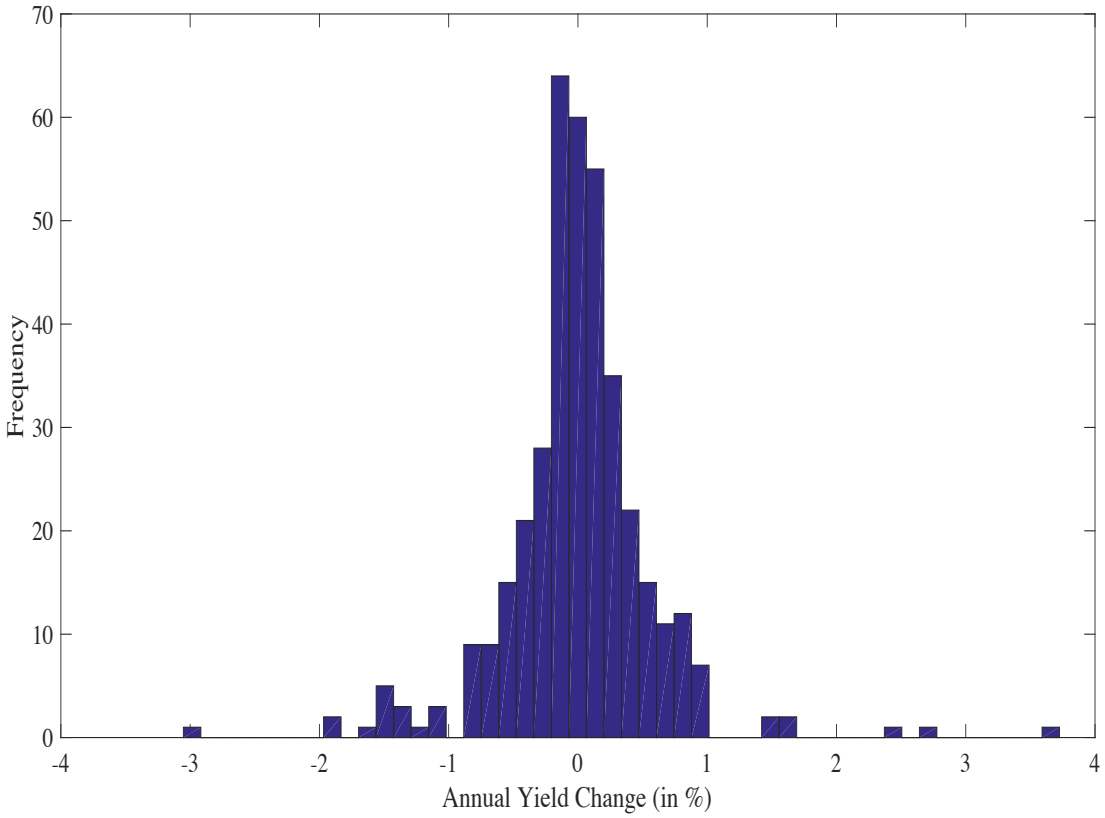

Figure 15: Duration Aggregate U.S. Corporate Bond Index.

Starting from the market portfolio, the total return impact of an extreme yield change on the Fund's bond portfolio performance, which is our proposed idiosyncratic tail risk exposure measure $\xi_{t,n,}$

can then be computed as:

$$\xi_{t,n} = -D_{t,n} \times w^M_{t,n} \times 1.63\%, \tag{3}$$

where $D_{t,n}$ is the modified duration of country n's government bond portfolio and the weight $w^M_{t,n}$ represents the market value of country n's government bond portfolio as a fraction of the total value of government bonds across all countries.

Next, we report the exposure measure for each country using bond duration and market weight from December 2017. The results are reported in Table 8, where we list the countries sorted by their exposure. That is, the country with the largest exposure is listed first. The table shows that the Fund's tail risk exposure to Japan would be the largest, followed by the US and the UK.

The idiosyncratic tail risk exposure numbers can now be combined with a policy variable that states that the maximum tail exposure for any particular country should not be larger than $\bar{\xi}$. The portfolio weight can then be scaled back by a concentration adjustment factor to achieve this objective.

The reason for why we use the same extreme yield change scenario for each country is that it is very hard to predict where in financial markets extreme movements will occur. If on the other hand one has strong views that these extreme scenarios are a function of country characteristics such as the credit rating or other measures of fiscal strength, then the relevant quantile of the $\varepsilon_{t,n}$ distribution can be modeled as a function of such characteristics, and the exposure measure ξ can be adjusted accordingly.

10.2 Adjustment factors

One can use the risk measures of the previous section to construct adjustment factors. More precisely, we consider a final benchmark

$$w_{t,n} = \frac{w^M_{t,n} A_{t,n}}{\sum_n w^M_{t,n} A_{t,n}}, \tag{4}$$

Table 8: Idiosyncratic Tail Risk Exposure ξ

Country	Modified Duration	Market Weight	Exposure measure (%)
Japan	9.55	0.30	-4.615
U.S.	6.23	0.29	-2.991
U.K.	11.79	0.07	-1.406
France	8.13	0.07	-0.930
Italy	6.80	0.07	-0.755
Germany	7.41	0.05	-0.601
Spain	7.17	0.04	-0.468
Belgium	9.34	0.02	-0.270
The Netherlands	8.01	0.02	-0.201
Canada	6.58	0.02	-0.162
Austria	8.89	0.01	-0.158
Australia	6.44	0.01	-0.157
Switzerland	11.74	0.00	-0.066
Ireland	6.82	0.01	-0.062
Denmark	8.97	0.00	-0.060
Finland	7.03	0.00	-0.050
Sweden	6.12	0.00	-0.033
Singapore	6.80	0.00	-0.032
Slovakia	7.92	0.00	-0.020
Slovenia	9.20	0.00	-0.018
Czech	5.50	0.00	-0.017
Norway	4.63	0.00	-0.015
New Zealand	4.86	0.00	-0.014
Lithuania	11.77	0.00	-0.005
Latvia	8.85	0.00	-0.004
Luxembourg	5.63	0.00	-0.004
Hong Kong	2.78	0.00	-0.001

where $A_{t,n}$ is the adjustment factor. We consider $A_{t,n} = f(\xi_{t,n})$. One option would be to limit the exposure at an absolute level, $A_{t,n} = \min\{\xi_{t,n}, \xi^\star\}$, where $\xi^\star$ is a policy variable.

We prefer a solution that is smooth in the risk factors. We therefore parameterize $A_{t,n} = \left(w^M_{t,n}\right)^{\alpha_1} \xi^{\alpha_2}_{t,n}$, where the coefficients are policy variables. To provide a possible way to calibrate these parameters, we choose them so that the resulting weights are close to the current GDP weights. In particular, we estimate a regression

$$\ln w^{GDP}_{t,n} = \alpha_0 + \alpha_1 \ln w^M_{t,n} + \alpha_2 \ln \xi_{t,n} + \epsilon_{t,n},$$

take the fitted value, and exponentiate it.

Figure 16 illustrates the results. Our adjustment procedure significantly reduces the weight on Japan, just as GDP weights do. However, it leaves the weight in the US unchanged compared to market weights (yet significantly lower compared to GDP weights).

10.3 Harvesting risk premia outside of the benchmark

Section 7 summarizes dynamic factor strategies based on carry, low-beta, momentum, and value that in some cases deliver significant improvements in the risk-return tradeoff beyond a market portfolio. While we consider these important strategies to consider as part of the overall investment strategy of the fund, these strategies are probably best implemented as part of the active mandate rather than the benchmark. The primary risks to be harvested via the benchmark are term and credit risks across the different fixed income segments.

To measure the risk of active strategies, we recommend complementing traditional risk measures with the same stress scenarios as discussed in the previous and the next sections.

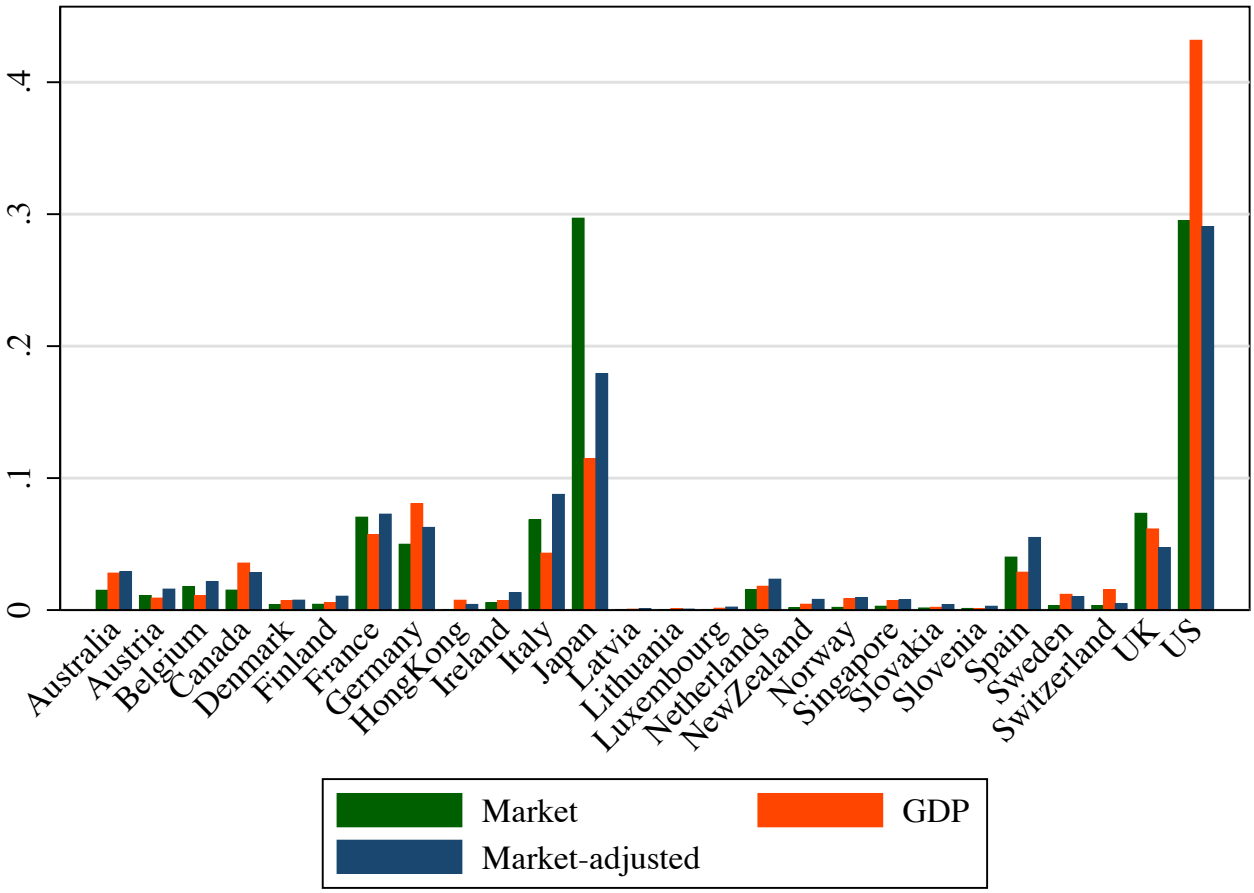

Figure 16: Benchmark weights in December 2017 for market weights, GDP weights, and adjusted market weights for the government bond portfolio in developed markets.

11 Stress testing

Several important secular trends have affected fixed income markets. First, interest rates in many areas in the world have experienced a long downward trend. In the U.S., bond yields have been on a downward trend for about four decades.

Second, and relatedly, after the recent financial crisis, several areas in the world had (and have) nominal interest rates that were stuck at (or even a bit below) the so-called zero lower bound. Interest rates at the zero lower bound have been the norm in Japan for two decades, but have been a relatively new experience for European countries and the U.S. The disadvantage of the presence of such secular trends is that the statistical moments derived from 50 years of historical data (means, variances and

correlations) may not be sufficiently representative for risk management purposes going forward.

Furthermore, as we have argued before, investment grade bond investments are the combination of risk free bond investing plus the writing of an out-of-the-money put option. By writing put options, one can earn the premium that the buyer of the option is willing to pay (the credit premium). This premium enhances the return of the bond investor for as long as the option stays far out-of-the-money (i.e. the issuer's financial condition remains healthy).

However, in rare occasions, the credit situation of the issuer deteriorates, leading to outsized negative returns. Such infrequent, very negative returns (often called a Peso problem) are an important challenge for traditional risk measures (such as tracking error volatility), as the available data may not feature the Peso event.

Because of this combination of the existing secular trends as well as the Peso problem inherent in risky bonds, we think it important to consider stress test scenarios for the fund's portfolio overall, and the fixed income portion in particular. By studying historical debt crisis in particular regions and applying them to other regions, insights can be obtained as to what the impact of such crisis would have on the overall performance of the fund as we have illustrated in the previous section. This risk measure can be combined with traditional risk measures such as tracking error. In addition to the stress scenarios that can be constructed using historical data, it may be desirable to add additional scenarios based on market dynamics.

We note that the same scenarios may not only be useful to determine the adjustment factors, but can also be used to assess the risk of active strategies, such as the factor strategies that we have discussed before.

The Ministry could decide to make the stress scenarios, and the Fund's performance during the stress scenarios, publicly available, which may by helpful in communicating the Fund's riskiness.

12 Conclusions

In this report, we have evaluated important considerations in the design of a global fixed income benchmark. We have used as a guiding principle that the proposed benchmarking guidelines are widely applicable as opposed to them being merely appropriate to current economic conditions. In our report, we have used the assumption that the 70/30 split between government and non-government securities is sufficient to maintain the overall liquidity of the benchmark.

Based on our analysis, we have three main conclusions and recommendations:

1. We recommend a benchmark that is broadly diversified across geographies and across fixed income segments to minimize risk and maximize investment capacity. In particular in fixed income markets, volatility and correlations can quickly change and, most importantly, in ways that are hard to predict based on historical data alone. The example of German and Italian yields provides a simple example. Hence, in the spirit of the fund's objective to be broadly diversified and to be exposed to the main risk factors in financial markets via the benchmark, we argue that a portfolio that covers a wide set of countries and the main fixed income segments (treasuries, inflation-linked bonds, corporate bonds, and covered bonds) not only diversifies risks during normal times, but also across changes in market regimes. In summary, a natural starting point for the benchmark is a portfolio that mimics the market portfolio. We have discussed particular reasons to deviate from market weights, such as market segmentation, but given that the research on these questions is relatively recent, we currently view such deviations as an active investment management decision.

2. We have documented the high concentration in fixed income markets in the different segments. It is worth noting that different countries play a prominent role in different fixed income segments. This high degree of concentration may expose the fund to idiosyncratic movements in a country or fixed income segment. Historically, the fund has addressed this concentration concern by using GDP weights in the government bond portfolio. While this measure has perhaps been

effective given the past and current distribution of GDP, there is no reason why this cannot change in the future. For instance, according to the International Monetary Fund (IMF) in 2017, in the U.S., nominal GDP is $19 trillion, it is $12 trillion in China, and $5 trillion in Japan. Bloomberg announced in March of 2018 that China will be included in the Bloomberg Barclays Global Aggregate Index as of next year, which would make it the second-largest position in the benchmark without further adjustment.

3. Risk measurement in fixed income markets is challenging as we discussed in the first item above as backward-looking risk measures may not reveal the potential risks going forward. While prices and / or quantities in credit markets may look unusual compared to historical data, such episodes are often also characterized by arguments in favor or against why "this time is different." Again, the convergence in European treasury yields provides a useful example. From a risk measurement and management perspective, this is a major challenge and perhaps more challenging for fixed income markets as we have seen a secular decline for almost four decades in interest rates and the path going forward is uncertain.

   To address the last two items, we propose an idiosyncratic tail risk exposure measure for each country that depends on three inputs: (1) the country's market portfolio weight, (2) the country's government bond duration and (3) adverse yield change scenarios. We have used the available historical data across countries to deduce a set of extreme scenarios that is the same for all countries. This approach recognizes that it is very hard to predict in which country or region the next extreme interest rate changes will occur, particularly over longer horizons, and it is therefore useful to apply the most extreme yield changes in the whole panel to generate the scenarios. It is of course possible to model the scenarios as a function of country or regional characteristics such as fiscal strength if one has a particularly strong view on where the next crisis is most likely to occur. We believe such explicit views belong in the active management discussion of the fund rather than in the discussions regarding the benchmark.

   In addition to using historical data, we strongly recommend additional scenario analysis and stress testing as a potential tool to highlight the potential risks in the current portfolio. Stress

testing is being used more extensively in the context of bank and insurance company supervision, and we think that the insights from these analyses are helpful to complement traditional risk measures, both in terms of understanding the benchmark risks and the active risks.

We end this report by formulating specific answers to the questions posed to us.

(a) *What are the considerations of various portfolio construction rules and which currencies should be included in the index?* We recommend that, to the extent possible, all currencies are included in the index. Even when historical correlations have been reasonably high between segments or countries, suggesting that a certain segment or country is spanned by others, there is no guarantee that this will continue in the future. In addition, when adhering to market weights, the large size of the Fund is less of a concern as countries or segments that are small will also receive a small portfolio weight in the Fund's portfolio. This is not necessarily the case when other weighting schemes are employed.

(b) *Which segments should be included in the benchmark?* We recommend that, to the extent possible, all segments are included in the index. The logic is the same as explained under item (1).

(c) *How should the duration (that is, the sensitivity to interest rate risk) of the bond portfolio be chosen?* We recommend that the duration of the portfolio is chosen in accordance with the duration that follows from market weights. We view the decision to increase or decrease exposure to the term premium as an active investment management decision.

(d) *Should inflation-linked bonds be included in the index?* We recommend that inflation-linked bonds are included in the index using market weights, following the same guiding principle that being widely diversified helps mitigate long-term idiosyncratic risk exposures and maximize investment capacity.

(e) *Are there risk premiums which should not be harvested through the benchmark index?* We have discussed the recent development of factor-based investment strategies. There is

growing evidence that factor based investment strategies are able to generate enhanced risk return tradeoffs, though many different implementation schemes are currently put into practice in the investment management landscape. We view the level of exposure to these factors as an active investment management decision.

References

Ang, A. (2014): "Asset Management: A Systematic Approach to Factor Investing," *Oxford University Press*.

Asness, C., T. Moskowitz, and L. H. Pedersen (2013): "Value and Momentum Everywhere," *Journal of Finance*, 68, 929–985.

Asvanunt, A., and S. Richardson (2017): "The Credit Risk Premium," *The Journal of Fixed Income*, 26, 6–24.

Black, F., and M. Scholes (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–654.

Brooks, J., and T. J. Moskowitz (2017): "Yield Curve Premia," *Working paper, Yale SOM*.

Cieslak, A., and P. Povala (2015): "Expected Returns in Treasury Bonds," *The Review of Financial Studies*, 28, 2859–2901.

Cochrane, J. H., and M. Piazzesi (2005): "Bond Risk Premia," *American Economic Review*, 95, 138–160.

Dahlquist, M., and H. Hasseltoft (2013): "International Bond Risk Premia," *Journal of International Economics*, 90, 17–32.

De Bondt, W. F.-M., and R. Thaler (1985): "Does the Stock Market Overreact?," *Journal of Finance*, 40(3), 793–805.

Ellul, A., C. Jotikasthira, and C. T. Lundblad (2011): "Regulatory pressure and fire sales in the corporate bond market," *Journal of Financial Economics*, 101(3), 596–620.

Fama, E. F., and K. R. French (1993): "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 33, 3–56.

Frazzini, A., and L. H. Pedersen (2014): "Betting Against Beta," *Journal of Financial Economics*, 111(1), 1–25.

Greenwood, R., and S. Hanson (2013): "Issuer Quality and Corporate Bond Returns.," *Review of Financial Studies*, 26, 1483–1525.

Greenwood, R., and A. Vissing-Jorgensen (2018): "The Impact of Pensions and Insurance on Global Yield Curves," *Working paper, Harvard Business School*.

Harvey, C. R., Y. Liu, and H. Zhu (2016): "... and the Cross-Section of Expected Returns," *The Review of Financial Studies*, 29, 5–68.

Israel, R., D. Palhares, and S. Richardson (2018): "Common factors in corporate bond returns," *Journal of Investment Management*, 16, 17–46.

Jiang, Z., A. Krishnamurthy, and H. N. Lustig (2018): "Foreign Safe Asset Demand for U.S. Treasurys and the Dollar," Working paper, Stanford University.

Koijen, R., T. Moskowitz, L. H. Pedersen, and E. B. Vrugt (2018): "Carry," *Journal of Financial Economics*, 127(2), 197–225.

Koijen, R. S., F. Koulischer, B. Nguyen, and M. Yogo (2018): "Inspecting the Mechanism of Quantitative Easing in the Euro Areal," *Working paper, University of Chicago, Booth School of Business*.

Krishnamurthy, A., and T. Muir (2016): "How Credit Cycles across a Financial Crisis," Working Paper, Stanford GSB.

Krishnamurthy, A., and A. Vissing-Jorgensen (2007): "The Demand for Treasury Debt," NBER Working Paper No. 12881.

——— (2011): "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy," *Brookings Papers on Economic Activity*, pp. 215–265.

——— (2012): "The Aggregate Demand for Treasury Debt," *Journal of Political Economy*, 120, 233–267.

Lustig, H. N., A. Stathopoulos, and A. Verdelhan (2018): "The Term Structure of Currency Carry Trade Risk Premia," *Working paper, Stanford GSB*.

Maggio, M. D., A. Kermani, and C. Palmer (2016): "How Quantitative Easing Works: Evidence on the Refinancing Channel," *Working paper, Harvard Business School*.

Maggiori, M., B. Neiman, and J. Schreger (2018): "International Currencies and Capital Allocation," *Working paper, Harvard University*.

Merton, R. C. (1974): "On the pricing of corporate debt: The risk structure of interest rates," *Journal of Finance*, 29, 449–470.

Mian, A., A. Sufi, and E. Verner (2017): "Household Debt and Business Cycles Worldwide," .

Nakamura, E., J. Steinsson, R. Barro, and J. Ursua (2013): "Crises and Recoveries in an Empirical Model of Consumption Disasters," *American Economic Journal: Macro-economics*, 5, 35–74.

Schularick, M., and A. M. Taylor (2012): "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008," *American Economic Review*, 102, 1029–1061.

Sharpe, W. (1964): "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425–444.

Vayanos, D., and J.-L. Vila (2009): "A Preferred-Habitat Model of the Term Structure of Interest Rates," *Working paper, London School of Economics*.

A Data availability

We list, per segment, the sample periods that are available for different countries.

Table A1: Data availability Treasury bonds in developed markets.

Country	Start	Date	Country	Start	Date
Australia	29-Apr-88	29-Jun-18	Lithuania	30-Jan-15	29-Jun-18
Austria	26-Feb-99	29-Jun-18	Luxembourg	26-Feb-10	31-May-18
Belgium	26-Feb-99	29-Jun-18	Netherlands	26-Feb-99	29-Jun-18
Canada	30-Jan-87	29-Jun-18	New Zealand	27-Feb-87	28-Feb-18
Denmark	27-Feb-87	31-May-18	Norway	31-May-91	31-May-18
Finland	26-Feb-99	28-Feb-18	Singapore	28-Feb-02	29-Jun-18
France	29-Jan-99	29-Jun-18	Slovakia	31-Jan-05	29-Jun-18
Germany	26-Feb-99	29-Jun-18	Slovenia	31-Jan-05	29-Jun-18
Hong Kong	30-Sep-04	29-Jun-18	Spain	29-Jan-99	29-Jun-18
Ireland	26-Feb-99	29-Jun-18	Sweden	27-Feb-87	31-Jan-18
Italy	26-Feb-99	29-Jun-18	Switzerland	29-Jan-10	29-Jun-18
Japan	27-Feb-87	30-Mar-18	UK	27-Feb-87	31-May-18
Latvia	28-Feb-14	29-Jun-18	US	30-Jan-87	31-May-18

Table A2: Data availability Treasury bonds in emerging markets.

Country	Start	Date	Country	Start	Date
Argentina	31-Jul-08	31-Jul-18	Nigeria	30-Apr-13	31-Dec-15
Brazil	31-Jul-08	31-Jul-18	Offshore China	30-Apr-13	31-Jul-18
Chile	31-Jan-05	31-Jul-18	Peru	31-Jul-08	31-Jul-18
China	31-Jul-08	31-Jul-18	Philippines	31-Jul-08	31-Jul-18
Colombia	31-Jul-08	31-Jul-18	Poland	31-Jan-05	31-Jul-18
Croatia	31-Jul-08	31-Jul-18	Romania	30-Apr-13	31-Jul-18
Czech Republic	31-Jan-05	31-Jul-18	Russia	31-Jul-08	31-Jul-18
Egypt	31-Jul-08	31-Jul-18	South Africa	31-Jan-05	31-Jul-18
India	31-Jul-08	31-Jul-18	South Korea	28-Feb-02	29-Jun-18
Indonesia	31-Jul-08	31-Jul-18	Taiwan	31-Jan-06	31-Jul-18
Israel	31-Jul-08	31-Jul-18	Thailand	28-Feb-02	31-Jul-18
Malaysia	31-Jan-06	31-Jul-18	Turkey	31-Jul-08	31-Jul-18
Mexico	31-Mar-05	31-May-18			

Table A3: Data availability inflation-linked bonds.

Country	Start	End	Country	Start	End
Australia	31-Aug-12	30-Apr-18	Japan	31-Jan-06	31-May-18
Canada	31-Oct-97	30-Mar-18	New Zealand	30-Apr-14	31-May-18
Denmark	31-Dec-12	31-May-18	Spain	28-Nov-14	31-May-18
France	28-May-99	31-May-18	Sweden	30-Jan-98	31-May-18
Germany	28-Apr-06	30-Apr-18	UK	30-Jan-98	30-Apr-18
Greece	30-Apr-04	31-May-10	US	31-Oct-97	30-Apr-18
Italy	31-May-04	31-May-18			

Table A4: Data availability other government bonds.

Country	Start	End
Agencies	31-Jan-01	29-Jun-18
Local Authorities	31-Jan-01	29-Jun-18
Sovereign	29-Sep-00	29-Jun-18
Supranational	29-Sep-00	29-Jun-18

Table A5: Data availability corporate bonds.

Country	Start	End	Country	Start	End
Australia	31-Jan-01	31-May-18	Mexico	31-Jan-01	31-Jul-18
Austria	30-Mar-01	31-May-18	Netherlands	30-Mar-01	29-Jun-18
Belgium	31-Jan-01	31-May-18	SouthKorea	31-Jan-01	31-Jul-18
Canada	30-Mar-01	29-Jun-18	Spain	28-Feb-01	31-May-18
France	31-Jan-01	29-Jun-18	Sweden	31-Jan-01	29-Jun-18
Germany	31-Jan-01	29-Jun-18	Switzerland	31-Jan-01	30-Apr-18
Italy	30-Apr-01	29-Jun-18	UK	28-Feb-01	29-Jun-18
Japan	31-Jan-01	29-Jun-18	US	28-Feb-01	29-Jun-18

Table A6: Data availability covered bonds.

Country	Start	End	Country	Start	End
Austria	31-Oct-03	29-Jun-18	Netherlands	31-Mar-99	31-May-18
Canada	31-Dec-07	31-May-18	Norway	31-Aug-07	29-Jun-18
Denmark	30-Sep-04	29-Jun-18	Portugal	28-Feb-07	29-Jun-18
Finland	30-Nov-05	29-Jun-18	Spain	28-May-99	29-Jun-18
France	30-Nov-99	31-May-18	Sweden	29-Mar-02	29-Jun-18
Germany	29-Jan-99	31-May-18	Switzerland	26-Feb-10	29-Jun-18
Ireland	31-Jan-00	31-May-18	UK	29-Aug-03	29-Jun-18
Italy	30-May-03	30-Mar-18	US	31-Oct-06	31-Mar-16
Luxembourg	29-Jan-99	29-Jun-18			

Appendix

Terms of reference for an expert group appointed to review the fixed-income investment framework for the Government Pension Fund Global

1 Background

The capital of the Government Pension Fund Global (GPFG) has its origin in accumulated central government petroleum revenues from the extraction of oil and gas on the Norwegian continental shelf. The market value of the Fund's investments was NOK 7,952 billion at the end of September 2017.

The governance structure for the Fund is based on a clear division of roles and responsibilities. Overarching management responsibility lies with the Ministry of Finance, whilst operational management execution is performed by Norges Bank (the central bank). Fundamental strategic choices have been endorsed by the Storting (parliament). Norges Bank manages the GPFG on the basis of a mandate laid down by the Ministry of Finance. The mandate expresses the key features of the investment strategy for the Fund, and includes, *inter alia*, provisions on benchmark indices, risk taking limits, reporting requirements and responsible management.

The GPFG is primarily invested in listed equities and fixed-income instruments. The Fund is also invested in unlisted real estate within limits stipulated in the mandate. The investment objective is to achieve the highest possible return at a level of risk that is acceptable to the Fund owners. The investment strategy for the Fund has been developed gradually on the basis of assessments of expected risk and return, as well as the purpose and distinctive characteristics of the Fund and the comparative advantages of the asset manager, along with the fundamental investment beliefs of the Ministry.

The investment strategy for the GPFG is premised on the following key principles:

- → Overall risk in the Fund can be reduced by diversifying the investments across a large number of securities.
- → The equity and fixed-income benchmarks are based on broad, global sub-indices and segments from recognized index providers with additional bespoke filters such as ethical screening. The indices shall be investable and shall as a main rule lend themselves to being replicated closely and at low cost.
- → Risk taking shall be clearly communicated and broadly endorsed. There is a high degree of transparency concerning the basis for the investment strategy and a moderate degree of active management.

Over time, the risk level in the GPFG has increased and the Fund's benchmark and investment universe have been expanded by, *inter alia*, increasing the equity share and permitting investments in additional countries, currencies, asset classes and financial instruments. In 2017, the Storting endorsed an increase in the equity share of the strategic benchmark index to 70 percent, with the phase-in being implemented over time. Fixed-income instruments will make up the remaining 30 percent of the index. In conjunction with the decision to increase the equity portion in the Fund, a review of the fixed-income benchmark index was announced in the white paper to parliament on the management of the Fund in the spring of 2017. As part of the ongoing review, advice and

assessments have been obtained from Norges Bank. In order to broaden the decision-making basis further, the Ministry has decided to appoint an expert group to assess the topics outlined in section 3.

2 The current fixed-income benchmark

The purpose of the Fund's fixed-income investments is:

- → to reduce the volatility of overall Fund returns,
- → contribute liquidity, and
- → provide exposure to fixed-income risk factors, such as interest rate risk and credit risk.

The current fixed-income benchmark index was adopted in 2012 on the basis of a strategic equity share of 60 percent, and reflects the said purposes. The benchmark is based on index products provided by Bloomberg L.P., and comprises a government bond portion (70 percent) and a corporate bond portion (30 percent). The apportionment between the two parts of the benchmark is fixed, with full monthly rebalancing back to the chosen weights. The fixed-income benchmark is exclusively comprised of investment-grade securities. Bonds from Norwegian issuers and bonds denominated in Norwegian kroner are not included in the benchmark index.

The composition of the government bond portion of the fixed-income benchmark is based on the securities included in the underlying indices at any given time, and consists of nominal government bonds, inflation-linked government bonds and bonds issued by supranationals. The country distribution within the government bond portion is calculated on the basis of the size of the economy of each country, as measured by gross domestic product (GDP) in US dollars. The country weights are rebalanced back to the original weights monthly. Within each country, sub-segments and individual bonds are weighted by market weights. Certain country weights in the government bond portion are supplemented by adjustment factors motivated by investability considerations (to avoid excessive ownership shares in countries with high GDP relative to the size of their government bond market)[1]. Investability is of particular concern, given the considerable size of the Fund. In addition, the management mandate requires Norges Bank to take account of differences in fiscal strength between countries in the composition of the government bond investments. The latter mandate provision implies that the asset manager cannot necessarily closely replicate the benchmark index. The corporate bond portion comprises corporate bonds and covered bonds issued in seven approved currencies[2]. See Section 3-2 of the Ministry's mandate for the management of the GPFG for a detailed description of the fixed-income benchmark.

The Ministry has in its management mandate defined an investment universe for fixed-income instruments which is broader than the benchmark index. The Bank shall seek to keep any deviations from the benchmark index in the composition of the actual portfolio within the mandated limit for expected tracking error of 1.25 percentage point. In addition, the Bank is required to have limits for the minimum overlap between the actual fixed-income portfolio and the fixed-income benchmark, separate limits for tail risk, as well as credit risk limits for individual investments and at the portfolio

[1] An adjustment factor of 0.25 is applied to the country weights for Chile, Hong Kong and Russia.
[2] USD, CAD, EUR, GDP, SEK, DKK, CHF.

level. In order to ensure that Norges Bank is not forced to immediately divest bonds that are omitted from the benchmark index as the result of their credit rating being downgraded below investment grade, it is permitted to hold up to 5 percent of the fixed-income portfolio in high-yield bonds.

3 The assignment

The expert group shall by [1 October 2018] submit a report containing analyses and assessments. The report will form part of the basis for the Ministry's assessments of a suitable fixed-income investment framework for the GPFG, including the fixed-income benchmark. The expert group shall in its analyses and assessments attach weight to the characteristics of the overall strategic benchmark index comprising equities and fixed-income instruments (70 percent and 30 percent, respectively). The group shall base its analyses and assessments on the Ministry's key investment strategy principles.

In a letter 1 September 2017 Norges Bank proposed several changes to the benchmark index, including a significant reduction in the number of currencies in the fixed-income benchmark, and to limit the composition of the index to nominal government bonds only. The bank further recommended to leave the investment universe unchanged. In light of Norges Bank's advice the Ministry requests the expert group to analyse and assess the following aspects:

a) Choice of currencies and index weighting principle

The group is requested to analyse how various country and currency compositions in the fixed-income benchmark may contribute to meeting the principle of broad diversification. The analysis shall be conducted on the basis of various weighting principles, including market weights (with or without adjustment factors) and GDP weights. Advantages and disadvantages of the various weighting principles shall be discussed. The significance of emerging markets for diversification shall be addressed. The group shall also assess whether the long investment horizon of the Fund suggests that the index rules should pay special heed to capturing any changes in the fiscal strength of government bond issuers over time.

b) Choice of segments

The group is requested to assess whether other segments than nominal government bonds from developed economies should be included in the benchmark index, including corporate bonds, covered bonds, bonds issued by supranationals and nominal government bonds from emerging economies issues in local currency. The group is requested to assess the expected credit premium in this context, and how one should ensure exposure to such premium.

c) Choice of duration and sensitivity to interest rate changes

The group is requested to assess the expected term premium, and how one should ensure exposure to such premium. The group is also requested to assess whether there is reason to assume segmentation in the bond market and the significance thereof for the index rules.

d) Inflation-linked government bonds

The group is requested to assess whether inflation-linked government bonds should be included in the fixed-income benchmark.

e) *Risk premiums which should not be harvested through the benchmark index*

If the group is of the view that there are bond market risk premiums to which the GPFG should be exposed, in addition to term and credit premiums, but which are not suited for inclusion in the benchmark index, an assessment is requested as to how the Ministry should stipulate risk limits in the mandate for any exposure to such premiums.

4 Relevant background information

- Report of 29 March 2011 from Stephen Schaefer and Jörg Behrens: NMOF Fixed Income Review – Final Report
- Letter of 9 June 2017 from the Ministry of Finance to Norges Bank on the fixed-income benchmark for the GPFG
- Letter of 1 September 2017 from Norges Bank to the Ministry of Finance on the fixed-income investments in the GPFG
- Letter of 26 October 2017 from the Ministry of Finance to Norges Bank on the fixed-income investment framework for the GPFG
- Letter of 14 December 2017 from Norges Bank to the Ministry of Finance on the fixed-income investment framework for the GPFG
- Mandate for the management of the GPFG
- Meld. St. 17 (2011–2012) Report to the Storting on Management of the Government Pension Fund in 2011, chapter 2.2
- Meld. St. 23 (2015–2016) Report to the Storting on Management of the Government Pension Fund in 2015, chapter 8