IRSG

INSURANCE AND REINSURANCE STAKEHOLDER GROUP

Advice on Consultation Paper on the Methodology on Value for Money Benchmarks

IRSG-24-16

European Insurance and Occupational Pensions Authority

1. IRSG SUMMARY

We welcome the opportunity to respond to EIOPA's Consultation on its methodology for developing Value for Money (VfM) benchmarks. In the past we have been active with this same topic. On early 2023, we provided an advice on the methodology to assess value for money in the unit-linked market (LINK) which already included many important views how to consider value for money aspect for UL products.

Giving its advice to this consultation, the IRSG believes that a proper Value for Money (VfM) assessment should holistically consider all the features and benefits of insurance products, recognise markets' diversity and respect manufacturers' autonomy in setting prices. Unit-linked and hybrid products, which meet diverse customer needs, already benefit from a robust regulatory framework and provides authorities with sufficient elements and powers to intervene where necessary. EIOPA has also provided extensive guidance in its recent Statement and Methodology, which needs time to bed-in.

Some of the members find that the proposal for VfM benchmarks is premature and would lead to unnecessary administrative burdens for product manufacturers. They also believe that benchmarks should not be disclosed. Publishing them could lead to misinterpretation, thereby misguiding consumers, harming their trust and nudging them towards less costly products without considering the risks and benefits.

Other members find that the VfM work is needed now and has been identified in the past as a supervisory issue by many NCAs and by EIOPA itself. They believe that benchmarks should be published and shown to the consumers once the process is finalised. This would be the only way consumers could have something in hand for comparison as no comparison websites exist currently locally. They would also point out that the performance benchmarks should be the relevant capital market benchmarks (those chosen by the providers themselves) instead of merely peer group benchmarks which do not guarantee the objective value for money of a given product, and target only "outliers" of each peer group every time. Moreover, they would strongly recommend not to use the performance and cost data of the PRIIPs KIDs, which are only future projections and based only on short term past performance (5 years), and instead use the actual historical data, as the French NCA is currently doing for its own VfM initiative with the insurance industry.

We find that the main risks in the VfM benchmark work are:

- The current work on benchmarks could unduly focus on costs and therefore encourage a “race to the bottom” for essential features, i.e. product manufacturers decreasing the quality and variety of the products to meet the benchmark (eg, products with lower or no financial guarantees, more limited biometric risk covers, less choices for consumers, etc.). This would not be in the best interest of customers. All of this could have detrimental effects on consumers, as it would become more difficult for them to find products that fully meet their needs, especially in terms of safety and protection. This would not encourage consumers to invest and would be against the objectives of the Capital Markets Union (CMU).
- Multi-Option Products (MOPs) are the preferred solution of many consumers for their adaptability and variety. Between the two suggested options there’s problems in both. Option 1 bears the risk to be oversimplistic and would lead to misleading comparisons. Option 2 puts more emphasis on the available options but could distract from the product features if the options only play a minor part. More work should be done if still aiming to come up with feasible methodologies. Some members propose a third option: clustering by unit, not by contract, to consider the fact that many MOPs offer hundreds of different units, and this is the option chosen by the French NCA for its current VfM initiative with the French Insurance industry.

Finally, we believe that there needs to be a clear distinction between supervisors’ work to identify outliers and companies’ internal work on product design. Supervisors may have needs to refine their tools to monitor the market, but these tools should not be used by companies as part of POG. In our opinion, this is the only way to maintain the POG flexible and focused on consumers’ needs, while avoiding additional burdens and unintended restrictions for insurers. Quoting EIOPA, “The product should provide value for money in itself, regardless of the comparison with the products sold in the EU market.” (p. 10 of the consultation paper).

IRSG ANSWERS TO SPECIFIC QUESTIONS ASKED BY EIOPA

Q1: Stakeholders are invited to provided inputs and views as to how value for money benchmarks should work and their usefulness for product comparability.

The IRSG is of the opinion that in theory, value for money benchmarks at a sufficiently but not excessively granular level should be very useful in allowing effective comparison between products of similar type. For instance, targeted KPIs can provide insights to comparative charging structures and enable assessment of where a product may not add value. Benchmarks should also be guided by a principle that, as a first indication of outliers, their purpose should not be to cover all elements of products addressed, but to give high level indication of the need for further investigation. We see examples in EU where it is not easy to compare products and it is still a complex exercise to make comparisons between products.

We also believe that there needs to be a clear distinction between supervisors' work to identify outliers and companies' internal work on product design. Supervisors may want to refine their tools to monitor the market, but these tools should not be used by companies as part of POG. In our opinion, this is the only way to maintain the POG flexible and focused on consumers' needs, while avoiding additional burdens and unintended restrictions for insurers. Quoting EIOPA, "The product should provide value for money in itself, regardless of the comparison with the products sold in the EU market." (p. 10 of the consultation paper).

We believe that the purpose of the benchmark work should also be clarified. As a first indication of potential outliers, it should not be such that it acts as a constraint on product innovation, variety of features, or inadvertently leads to customer detriment in any way through the behaviour it encourages. It is important that benchmarks encompass all quantitative and qualitative features of the product like life protection, disability cover, flexibility etc. In this way product diversity and innovation could be maintained. Some members point out that, since there are lack of comparison websites (Malta being one good example), these benchmarks will finally overcome the complex analysis required to compare products which the non-savvy consumer is not able to do.

The extent of product parameters/charges can theoretically vary from product to product, company to company, channel to channel, country to country and for MOPs from fund to fund. This means that there could be an excessively complex set of KPIs and clusters. A further clear guiding principle of this work, which will serve well in the long term, should be that benchmarks will be simple, and that quest for simplicity should be at the heart of all the work done. For instance, a cost indicator could be one of the elements to effectively detect potential outliers without unnecessary burden. Some members would bring out that RIY as an example, has the potential to be a very useful cost benchmark as it brings the full charging structure together in one simple measure, with the potential to vary the terms over which it is measured.

We find that to be properly comparable, benchmarks should include all charges which are made to the customer, including all distribution related charges. Furthermore, product profitability is a good guide as to the level of value being provided to customers in return for the charges imposed, though clearly profitability on its own does not give an insight into the competitiveness of a company's cost base. POG processes should all include assessment of profitability. It would not be meaningful to seek to include profitability as a benchmark, but this should be an important measure where company specific aspects come to be assessed for outliers.

Some of the of the member believes that to be meaningful, product comparisons via benchmarks should not seek to incorporate investment performance measures in their remit. At least, if investment performance is incorporated, it should be done in a way which clearly distinguishes between the impact of charges and the impact of performance. Failure to clearly attribute in this way would likely lead to confusion in assessment of value.

We suggest ensuring a clear assessment process and avoid blurring the concept of VfM with the limited scope of benchmarks, by emphasizing:

- the supervisory nature of the VfM benchmarks, a tool with the purpose of supporting NCAs in their risk-based supervisory approach
- benchmarks cannot assess product quality and are not useful for product comparability in the POG process and
- some members state also that they must not be used as a public comparison tool aimed at consumers, but others state they should be disclosed.

As an observation, we disagree with the generalized statement that "unit-linked …products are overly complex with limited direct benefits for customers". This may well apply to some products, but it is wrong to generalize in this way.

We would also like to point out that it might be valuable to EIOPA to compare to supervisory experiences outside of the EU 27 by other jurisdictions, in particular the value assessments performed by the UK NCA in 2020-21.

Finally, some of the IRSG members would like to bring out that even if limited to a few products in the insurance sector, the work relating to value for money concerns a considerable number of consumers. Although the insurance and financial industry and the authorities consider governance concerns to be the highest, placing greater emphasis on value for money considerations as part of the POG process will benefit consumers and facilitate the integration of retail investors in CMU. These members would also like to point out that the value for money benchmarks should be designed with a holistic approach, considering not only the immediate cost but also the long-term benefits and overall quality of the product or service. For these members, transparency in the methodology of establishing benchmarks is crucial to gaining the trust of consumers and stakeholders alike. Relevant factors that influence the overall value proposition should be considered like durability, performance, environmental impact, and customer satisfaction. Stakeholders should be involved in the identification and weighting of these factors to reflect diverse perspectives. Visual representations and clear metrics should be employed to enhance understanding. There should be periodic – but not too frequent – reviews to stay relevant. There should also be a commitment to ongoing consumer education. These members believe that by incorporating these considerations, value for money benchmarks can serve as valuable tools for enhancing product comparability and fostering a market environment that prioritizes consumer interests. Currently the situation regarding costs and charges has remained the same i.e. nor easy to compare products and it is still a complex exercise to make comparisons between products.

Q2: Stakeholders are also invited to share whether they agree on what the benchmarks are and are not.

The IRSG finds that use of many products and potentially geographic clusters, combined with "multiple qualitative and quantitative indicators" presents very likely an unduly complex system. Therefore, the intent of the proposal may, as currently set out, be excessively wide.

We believe that the definition of 'benchmarks' should go hand-in-hand with how benchmarks are used. Monitoring of a possible market misconduct requires surely different methods than for the savings product offering. In order to tackle the supervisory market monitoring needs, it might not even be that important to have well defined 'benchmarks'. Any market offering analysis within the supervisory college needs to meet only its own needs and level of understandability and simplicity, as all the users are experts. Moreover, detecting possible abuses does not require a very refined set of indicators - focusing on costs should do the job in many cases. Afterwards, a specific analysis, to be carried on case by case by the supervisory authorities on the outlier detected, should be considered - but this would take place completely out of the framework of market wide benchmarks. For other 'benchmark' purposes, we believe that further work with the already comprehensive and good costs & past performance report would already do a lot.

Overall, we see merit in the following key elements of the VfM benchmarks:

- The intention to facilitate and enable a more risk-based approach to supervision and assertion that benchmarks are not intended as a consumer disclosure tool, reflecting the requirement of in-depth knowledge of the VfM methodology which is not designed with consumers use in mind.[1] Benchmarks used by customers for product comparison would be highly misleading. The benchmarks are intended as a supervisory tool to identify unusually expensive products. Products within a benchmark do not have necessarily a low/an average price or good quality. However, benchmarks might imply to customers that products within the benchmarks are particularly favourable since they (over)achieve the requirements of NCA's
- the acknowledgement of the multi-dimensional angle of VfM, as both quantitative and qualitative indicators are important and
- National specificities might be required to give enough detail on key aspects.

Some of the members would emphasize that benchmarks don't seem to be useful for product manufacturers in the POG process. The POG process requires testing that products align with the target market's needs, objectives, and characteristics. Also, EIOPAs 2021

[1] Paragraph 3.6 of consultation paper

supervisory statement on value for money requires extensive cost identification as part of the POG, which should lead to better transparency, where needed. Being within the benchmarks that are supposed to identify outliers (products significantly more expensive than the average) does not indicate any alignment with the customer needs. EIOPA emphasizes for this reason, that benchmarks are no safe harbor. Benchmarks that do not indicate quality or a good price-level and that are no safe harbor are not helpful in the POG process. Also, manufacturers do not need benchmarks to determine the price in relation to other products on the market. These members would also like to point out that POG rules, the benchmark statement and the methodology must be designed in a way to ensure that consumer interests are considered not only during the product design stage, but also throughout the entire life cycle.

Other members would like to point out the importance of sustainability features. Taking into account the sustainability in the benchmark methodology would allow market participants to make well-informed choices with the overall objective of protecting consumers and investors. Consideration of climate impact might need a long-term perspective from both the manufacturers and the customers with might sometimes be contradictory to the short-term objectives. Consumers should be able to understand well the sustainability features and their impact to the products.

The IRSG finds that while authorities might find it useful to further develop their supervisory tools to identify outliers in the market, this should neither result in additional administrative burdens and reporting for insurance companies, nor lead to convergence and standardisation. Product manufacturers should not be asked to make use of new product clusters, indicators and benchmarks in their product testing or pricing, to prevent product homogenisation, dampened competition, and limited choices. Consumers' needs should remain a priority, and insurers' freedom to design products and set prices must be respected. In this light, some of the IRSG members believe that benchmarks should not be used by product manufacturers, distributors and consumers. Besides, value assessments should not be based on rankings of products.

Some of the members are of the opinion that the performance "benchmarks" considered by EIOPA (and by the French NCA) are merely peer group benchmarks, not objective ones, on the performance of the underlying assets. Unit managers themselves have stated investment objectives and therefore explicit (mostly) or implicit objective benchmarks. They believe that those are the ones which should be used. And if not, one may well have a “cluster” average performance that is below the underlying market performance, which means in that case that most existing units in that cluster do not provide adequate value for money. An example of this is when all money market units generate lower performance than the money market itself, or it can be the reverse, and then EIOPA may exclude units that are however providing adequate VfM. These IRSG members find that this is also why it is crucial not to create separate clusters for index-managed units, as citizens may and should have the choice – for a given asset class and SRI level – between an actively managed and an indexed unit. Taking out the index ones from the cluster introduces a major flaw in the VfM Assessment (cost/benefit assessment for the same asset and risk class). As brought out in question 1, EIOPA would benefit from assessing the “Value Assessments” conducted by the UK FCA in this regard.

The IRSG finds that to this stage the way how benchmarks are planned leaves it open whether the benchmarks are supposed to be pan-European benchmarks or national benchmarks. Given the fact that national markets substantially differ from each other, eg in terms of insurance premium and capital gains taxation, mandatory integration and extent of biometric risk, interest rate, authorized investment options, different capital protection mechanisms, etc. pan European benchmarks would not be feasible. However, national benchmarks set by a European authority like EIOPA would infringe the principle of subsidiarity pursuant to Article 5(3) of the Treaty on European Union (TEU). We also question the value for customers of comparing products offered to them against any pan-European benchmark that is possibly balanced to markets very different than their own.

Furthermore, we believe that “benchmarks” should provide supervisors with a common set of data to objectivize as much as possible their assessment of VfM. In that sense, benchmarks must not lie only on cost factors, but consider all elements of performance, including risk protection techniques such as guarantees or qualitative features. This implies that product clusters are specific and homogenous enough to avoid biased comparisons or conclusions. Given the diversity of products across national markets, much more important for insurance than for asset management, it may lead to adopt a European methodology for benchmarking, leaving at NCAs level the detailed implementation.

Q3: Do you already have similar tools in your market that would serve the same purpose?

We see that currently, some national competent authorities (NCAs) are tailoring VfM approaches to fit their markets, showing that EU-wide uniformity is not feasible and VfM supervision needs flexibility. One finds examples in Germany by the BaFin-Merkblatt zu kapitalbildenden Lebensversicherungen and in France by a marketplace initiative resulting from a discussion between ACPR, France Assureurs. NCAs have been driving the Value for Money assessment, leveraging on the concept of benchmarks, with market players for several years. The experience built on such initiatives is valuable and should therefore be recognized in the ongoing EIOPA discussion. As such, we emphasize the need to safeguard the national approaches in VfM benchmarks already in place in several markets and that NCAs need adequate leeway to take national specificities into account.

The product information is generally available in the Irish market, for instance on product charges. This information could be used to establish charging related benchmarks. Equally, PRIIPS information, e.g. as to RIYs, is available. And then in Malta, each manufacturer provides the charge structure in the KID yet they are all different and difficult to compare as they do not use the RIY. Some have hefty upfront charges, some hidden charges which can only be estimated from the projected maturity values. No comparison is possible as no common charge structure is used in Malta.

In France, the common initiative of ACPR and insurance companies, launched end 2022 and developed over 2023 to identify “outliers” in terms of VfM among UL products has led to a form of consensual benchmarking and to measures taken by companies and reported to the NCA to effectively increase VfM for customers. On this case, it was only on the UL part of the French MOPs market. ACPR has already identified 27 “clusters” crossing asset classes (mostly ECB classification, but with some significant deviations from it) with the SRI level, in a more granular way than EIOPA. For some members, EIOPA’s methodology could be simplified by dropping down the first clustering criterion; UL / hybrid product. Indeed, in France, the VfM process looks only at unit-linked, leaving aside the capital guaranteed component (the UK’s “with profit” policies) of the MOPs. EIOPA should look at the French ACPR which (despite deviations and flaws from its own approach) seems more realistic.

Some of the IRSG members believe that a step-by-step approach starting with standard MOPs; UL contracts with or without optional and separately priced biometric coverage would be preferred. The reason for this is that the UL and Hybrid products offering in EU is very fragmented with a lot of diversity and very limited cross-border offering.

Q4: While EIOPA indicated that initially it will not publish the benchmarks, stakeholders are also invited to share views as to whether the benchmarks should be published or not already in the first initial phase.

Some of the IRSG members believe that the benchmarks should be published and shown to the consumers once the process is finalised. This would be the only way consumers could have something in hand for comparison as no comparison websites exist currently locally. Consumers tend to buy on trust. As a remark, these members would like to bring out an example from France, where the life insured have protested that the NCA (ACPR) has not only not consulted them or their representatives on its VfM initiative but did not even provide any information to them. Now, as EIOPA is consulting publicly on the topic, the industry would most probably be informed on the finalized and adopted methodology and results. This then helps ensuring that EIOPA is not designing too complex methodology and rules and would hopefully reverse the too often paternalistic behavior from the EU supervisor, which – for example – still labels EU citizens as pension savers as "beneficiaries", versus the much more adequate "participants" in the US. These members believe that once the methodology and benchmarks are finalized, they must be disclosed to the life insured citizens. It is a fundamental transparency and life insured protection matter. In most other consumer markets, citizens have access to comparable price and performance data.

Other IRSG members believe that VfM benchmarks should not be disclosed. Publishing them could lead to misinterpretation, thereby misguiding consumers, harming their trust and nudging them towards less costly products without considering the risks and benefits. If developed, benchmarks should be for internal supervisory use only. These members finds that the VfM assessment is a matter for expert discussion at NCA vis-à-vis manufacturer level and benchmarks should be treated as such as a result. Also EIOPA has stated that the nature of the indicators and the product clustering process require an in-depth knowledge of the VfM methodology which is not targeted for consumers[2]. Benchmarks are neither suitable nor intended to be used in or referred to the advice process, as they can only capture a part of the products' benefits and features. They should therefore be excluded from publication. This assertion is valid irrespective of the phase of development and/or refinement of the methodology, given that an in-depth knowledge of the VfM methodology is not aimed for consumers. These same members would like to emphasize that

[2] Paragraph 3.6 of consultation paper

benchmarks, if used as a disclosure tool, would – de facto - result in price and competition control and contradict the insurers' freedom to set prices as stated in Articles 21 and 182 Solvency II Directive. The European legislator confirmed in Recital (8) of the IDD Delegated Regulation on POG that the POG requirements should not be understood as an interference with the manufacturers' freedom to set premiums or as a price control in any form. They would like to point out also that benchmarks carry the risk to limit competition in two ways:

First, benchmarks may work as a barrier for (smaller) new market entrants that are not able to meet these benchmarks due to the lack of economies of scale. This would infringe the fundamental freedom to provide services in Article 56 of the Treaty on the Functioning of the European Union (TFEU). Second, already established providers may decide to leave the market segment of retail investments which would further increase market concentration.

Finally, these IRSG members find that consumers information is important but should remain separate from the value for money work.

Q5: Stakeholders' views on the approach to product clustering are sought.

We find that the variety of Insurance-based Investment Products (IBIPs) in Europe and the differences between national markets (eg, different rules and investment portfolios) makes categorisation difficult. On the one hand, all the features and services that products can offer should be considered; on the other hand, having too many clusters or too few products for each cluster would make the exercise impossible. This questions the feasibility and usefulness of clustering at a European level. In any case, to avoid biased conclusions, a sufficiently detailed granularity must be ensured, even if the clusters contain a small number of types of contracts managed in a small, even unique number of Member-states.

Also simplicity is important in product clustering. Although certain unit linked products are pension products and therefore can offer tax benefits, at least locally. These products are partly based on the same principles as other unit-linked products but also may have different features for pre- and post-retirement life stage.

We believe that simplicity should be a key principle in establishing clusters. We would also like to draw EIOPA's attention to inherent challenges in the boundary of product clustering. For example, is there a "bright line" separating "significant" and "not significant" biometric coverage? The customer preferences and circumstances would determine the importance of this product feature. Arbitrary clustering could lead to products that could plausibly lie in multiple clusters, leading to inappropriate conclusions being made, or gaming the system. Furthermore, challenges are expected with respect to flexible products that fit into several clusters since they often offer inter alia different holding periods, both single and regular premiums, different guarantee levels, for which a solution is not yet envisaged.

Finally, as products offered in EU differ in many subtle ways, simplicity however important it is, may just be impossible to achieve in a pan-European level. It is important to recognize that supervision of specific cases is best done through bespoke monitoring and supervisory dialogue rather that with benchmarking against average cases that will fall short and far of unique matters. One possibility we see is to allow benchmarks made at national level rather than at EU level (to ensure better comparability). The benefit of this would be (i) taking away material product differences which arise because of different country features and (ii) putting the use of benchmarks in the hands of those that are relatively familiar with the underlying product structures and country specific features. As the market offering in EU is still very much national, it is not very informative for a policyholder to try to come up with findings on whether the product in question gives value or not, compared to some benchmarks.

Q6: Do you agree with the essential and additional criteria for product clustering? Should additional criteria be collected?

The IRSG is of the opinion that all product benefits should be considered, as well as the variety of products in different markets. It is not clear how EIOPA would make use of the "additional criteria" but treating them as secondary can mislead the assessment. Balancing the inclusion of various features with the need to keep the exercise simple will be extremely challenging. The additional benefits should be given same importance as the product could have been chosen for these specific additional features.

We would question whether it is of value for the first phase of consideration of value to look at all investment options and risk levels within a multi-option product. For some members, comparison of charges for a standardised risk level would provide insight to relative value from different products. Outlier products identified through this high-level analysis could then be examined in more detail including in relation to funds with different risk levels. Some market monitoring would then be required as an overlay to ensure that there were no efforts to subvert the system, e.g. by staying close to benchmark at standard level but with inflated charges at different risk levels. Also, recommended holding period for retirement- based products should also include term to expected retirement. Where biometric benefits are offered alongside investment benefits, the approach to benchmarks should be as far as possible to disaggregate the various benefits. This will lead to the most effective insights to value – benchmarks which seek to make sense of bundled benefits are likely to lead to confusion.

We believe that benchmarks only capture certain product features via quantitative indicators. Many of the core features like quality, insurance cover and other safety features (e.g. risk reduction), cannot be accounted for by quantitative indicators. Therefore, it is basically impossible to create a holistic view of insurance products solely by using benchmarks. In our understanding, benchmarks can be a useful tool for experts to give an orientation on certain quantitative measures. Anyway, even for experts, it is important to understand the indicators used and their shortcomings very well. For instance, any indicator based on performance (e.g. yield, surrender values, etc.) has limitations. It either depicts the past, which will not be repeated in the future or it is model based, which means results will vary depending on the model and parameters chosen and remain based on the past too. Furthermore, on quantitative indicators, the biometric risk coverage seems to be complex. We do not believe that a contract with significant biometric or provident cover can be reduced to a statistical analysis based on indicators. Also, it is not self-evident how a biometric risk cover, e.g. mortality or morbidity, should be interpreted in terms of ‘value’ or ‘non-value’. Furthermore, advice should also qualify as an essential factor:

- A product which is designed for distribution by fee-based advice does not contain distribution costs. All indicators will look preferable to ones of products where advice is included. However, in reality the fees for advice can be rather high, such that the product would not necessarily appear preferable if the fees were contained in the indicators.
- A product without advice will also appear preferable in many indicators. But of course, there is a significant difference in services provided.

From the French experience on product clustering, some members would like to point out that for the “Essential features”:

- EIOPA uses only 3 intervals of SRI classes, the French up and running VfM tool has up to 6 depending on the asset class of the unit.

- The premium frequency clustering does not match the reality: it is not only single vs regular, but also not single but not regular, this omitted option being probably the most used in the French market for example.
- Death coverage: low/high does not cover the option of no coverage at all, which seems relevant for the French UL market for example.

And for "Additional features":

- "Asset type: equity, bonds, MMFs, AIFs, other". For the French VfM this is an essential feature, and rightly so, given its importance on the performance of the product. Also, the French VfM includes also "mixed assets", which is very widely used, and uses mostly the ECB fund categorisation. And "AIF" is not an asset class but a legal category of investment funds alongside UCITS. Maybe EIOPA means hedge funds? The French VFM also includes real estate units.
- Level of guarantee: the 100% level is missing, although the most widely used in France.

Q7: Do you agree with the proposed approach to use the additional criteria to either develop more detailed clusters or to provide qualitative considerations on how to take these elements into account when looking at the benchmarks?

We find that dividing features into "essential" and "additional" overlooks the importance of a holistic assessment of all the elements from which consumers derive value. Including more features complicates evaluations, but qualitative features like the level of advice or services cannot be overlooked. This might not be in the first level of analysis but should be considered in the overall process. Finding a balance between comprehensive assessment and manageability is tricky, questioning the usefulness of clustering at a European level.

While quantitative data provides a foundation, qualitative insights can add depth and context to the evaluation process. Qualitative considerations are particularly valuable in assessing aspects that may not be easily quantifiable but are essential for a comprehensive understanding.

Q8: Do stakeholders think that for MOPs Option 1 would suffice or that Option 2, which would be more substantial in terms of reporting but also more precise and granular, should be preferred?

We find that multi-Option Products (MOPs) are the preferred solution of many consumers for their adaptability and variety. However, VfM benchmarks for MOPs are particularly challenging as MOPs can be built in different ways and can allow consumers to choose among many different combinations of funds. EIOPA's proposals for MOPs benchmarking do not seem workable, as they either oversimplify or demand too much data, misrepresenting MOPs' qualities and making their offering complex.

Between the two suggested options, the members find problems from both. On the one hand, option 1 could be more appropriate than Option 2, even though option 2 offers more realistic level of clustering. On option 1, if only considering one standardised investment risk level in performing a comparison for MOP's, this would give insight into products which were out of line based on the standardised risk level and further analysis of those products would be required, including into different risk levels. On the other hand, option 1 bears the risk to be oversimplistic and would lead to misleading comparisons. Option 2 is complex, but its greater granularity could make it more adapted, flexible and pragmatic in its implementation.

Another way of looking the options is that option 1 could be modified and a more representative fund selection should be used. Here using the system from the costs and past performance report where the most sold options are used might be helpful. Taking the most expensive, the cheapest and the average option might use rather exotic funds as representatives. Option 1 with the most common options put the focus on the product itself. The available options which are most often funds could be evaluated separately. Option 2 puts more emphasis on the available options. This could be preferable in some markets where the emphasis is on the underlying investment. It could however distract from the product features if the options only play a minor part, e.g. if only the profits are invested in funds.

For some members, if the clustering were the same for all unit linked products, both options could be used as alternatives based on the specificities of the national markets. As the vast majority of unit linked products are MOPs, the clustering should be the same for all products. Some clustering factors described for option 2 should be used for both options. Especially more granularity on the risk class and where possible also the type of investment (not always clear with hybrid products).

We find that both option 1 and 2 seems to ignore the reality of the MOPs markets at least in France, Germany and Belgium where MOPs often include hundreds of units.

Some of the members would like to point out that the French VfM Supervisory tool uses quite effectively a third option: clustering by unit. This method should also allow including the wrapper cost as one cost feature.

Q9: For Option 2 do you think the clustering approach should be revised by focusing more on the underlying options and less on some of the other essential product features?

We see a special challenge in incorporating investment performance for benchmark of products in a meaningful way. Extreme care should be exercised when resorting to investment performance for several reasons. Firstly, depending on time boundaries very significant contrasted outcome may be revealed. Secondly, depending on idiosyncratic choices and objectives, the performance of investments may differ significantly, which, in itself is not good or bad. For instance, social housing programs do not provide the most rewarding financial returns, together with entrepreneurship securities or securities in innovative sectors, even in relation with sustainability goals. Thirdly, the asset mix plays a determining role and lastly, it should be kept in mind that markets have volatile behaviors for all sorts of reasons that may hit at different times and with variable correlations, market timing may be decisive. Additionally, insurers and consumers may have views and intentions of their own that cannot and should not be judged against theoretical or average benchmarks.

We also believe that for hybrid products the products features but also important indicators like risk and performance can differ quite significantly from those of the underlying. For POG purposes it is important to have a holistic view of the product.

Finally, we believe that the proposed benchmarking options are unrealistic and do not match market realities. IBIPs offer diversification and unique asset access, which is not captured by comparing single funds. At the same time, it would be impossible to assess all combinations of funds that MOPs can offer.

Q10: For Option 2 do you think that the inclusion of the profit participation investment option in the asset class feature is appropriate for a correct interpretation of hybrid products?

Not in general, as there are different types of hybrid products. There are types of hybrid products where the customers can treat the with-profit fund like an option. But there are others where the with-profit fund is used as part of the risk mitigation and customers do not have a choice regarding the with-profit fund. In these cases it would be misleading to present the with-profit fund as an option.

We find that Hybrid products are a good example bringing value in many ways and especially via the the guaranteed and non-guaranteed component. Here it seems inappropriate to assess parts of the investment separately. Some of the member would point out the example from France which was offered in earlier questions on ways to deal with this issue.

Q11: Stakeholders are invited to provide feedback on the use of VfM Methodology Level II indicators, are these a good fit for the benchmarks? Should Level I indicators be used?

Some of the IRSG members has the opinion that the PRIIPs KID indicators must be avoided due to their very poor quality and misleading nature:

- There are no actual and long-term performances in the KID, and no actual performance of the benchmark chosen by the manufacturer of the unit. Only future performance scenarios based on the last five years actual performances, leading to highly incorrect and misleading data. Therefore, the IRR PRIIPs methodology is to be avoided.
- Similar problems for costs: there are no actual costs in the PRIIPs KID, only the "reduction in yield " estimated on one of the four future performance scenarios, and only for the specific recommended holding period of the unit's manufacturer.

They see that EIOPA should use (like the French VfM) the actual historical performances and total costs of the units. For the units themselves, the historical performances of the units and of their benchmarks are already available and disclosed in a standardized format by all EU unit managers in compliance with EU Law (Delegated regulation (EU) 2021/2268 of 6 September 2021 amending the Delegated Regulation (EU) 2017/653 on the PRIIPs KID).

- The ratio cost/premium is not so relevant as cost/ total value of the product's unit, as often the biggest costs are trailing and based on assets, not on premiums.

Other members of the IRSG believes that the POG framework should remain flexible and adaptable, allowing product manufacturers to freely design products and set prices. EIOPA's

Methodology Level 1 indicators like the Key Information Document (KID) and Solvency II elements are enough to perform a first market screening and spot potential anomalies that require further investigations.

Moreover, the same members find that Layer I refers to a market wide assessment to identify relevant outliers, undertakings or products - based on a trade-off between costs and returns stemming from PRIIPS KID data, national reporting, SII retail risk indicators etc - that are then subject to higher supervisory scrutiny under layer II. In other words, layer II was meant as an additional step for products that somehow stood out in the essential step driven by layer I. Using benchmarks on all products only with layer II indicators means submitting all products to layer II observation and rendering layer I obsolete. We would therefore refrain from using layer II indicators without the originally intended perspective with layer I.

Some of the members believe that the VfM Methodology Level II indicators provide a more comprehensive and nuanced understanding of value for money, capturing a broader range of factors that contribute to the overall assessment. These indicators, which often include qualitative aspects and long-term considerations, allow for a more robust analysis of the value proposition. Level II indicators are a good fit for benchmarks, especially when aiming for a holistic evaluation of products or services. The inclusion of these indicators enhances the depth and accuracy of benchmarking processes, providing consumers and stakeholders with a more informed basis for decision-making. While Level II indicators offer a more detailed analysis, we recognize that Level I indicators, focusing on more straightforward and easily quantifiable metrics, can play a complementary role. Depending on the context and the specific benchmarks' objectives, a combination of both Level I and Level II indicators may be beneficial. The choice between Level I and Level II indicators should be flexible and context-sensitive. Different sectors and markets may require different approaches based on their unique characteristics. A flexible framework allows for adaptation to diverse consumer needs and industry-specific considerations.

Q12: Stakeholders' views on the proposed indicators are sought, including on the intervals at which the indicators need to be assessed.

Some of the IRSG members find that the VfM assessment for insurance products should occur at the end of their intended term, reflecting their long-term investment nature. Early evaluations could overlook the benefits of sustained investment and misjudge the efficiency of products. Should these indicators be introduced, we support their re-assessment after a reasonable period on time, in order to allow enough time to the industry to adapt to this exercise.

Some of the members would like to point out that where entry costs are paid at policy subscription and they significantly impact the amount that will be invested into the unit-linked funds, it is important to understand their relevance with respect to the total costs. This becomes even more important in case of business models where most of charges are being taken upfront reducing substantial performance on the gross premiums invested.

Finally, some of the members would also emphasize that it is not appropriate to consider vfm only over the “intended term” or RHP. This would allow customers surrendering policies at different terms earlier or later) to be disadvantaged. Those customers also have a right to be protected - often, although there may be a RHP, customers will have different time horizons at outset and of course life events may change their perspective as their policies progress.

Q13: Stakeholders are invited to also provide feedback as to which indicators works best

We believe that initial screenings should use only EIOPA’s Methodology Level 1 indicators, with deeper analysis reserved for selected cases post-screening.

Q14: Do you believe additional indicators should be measured?

No, we see the POG process aims at ensuring that the product is consistent with the target markets’ needs. To achieve this, POG needs to remain flexible and product manufacturers are best placed to determine how to perform the product testing. In case there was any additional indicators, we would also call for inclusion of additional aspects, for example, safety protection, quality, service level of advice, and product innovation to mention a few as additional qualitative factors when assessing the products. We emphasize that these cannot be benchmarked and are rather aimed to be considered when taking a holistic view at the value for money of the products.

Q15: In case option 2 for MOP is chosen, do you think that more appropriate indicators applicable only to the single investment options should be identified?

We believe that both options for MOPs are unfeasible, and there is a need to maintain the POG framework flexible and proportionate. Manufacturers should not be constrained by fixed indicators in their product design and should not be subject to increased reporting burdens.

Q16: Do you agree with the proposal of using PRIIPs KID assumptions for the moderate scenario for the calculations of the indicators? Should an additional scenario (point in time) being included to evaluate the current performance of the product?

Some of the IRSG members believes that for initial screenings, the PRIIPs KID could be sufficient, identifying areas for further analysis. Moreover, it would make sense to be consistent with assumptions shown in the KID document, to promote simplicity and consistency. Only looking at the moderate scenario means that only average outcomes will be observed. It might also be relevant to look at outcomes that represent the risk of the product.

Other IRSG members don't agree to this and would refer to the points raised in Q11 on the reasons why PRIIPs KID indicators must be avoided.

Some of the members would also like point out that while moderate scenarios provide a baseline for analysis, an additional scenario that captures real-time data can offer insights into how the product has performed under current market conditions. This can be particularly valuable for stakeholders seeking a more dynamic and up-to-date assessment. It is essential to consider the dynamic nature of financial markets. Including an additional scenario that reflects the product's current performance allows for a more nuanced evaluation, considering any recent market developments or fluctuations that may impact the product's standing.

Q17: Do stakeholders agree to use percentiles to define benchmarks?

Some of the members agree to this as they find that this clearly defines the position of the product compared to others.

Other members find that the goal of POG is to ensure that the product is consistent with the target markets' needs. Product manufacturers should not be required to adhere to benchmark percentiles, nor to prioritise cost over quality. Also, while the use of percentiles could be one valid approach to define benchmarks, it has to be done in a meaningful way to deliver the results aimed at by EIOPA. Even in markets without any problematic products the percentiles would be populated. As such, no automatism should exist in flagging products as problematic. Percentiles may not work depending on the price distribution, e.g., when all products are very close in price, or when there is a single outlier. They should be seen as only one solution as might not always be the best for all markets (e.g. measuring the distance to the average might be better in some cases). Also, it is difficult to set the percentiles in advance.

Q18: Do stakeholders agree that percentiles should be defined once the data is available and that such percentiles should be adjusted as relevant?

See answer to Q17.

In addition to that, the IRSG members believes that one should avoid defining the outliers levels (e.g. 20%) exclusively based on collected data otherwise there will always be a pull towards the mean and it would not be meaningful. We suggest defining the outliers after examining collected data and based on what is making sense economically as a matter of principle. We see merit in raising a deep understanding of the market (including the peripheral areas) by means of the envisaged data collection pilot, prior to defining values. We assume that on basis of a good market oversight it will be easier to determine what an outlier is and what a "normal" product with comparable high indicator values is.

It should be clear, that there might be clusters without any percentiles determined as benchmark. It is likely that there will be specific homogeneous markets or clusters without any outliers. In those cases, there is no need to determine any percentiles.

Q19: In stakeholders' views are there some minimum/maximum percentiles which should be used?

See answer to Q17.

In addition to that, we find that an outlier should be defined based on common sense i.e. based on what makes sense economically. There must not be a minimum/maximum percentile-benchmark. In homogeneous cluster without any obvious outliers, there should not be a percentile otherwise, benchmarks would not serve to identify outliers. This would lead to successively stricter cost-control, which is not a justified purpose of benchmarks.

Q20: Do stakeholders think that the data collection should be expanded?

No, in our opinion any additional reporting should be avoided, as some data already exists via current regulatory channels. Increased reporting requirements would be against President Ursula von der Leyen's commitment to reduce reporting efforts by 25%.

Q21: If yes, which data collection principles should be used?

See previous responses.

Q22: Do stakeholders foresee a significant impact in the data collection in terms of resources and time in comparison to the current Cost and Past Performance data collection?

We believe that there is a concern over the need for more resources, time, and system overhauls if new types of data are collected as part of Costs & Past Performance reports. EIOPA also refers to a pilot survey for MOPs, whose level of complexity is still unclear.

Q23: How would you assess the impact that the benchmarks methodology would have in your organisation? Please consider both the data collection and the use of the benchmarks when they will be available.

Although the full impact of the VfM benchmarks is yet unknown, our understanding is that the extra reporting load and bureaucratic increase are concerning implications.

Q24: Do stakeholders agree with benefits of the proposed approach?

Some of the IRSG member find that YES, especially if the benchmarks are made public and then it will definitely help consumers understand the impact of charges and fees being imposed in the products by comparing to others in the market and also across borders.

they find it important that benchmarks will also help compare products like for like not just on cost and performance.

Other IRSG members are of the opinion that NO, as the potential consequences of benchmarks on market innovation and competition, as well as the increased industry burdens, are concerning. The limited success of pan-European personal pension products (PEPPs) shows that over-regulating products and setting cost caps do not work. They feel that benchmarks are just one tool for supervisors out of a bigger toolbox which is necessary to gain a holistic view on the VfM of a product. Benchmarks on their own have a rather limited value as they need to be put into context by experts. Moreover, while agreeing in principle with the rationale of the proposed approach outlined by EIOPA, one observes that many elements of the methodology are still subject to discussion and do not allow for a meaningful impact assessment at this stage. Furthermore, these members remain concerned about the impact of the ongoing discussion on VfM under current IDD provisions on future regulation under the Retail Investment Strategy (RIS). While EIOPA highlights that current work is to be considered independent from the RIS and fully entrenched and based on existing IDD requirements (paragraph 1.7 of the consultation paper), section 9 of the consultation paper aims to bring the two discussions into perspective. The present input refers to benchmarks as a tool for supervisory purposes to inform a more risk-based approach under current IDD. As a result, one should refrain from using the current consultation out of the due context to inform the discussion of future regulation in RIS, under a very different use of benchmarks and wish to safeguard the relevance of the present contribution in the context of the current EIOPA consultation. These members also believe that the approach taken to possible implementation will be important in driving benefits. A system which is too complex or which somehow constrains value and innovation in product design, e.g. through being too focused on particular benchmarks, would be likely not to lead to increased consumer value, which must be the key aim here.

Q25: Are there additional benefits in stakeholders' views?

We find that if there is a benchmark that leads to the selecting of insurance products in the advice process based solely on costs and performance it will neglect many of the core values of insurance. This would indicate for instance that a cheaper savings product with a higher risk profile is better than one with more features and limited risk. Usually, riskier products have a higher performance potential than safer products. However, the risk can always materialize in which case the performance is potentially much worse than that of a safer product. Customers obviously need to understand the risk profile of the product but also that an underperformance against some benchmark doesn't necessarily imply bad value for money. A product can always be made cheaper by reducing the quality of the product, the service or the advice, but many times this is against the customer need. Ranking products in public benchmarks will easily create a danger of misleading impression and communication. We believe that EIOPA's VfM benchmarks bring higher reporting demands, complexity, product standardisation and reduced innovation, while also bypassing the political negotiations surrounding the Retail Investment Strategy (RIS). Therefore, they would lead to severe drawbacks for both the market and consumers, rather than actual advantages.

As an observation, for some members currently there seems to be a low take-up of unit linked products and where this is the case, these benchmarks could bring in trust and confidence and hopefully more investment in the capital markets.

Q26: What could be the costs of implementing Option 2?

We believe that the lack of sufficient details makes it challenging to estimate the cost and no impact assessment is provided by EIOPA. Nonetheless, the European Commission (EC)'s impact assessment on RIS indicates elevated costs for supervisory reporting and this happens whether Option 1 or 2 would be chosen.

As there is no general answer to this the costs would in some markets be a lot higher than for option one. It is not always easy to automatize all the required calculations. So, for products with many options the costs are multiplied by the number of funds.

In case Option 2 is implemented, it is already clear that a lot of complexity will result at two levels of data at single fund level and at product level. EIOPA should focus on strategies with easily available data to automatize the information flow and prevent unnecessary efforts chiefly at single fund level, using the reference data to the extent possible.