EVALUATING THE POSSIBILITY OF USING CONSENSUS METAL PRICE FORECASTS IN THE NATURAL RESOURCE INDUSTRY

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor of Business Administration by

John Chris Lamprecht

Key Words: collective, consensus, cooperative inquiry, crowdsource, decision-making, forecast, metal price

April 11, 2023

ABSTRACT

The action research project, structured as a cooperative inquiry, examined whether the reliability of metal price forecasting could be improved using a collective decision-making approach. Forecasting metal prices is a crucial practice for natural resource organizations that can impact their sustainability. The usual practice is for an organization to make metal price forecasts in isolation without the certainty that their predictions will be reliable. Prior research has advocated using a collective approach for determining a consensus outcome that captures the diversity of insights. As a collective decision-making approach, the Delphi Method uses an anonymous deliberation process, and usually, the consensus average is communicated through a moderator to surface a consensus over several cycles. Based on the Delphi Method, a cooperative inquiry as a case study approach was used to collect the metal price forecasts and supporting justifications from invited participants on a panel website that allowed asynchronous interactions. Secondary gold price forecasts were sourced from London Bullion Market Association's annual precious metals survey between 2000 and 2020 to triangulate with the panel data collected. Collective decision-making can be divided into two elements, participation and aggregation. Participation deals with who is involved and how they interact to make the collective decision. Two generic methods can be identified, deliberation and collaboration. Deliberation combines participants' views to determine a consensus outcome that reflects the majority preference. Collaboration is more encompassing, aiming to share participants' insights and surface a consensus outcome that incorporates the prior insights and those that emerge from the participants' interactions. Aggregation, the second element in collective decision-making, concerns how the collective consensus outcome is determined. The common practice is to use the statistical average or median as representative of the majority view. As seen in prediction markets, an alternative aggregation method is the market price mechanism to determine the outcome. All the metal price forecasts sourced were analyzed quantitatively to evaluate the reliability of an individual participant compared to the consensus median. The forecasters' justifications from the two data sources were qualitatively analyzed to examine if keywords could be used to identify reliable forecasts. The outcome of the quantitative analysis was neither an individual nor the consensus median was consistently a reliable forecast. Depending on the market metal price direction, for declining metal prices,

the lower quartile, and rising metal prices, the upper quartile represented a more reliable aggregated consensus estimate. The qualitative analysis from both forecasting forums showed a poor relationship with forecasting reliability. The keyword usage among forecasters was high, indicating the use of common justification for the forecast made. In evaluating the outcome reliability of the cooperative inquiry, the possibility of structuring the forecaster's inputs metal price and keywords as probabilities were proposed. Given the importance of correctly predicting the future metal price direction, it is recommended to focus the explanations provided on the anticipated market price direction. Reciprocity was identified as an essential motivation for obtaining participation and continued contributions from participants and is a factor to consider in future related research.

Statement Of Academic Integrity

I hereby declare and confirm with my signature that this doctoral thesis is exclusively the result of my own autonomous and honest work based on my research and the literature published, as seen in the bibliography.

I also declare that no part of the paper submitted has been made in an inappropriate way, whether by plagiarizing or infringing on any third person's copyright.

Finally, I declare that no part of the paper submitted has been used for any other paper in another higher education institution, research institution, or educational institution.

Johannesburg and Liverpool, April 11, 2023

Acknowledgments

This thesis results from a Doctorate of Business Administration (DBA) program at the University of Liverpool, United Kingdom. As the culmination of the DBA program, I would like to acknowledge the role of the following people who contributed to my getting to the point of submitting my thesis:

- My family for supporting my decision to pursue the DBA and making the necessary sacrifices.
- The help to set up and get the Consensus Metals website functioning for the research project, Ringo Mashukano, Andrew Baker, and John Lamprecht (Jnr).
- The participants on the Consensus Metals panel who gave up their time to participate and shared their expertise freely with their fellow cohorts, allowing the research findings to be identified.
- Aelred Connelly and Edel Tully from the London Bullion Market Association (LBMA) in assisting with sourcing the annual LBMA precious metals surveys, despite working from home.
- Dr. Jim Hanly, my research supervisor, for his guidance in completing my thesis.
- The lecturers throughout the DBA presented the coursework and helped me understand the possibilities of conducting my research and the tools with which to do it.

TABLE OF CONTENTS

List Of Diagrams

List of Tables

Abbreviations

APE – Absolute Percentage Error

AR – Action Research

CM – Consensus Metals

LBMA – London Bullion Market Association (LBMA, 2020)

LBMA-PMS – LBMA Precious Metals Survey

LME – London Metal Exchange (LME, 2019)

Lower Quartile (Q1) – The first quartile is the middle value between the smallest number (minimum) and the median of the data set. It is also called the lower or 25th empirical quartile, as 25% of the data is below this point.

MSE – Mean Square Error

RelMAE – Relative Mean Absolute Error

RMSE – Root Mean Square Error

Theil U(II) – Theil U(II) Index

Theil Uc – Covariance element

Theil Um –Mean-variance or bias element

Theil Us – Variance element

Upper Quartile (Q3) – The third quartile is the middle value between the data set's median and the highest value (maximum). It is also called the upper or 75th empirical quartile, as 75% of the data lies below this point.

CHAPTER ONE - INTRODUCTION

1. Introduction

Natural resource organizations sell their mine production at prices linked to the market metal prices. The primary metal producers do not know in advance what the future market metal prices will be when they expect to make further metal sales. The future market metal prices will depend on the subsequent transactions of thousands of buyers, sellers (including the primary producers), traders, and speculators on the international commodity exchanges. Despite natural resource organizations' uncertainty about future metal prices, they must still make regular metal price forecasts for short-term operational decisions and longer-term strategic purposes. On occasion, the metal price forecasts made within the natural resource sector have not materialized as expected, resulting in significant adverse impairments for these organizations (PwC, 2021). Making metal price forecasts in the natural resource sector opens the possibility of improving the estimates' reliability by considering an alternative workplace approach that utilizes diverse opinions (Armstrong, 2001a; Batchelor and Dua, 1995). Are there wise metal price forecast gurus who know all, or is this a task better performed collectively by multiple experts?

When making judgmental forecasts dependent on a myriad of intervening influences, the forecaster cannot be sure of anticipating all the events that cause the outcome. A similar challenge of trying to foresee all unknown future events exists when forecasting metal prices. The information influencing future metal prices is emerging, as opposed to the historical and known information that underpins the current market metal prices (Fama, 1970). The information determinants of present versus future metal prices reflect a contrast of market participants' available information and expectations of prospective market participants about forthcoming events that are not certain to occur. The information dichotomy between the currently known and possible future events creates the opportunity to leverage various collective insights to more informed consensus metal price forecasts. Could including participants with diverse expectations potentially lead to more reliable metal price forecasts for natural resource organizations?

The theory underlying collective wisdom is premised on the concept that by accessing a diversity of insights from multiple participants, the consensus outcome can

exceed the reliability of most group members because the resulting aggregated estimate includes more insights (Armstrong, 2008; Batchelor and Dua, 1995; Hong and Page, 2012; Landemore, 2012a). Collective decision-making can be distilled into the fundamental elements of accessing and aggregating diverse insights. Accessing the participants' insights relates to who participates and to what extent the group members interact in determining the consensus outcome. Aggregating the participants' estimates refers to how involved the group members are in choosing the aggregation process and what constitutes the agreed aggregated consensus outcome. Each of these elements has alternative approaches, which may affect the reliability of consensus outcome, requiring consideration when structuring a collective decision-making approach. An implicit underpinning for collective wisdom is the expectation that the group members participating are "wise" (Andler, 2012; Flyvbjerg et al., 2012; Page, 2007; Rowe and Wright, 2001; Surowiecki, 2004; Tetlock and Gardner, 2016). What constitutes wisdom, who should participate, the nature of interactions between group members, and the aggregation method occurs in the collective wisdom literature (Andler, 2012; Armstrong, 2001a; Aristotle and Everson, 1988; Baker, 1976; Galton, 1907a; Hong and Page, 2012; Landemore, 2012a; Rowe and Wright, 2011; Sunstein, 2006).

Accessing the participants' views could range from passive to active and repetitive interactions during the collection phase of the group members' estimates. For passive interactions, deliberation would be an individual activity informed by the participants' current views. As the group interactions become more active, the prior opinions of group members could be altered by the insights shared during their joint deliberation. When the participants' interactions tend towards active and repetitive, the group members would likely aim through collaboration to achieve a consensus outcome mutually agreed upon by the majority of participants. Essential aspects concerning participants' interactions are the possibilities of undue influence by some group members, the inherent biases of participants, and participants' willingness to be informed by other group members' insights (Bazerman and Moore, 2008; Drummond, 2001). From a consensus metal price forecasting perspective, it opens the question of how significant the degree and timing of interactions are to the reliability of consensus metal price forecasts.

A process for aggregating the group members' views is required when combining the estimates of multiple participants to determine a consensus outcome. As with the

case of interactions, this could range from passive to active involvement of the group participants. In many instances of collective decision-making, the consensus outcomes are calculated as either the statistical average or median of participants' views, reflecting a passive involvement of the group members in the aggregation process. In a variation of the typical passive average and median aggregation approaches, the threshold for reaching a group consensus can be set at a level either below or above the central statistical measures to consider the views of minority stakeholders. The departure from the usual main statistical measures reflects a greater degree of involvement by the group members in deciding the consensus outcome, albeit calculated similarly to the statistical average or median. Under those circumstances, when the extent of interactions between group members is active and repetitive, the consensus outcome could be a negotiated compromise of their views, reflecting a collective result acceptable to most participants. From a consensus metal price forecasting perspective, it opens the question of the importance of allowing repetitive interactions to reach an agreed consensus outcome versus taking an aggregate statistical measure as the most reliable estimate of the group members' views.

The Action Research (AR) project underlying this thesis was structured as a case study investigating the aspects influencing collective decision-making identified from the relevant literature. Participants were invited to submit metal price forecasts over six months on a consensus forecasting panel, to investigate the viability of using a consensus approach for forecasting metal prices as an alternative workplace practice in the natural resource industry. The forecasting results of the consensus panel were triangulated with a similar analysis of the gold forecasts from the annual London Bullion Market Association precious metals survey (LBMA-PMS) over twenty years. The AR project's findings showed participants' low willingness to be influenced by other group members' views, possibly implying participants favored their insights when making metal price forecasts. As far as the aggregation of the participants' estimates was concerned, the findings of the action research project indicated either the lower or upper quartile were more reliable consensus predictions of expected metal prices, depending on the future direction of the market metal prices. As a pilot case study, the findings of the AR project are not generalizable, requiring further related consensus forecasting analysis and research to verify the reported results of this thesis. Anchoring bias may explain participants' preference to rely on personal views when

making decisions on metal price forecasts, perhaps indicating the solution, as presented in this thesis, is to focus on diversity through sourcing and combining multiple estimates. Traded markets are directional, down, neutral, or up, so the proposed alternative use of the lower or upper quartiles as more reliable consensus aggregation measures is logical. The issue is more when the central statistical measure will be the best.

1.1. **Statement of the Problem**

When market participants decide on an outcome, the possibility that any participant could have complete insight into the planned prior actions of all other market participants is improbable. The emergence of metal prices can be seen from a relativist ontological perspective, based on the ongoing market interactions representing a socialist constructionist epistemological perspective (Easterby-Smith et al., 2012). Commodity price forecasts are a crucial input variable in the planning and operation of natural resource organizations. Given the importance of metal price forecasts, mining organizations often source metal price estimates from third-party research organizations to have another perspective to supplement their internal projections.

Some organizational future metal price estimates combine other sourced estimates determined using alternative approaches, with the final forecast compiled by applying an aggregation method chosen by the organization. As the predictions are a culmination of the thought processes of several people, albeit with varying degrees of participation and sometimes aided by internal statistical analysis, the final derivation of the metal price forecast effectively reflects the perspectives of all the people involved aggregated into a consensus prediction. The use of aggregated consensus estimates is not an unusual practice in the mining industry, but seldom are such forecasts derived through open discussion and debate among industry experts supplying the prediction.

For organizations using multiple forecasts from either internal or external sources or both, the consensus metal price determined through the organization's decision-making process will have given different emphasis to the various sourced estimates, with the outcome a combination of all the views and insights of the industry experts considered. Despite the consensus metal price determination, the probability of specific organizational predictions proving repeatedly reliable is not insured,

particularly under high metal price volatility and for longer-term forecasts. Some organizations may be more capable than other market participants in forecasting metal prices. Still, such a positive trend is unlikely to be consistently maintained in an environment with many exogenous influencing factors.

The question arises, could the Consensus Metals (CM) panel participation be reduced without compromising the reliability of the estimate by avoiding the "uninformed crowd" and explicitly focusing on the unique insights of a select diverse group of industry experts (Page, 2007; Surowiecki, 2004)? Select a small group of industry experts who understand the metal market dynamics and encourage them to share their insights on a consensus survey panel to collectively surface metal price forecasts that are more reliable than their individual estimates. Focusing on industry experts with unique and divergent insights while fostering their interaction to instigate joint reflection and action on a consensus survey panel could surface more reliable forecasts for the participants' mutual benefit. Their collective actions and reflections combine their shared insights to reduce the probability surrounding uncertain future metal prices. Historical precedents of combined or consensus forecasts proving more reliable than individual estimates over repeated events have given rise to the research on the "*wisdom of crowds*" (Armstrong, 2001a; Batchelor and Dua, 1995; Hong and Page, 2012; Surowiecki, 2004). The Action Research (AR) project aimed to evaluate whether a similar collective approach could be used in the natural resource industry to improve the reliability of metal price forecasts influencing their organizations' decisions. Could a collective method for combining multiple estimates from several organizations yield a more reliable metal price forecast than the specific estimate of one organization?

1.2. **Workplace Context of the Problem**

As a business consultant in the natural resource sector, my prior experience included the usual industry tasks where reliance was placed on forecasted metal prices. Mining organizations make long-term investment decisions based on an expectation of future metal price levels. The future metal price levels impact the expected economically recoverable minerals, the investment pay-back period, and the net present value (NPV) of discounted future cash flows. Mining organizations also prepare annual budgets to plan debt repayments and provide markets with indications of expected earnings and dividends based on forecasted future metal prices. For those

industry experts working in the natural resource sector and involved in some or all the mentioned activities, it requires making or sourcing metal price forecasts as part of their workplace responsibilities. Given the importance of the sustainability of the mining organizations and the low likelihood of always correctly forecasting metal prices, which can be attested to from personal experience, it was considered a workplace activity that called for further research. The AR project aimed to evaluate the possibility of collectively making reliable metal price forecasts. Recognizing the limitations of a specific expert's insights could be overcome by tapping into the knowledge of multiple industry experts and pooling their expectations to surface consensus metal price forecasts for their collective benefit.

The motivation for considering a collective consensus approach arose from two events around the same time. I was advising a client on the sale of an interest in a joint venture mining project, with the primary value debate focused on the forecasted metal prices for the life of the mine. The buyer's advisors decided to use several industry metal price forecasts, using the median of the sourced forecasts as the most reliable estimate (Galton, 1908; Levy and Peart, 2002). It was impossible to argue the approach advised, as the practice was typical within the mining industry, and the forecasts of the two transacting parties were significantly different. The metal prices were recovering from a price slump but had not increased significantly.

During the DBA coursework, the topic of biases was discussed. In the literature reviewed, mention was made of experiments in which the average of a group would surpass that of the individual participants' accuracy (Bazerman and Moore, 2008; Drummond, 2001). Surowiecki (2004) cited the aggregation accuracy phenomenon in "*The Wisdom of Crowds: Why the many are smarter than the few and how*," leading me to question the relative accuracy of the metal price forecasts used in the mining deal. About a year after the sale, the metal prices exceeded the consensus median used in the final valuation. The actual outcome went counter to the expectation of the literature on consensus estimates. However, the forecasts were more accurate than those of both parties. The assumption was that the limited data of one transaction was not representative of the concept of consensus forecasts. If collecting metal price forecasts over several periods were repeated, the consensus median or average's reliability would be the most reliable, as predicted in the relevant literature.

The AR project was conceived to collect metal price forecasts from industry experts to test the assertion that consensus metal price forecasts can be more reliable over multiple forecasting periods, as it potentially pertains to use by mining organizations. The expectations were that the reliability of the collected predictions could be used to measure the reliability of the consensus average against the subsequent actual market metal prices. Conceptually, the idea for the AR project was simple, but the actual implementation proved to be more complex. For the AR project, certain aspects required consideration to ensure the outcome yielded verifiable results that could be replicated consistently as a sustainable business practice.

Frequently in metal price surveys, it is recognized that metal prices vary over time, so participants are asked to provide their minimum, maximum, and average metal price indications for the forecasting period. Some forecasters prefer to indicate ranges when making forecasts, as they base their estimates on the possibility of certain events occurring. The value of consensus measures, such as the average and the median, is associated with aggregating diverse opinions to achieve a consensus result that reflects collective insights. To address the preferences of the forecasters, and enhance the information collected, the possibility of using probability distributions needed to be considered instead of point estimates. A further consideration when collecting metal price forecasts was whether the rationale of the forecasters for their projections should be included in the AR project data collected. The literature supported the supposition that requiring explanations reduced the extent of the participants' biases. An aim in collecting metal price forecasts to aggregate and calculate either an average or median was to ensure diversity of opinions to limit the possibility of groupthink (Janis, 1973). For the forecasting process to include diverse participants, it needed to allow for asynchronous forecasting, allowing multiple forecasters to take part at a convenient time.

In the final structuring of the AR project, the issues of how the metal prices forecast collected would be structured, either as point estimates or as probabilities distributions, and whether required justifications for the forecasts submitted needed to be decided. The structural elements had to be considered as a balance between imposing on the participants and collecting data that could lead to reliable metal price forecasts, recognizing the research aimed to find an alternative workplace practice. It was also hoped the results from the AR project could be used within my consulting practice to

provide a better service to existing clients and attract some new clients through the reliable metal price forecasts the approach could yield.

1.3. **Action Research Approach for Solving the Problem**

The AR project evaluated the possibility of sourcing aggregated consensus metal price estimates that could be used in the workplace as reliable forecasts, either in conjunction with or as an alternative to existing practices (Coghlan, 2007; Coghlan and Brannick, 2014). Based on my prior experience of making metal price forecasts for use by clients within my consultancy practice or sourcing consensus metal price estimates from third parties for specific projects, the reliability of the forecasts has proven concerning. The metal price forecasts I had made alone were prone to my personal biases, explaining their inconsistent reliability (Bazerman and Moore, 2008; Drummond, 2001). Conversely, the third-party consensus metal price forecasts sourced should not have been as influenced by bias, and through aggregating the individual estimates, the combined consensus median or average should have been reliable (Armstrong, 2001a; Galton, 1907a; Hong and Page, 2012; Levy and Peart, 2002; Surowiecki, 2004). Somewhere between the theory of aggregation reliability and the actual outcome of the practice, some elements were not ensuring the approach's effectiveness.

Consider that some of a consensus approach's elements can influence the estimates' reliability. In that case, it requires identifying the main structural aspects to consider their possible impact. Two primary contributing factors are seen as the most significant in the reliability of consensus estimates: the forecast's quantification and the rationale for the predictions.

In a forecasting competition, there are going to be winners and losers. The implication is that not all forecasters are equally capable, and the outcome will depend on the most competent experts. There are three factors underlying the assumption of individual superiority. The most obvious is the requirement that the best expert is a group member making a forecast. Secondly, the expert's opinion will be the deciding forecast. Thirdly, the other group members' forecasts will be dispersed and linked to the preceding two issues. The positive and negative forecasting errors offset each other, leaving the expert's forecast as the outcome. The presumption of individual ability exceeding the joint capability of several group members when making

judgmental metal price forecasts contradicts the prevalence of consensus forecasts in practice provided by institutions such as Consensus Economics (2020) and Thomson Reuters GFMS (2020). A consensus aggregation approach opens the possibility that the forecasts of the less accurate group members are not evenly distributed around the winning expert's prediction. An alternative possibility is that the most reliable forecast does not come from the same expert on all occasions. Suppose the forecasting errors are not evenly distributed, and the most reliable forecaster is uncertain. In that case, it opens the possibility that a consensus approach might be more reliable than an individual expert.

When identifying an expert in making metal price forecasts, there is an assumption the individual has unique insights that allow them to make reliable forecasts. It presumes the individual can communicate the insights to others, allowing them to appreciate the wisdom of the expert. Ascribing unique insights to experts assumes that their reliable forecasts were not just lucky guesses but that they have specific knowledge linked to their capabilities. The expert's unique insights would allow them to make better-informed judgments than other forecasters, reflecting Hayek's (1945) idea of market efficiency based on special circumstantial knowledge. An assumption of all-knowing is disproven by markets where successful traders do not dominate over unsuccessful traders continually, reflecting that experts cannot always have complete knowledge. The implication is that experts have unique insights sometimes that make them more reliable forecasters than others on occasion, but not always. Experts can both inform the actions of others by communicating their insights and learning from the wisdom of others to inform their actions. The outcome of all participants' actions could be better informed through interaction.

Consensus metal price forecasts are available in practice but are not as widely used as expected. After researching the issues that may be limiting the reliability of consensus forecasts, some aspects could be addressed to improve their reliability. How do you collect, aggregate, and distribute consensus metal price estimates among the contributing participants, based on accessing and sharing their combined expertise, to improve the reliability of the metal price forecasts available for use in their workplace? The challenge is structuring the AR project to quantify the metal price forecasts of several industry experts while simultaneously collecting and sharing their expertise to inform other participants' contributions, resulting in a reliable consensus estimate.

Fundamental to the AR project's aim of making reliable consensus metal price forecasts available to industry experts is the understanding that the approach is based on their collective contributions. The AR project recognizes that each participant contributes to the consensus metal price estimate and is a beneficiary by being able to use the outcome in their workplace. In making their contributions to surface a reliable consensus metal price estimate, the participants each have an equal status relative to their cohorts. The AR project aims to find a metal price forecast representing a workable alternative to their workplace practices. Essential to the aims of the AR project is the recognition that no industry expert can have complete insight into all the influencing factors affecting future metal prices. It is through observing the wisdom of other contributors that it can inform their subsequent actions. Based on the insights gained from the group interactions, the cycle of forecasting, observing, reflecting, and forecasting again is limited to the participant's choice.

Considering the aims of the AR project, the purpose of the participants' interactions, and how they interacted, it posited the AR project within the modality of being a cooperative inquiry. The AR project was undertaken "*with people*" experiencing a similar challenge to find a workable alternative "*for the people*" involved to apply in their workplace in the future (Heron and Reason, 2008).

- The participants on the CM panel had a shared requirement to identify a workable alternative metal price forecast.
- Each participant on the CM panel was an equal co-subject in making forecasts and being able to observe, reflect, and make subsequent forecasts based on their experiences on the CM panel.
- Despite the intended aim of the AR project of identifying a reliable alternative consensus metal price forecast for use in practice, no assurance was given or taken by the participants that the outcome would meet their expectations by participating on the CM panel.

The structure of the CM panel looked to investigate the two aspects of using consensus metal price estimates in the workplace. None of the participating experts on the CM panel were considered to be the sole expert. The participants on the CM panel could observe, reflect, and act based on the actions of their cohorts in repeated cycles as they chose (Coghlan, 2007; Coghlan and Brannick, 2014). The CM panel structure aimed to facilitate active collaboration between the forecasting participants.

Regardless of the level of collaboration observed amongst the CM panel participants, the AR project looked to investigate the reliability of aggregated consensus metal price estimates. The AR project focused on identifying a feasible alternative workplace practice without excluding approaches that did not fit an expected level of collaboration.

1.4. **Overview of Thesis**

This thesis has five chapters, as follows:

- Chapter One (Introduction) outlines the research area, the motivation underlying the AR project, and an overview of how the AR project was undertaken.
- Chapter Two (Literature Review) discusses the theoretical foundations of collective decision-making which underpins the AR project and proposes a classification framework for assessing the nature of consensus decision-making.
- Chapter Three (Research Methodology) describes the epistemological and ontological assumptions framing the AR project and explains the AR project as a cooperative inquiry within the AR arena. The research design, methodology, and data collection processes are described.
- Chapter Four (Results & Findings) discusses the quantitative and qualitative data analysis, the research findings, and the conclusions drawn from the AR project's results.
- Chapter Five (Conclusion) concludes the thesis with a review of the AR project and findings, discusses the actionable knowledge learned from the AR project, and considers possible further actions that could be taken based on the AR project.

CHAPTER TWO – LITERATURE REVIEW

2. Literature Review

2.1. Introduction

The theory underlying collective decision-making dates to Aristotle's 4th century BC summation argument that "*many are better judges than a single man"* for *"among them they understand the whole*" (Aristotle and Everson, 1988, pp.66). The phenomenon of considering the opinions of many participants to determine a consensus estimate is observable in the modern-day practices of numerous institutions where the outcome relies on the views expressed. For example, reporting expected results for political elections by polling institutions such as Rasmussen (2021) that combine the opinions of many participants who had limited or no direct shared interactions. Similarly, the consensus earnings forecasts for organizations are widely reported in the financial markets by news services such as Thomson Reuters GFMS (2020), reflecting the opinions of multiple analysts who acted independently. A question that underlies the use of consensus results is whether the outcome could be improved through the prior deliberation by the participants of their insights. Alternatively, can the practice of aggregating the independent opinions of participants without them interacting yield a reliable consensus estimate?

In the first decade of the 21st century, collective decision-making was popularized by the publication of books such as:

- *"Smart Mobs: The Next Social Revolution" (*Rheingold, 2002).
- "*The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*" (Surowiecki, 2004).
- "*Deliberation Day" (*Ackerman and Fishkin, 2005).
- "*Infotopia How Many Minds Produce Knowledge*" (Sunstein, 2006).
- "*The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*" (Page, 2007).

The phrase "*the wisdom of crowds*" became more commonplace following the publication of Surowiecki's book and started to appear in the academic literature concerning stock market returns (Chen et al., 2013; Da and Huang, 2018; Ready-

Campbell and Hill, 2011; Lee et al., 2016; Hong et al., 2016). In 2006, Howe coined the phrase "*crowdsourcing*," a portmanteau of crowd and outsourcing, to reflect the growing significance of online collaboration for solving problems (Brabham, 2013). Prediction markets based on the market pricing mechanism for forecasted future events came to the fore through forums such as the Iowa Electronic Market, established in 1998 (IEM, 2019). For 41 elections held in 13 countries between 1988 and 2000, the Iowa Electronic Market achieved a lower average absolute error rate (1.54%) than the average of the comparative election polls (1.91%) (Berg et al., 2008). Estimize (2020), an online forum for sourcing financial forecasts, was founded in 2011, building on the concept of accessing the "*wisdom of the crowd*" by using a crowdsourcing approach (Drogen and Jha, 2013). From the early collective decision-making discussions of Aristotle to the more recent "*wisdom of crowds*" and "*crowdsourcing*" practices, the underlying theory has been researched to explore further the elements of access to and aggregation of the participants' contributions and what constitutes "*wisdom*" in the context of determining a consensus outcome (Armstrong, 2008; Batchelor and Dua, 1995; Galton, 1907a; Landemore, 2012a; Page, 2007; Sunstein, 2006; Surowiecki, 2004; Tetlock and Gardner, 2016).

Participation by participants in a collective decision-making process is broadly referred to as access, relating to all the actions their participation allows and imposes on them as group members. Access includes, among other things, the following aspects:

- Who is allowed to participate in the collective decision-making process.
- How the group members interact to share their insights and jointly reflect on the group's interactions.
- The ability of participants to influence other group members.
- The nature of the relevant insights held by individual participants relative to the collective decision the group is aiming to make.
- The impact of a leadership role by one participant relative to the other group members participating in the collective decision-making process.

Access can range from limited to significant group cooperation for collectively achieving a consensus outcome. At the lower range of cooperation, participation is

more akin to a "*crowd*" survey than a group explicitly working together to achieve a result. For example, when casting votes in an election, the outcome reflects the collective decision of the group acting independently rather than a combined effort by the voters to cooperate and act in a specific coordinated manner. Typically, participation is open to many participants, and the group's interactions before the collective decision-making are limited, other than canvassing by a leader to gain the votes of group members (Baker, 1976; Galton, 1907b; Mueller, 2003). The deliberation about the collective decision under consideration is more prone to be a personal reflection of the individual participants about their preferences rather than based on their shared interactions. The collective decision-making with minimal shared deliberation and intra-group influences will reflect the summation of the participants' opinions while surfacing limited new information (Aristotle and Everson, 1988; Batchelor and Dua, 1995; Keyt and Miller, 1991; Waldron, 1995). Participation is not dependent on the "*wisdom*" the group members can contribute to the collective decision-making outcome; instead is determined by their right to participate (Baker, 1976). It is well suited to situations with few outcomes and usually in which the participating group members have a vested interest, such as the elections discussed by Condorcet (Baker, 1976; Mueller, 2003).

If the group members' actions are to deliberate their choices, the collective decision-making approaches with low levels of collaboration can more appropriately be seen as individual deliberation with summation (Aristotle and Everson, 1988; Hong and Page, 2012; Sunstein, 2006; Tetlock and Gardner, 2016). For the participants, it is more like sense-making of their choices by cross-referencing other views rather than surfacing new collective "*wisdom*" (Argyris, 1977; Weick, 1988). The consensus outcome represents the summation of dispersed opinions into a single collective group estimate, i.e., summation (Galton, 1907a; Keyt and Miller, 1991; Surowiecki, 2004; Waldron, 1995).

When the selection of group members participating is restricted to only those perceived as having some "*wisdom*" about the collective decisions to be made, the participating group size can be smaller (Armstrong, 2001a; Batchelor and Dua, 1995; Surowiecki, 2004; Tetlock and Gardner, 2016). With the reduced number of participants because of their specific expertise, the level of interaction between the group members can be expected to be more extensive, and their reflections on the

group members' shared insights more significant (Argyris, 1977; Coghlan, 2007; Page, 2007; Revans, 1998). Examples of situations that exhibit more significant levels of interaction and deliberation when making collective decisions are committees, such as a board of directors and juries (Rijshouwer, 2019; Sunstein, 2006; Surowiecki, 2004). Although the interaction between participants is more significant, the interactions of the group members are typically aimed at agreeing or disagreeing on a limited number of alternatives, e.g., invest or not, guilty or not guilty. The role of the group leader is to allow the group members to deliberate their shared insights and, if necessary, have a deciding vote in achieving a definitive consensus outcome. The group aims to deliberate the identified possible outcomes jointly rather than for the group members to collaborate in exploring new alternatives. The potential benefit of collective deliberation when making a decision was identified by Ackerman and Fishkin (2005) and Schkade et al. (2007) as an opportunity for improving group decisions. Despite the group members' more significant interactions, the participant's actions and the group's functioning are primarily to deliberate on specific alternatives and reach a consensus outcome based on a summation process predefined by the group or its convener.

In some collective decision-making situations, the group members must deliberate on all possible alternatives jointly. The consensus outcome represents the collaborative outcome of their shared insights and reflections. In the everyday instances of collective decision-making, the group members have limited interactions and sharing of views. The consensus outcome represents an aggregation of their opinions relative to a predefined threshold, e.g., a simple majority. Suppose the level of collective deliberation between participants is more encompassing. In that case, their collaboration will aim to explore all possibilities and agree on a consensus outcome reflecting their shared opinions. In practice, the functioning of a jury that strives to reach a unanimous verdict is an example where the result could surface other alternatives (Sunstein, 2006; Tetlock and Gardner, 2016). For Hayek (1945), the functioning of the market price mechanism is similar to a collaborative consensus outcome, factoring in the participants' opinions to realize a market price that is potentially different from the participants' expectations.

For collective decision-making, how the group members' opinions are aggregated to reach the consensus outcome forms part of the access aspect for the participants.

Regardless of the levels of deliberation amongst the participants, the collective decision-making approach aims to aggregate their individual opinions to report an aggregated consensus outcome. In many instances of collective decision-making, the group members' opinions aggregation occurs without the participants having had any control or input about the aggregation process, e.g., the winner in a political election. With smaller groups that have selective access, the level of deliberation is likely to increase, opening the possibility that the participants may have some control over how the aggregated consensus outcome is determined. For the selective collective decision-making groups, the threshold set for a consensus outcome could be predefined by the group convener with possible input from the group members. For example, the articles of association for some organizations require a minimum of three-quarters of shareholders to vote in favor of specific corporate resolutions. Regardless of the aggregated threshold necessary to achieve a consensus outcome, the role of the participants in the aggregation process is passive, as the mechanism is predetermined.

The market price mechanism, as described by Hayek (1945), represents an aggregation of the choices of dispersed market participants without them specifically interacting to achieve a consensus outcome reflected in the ruling market price. The interactions of the sellers and buyers influence the market price, with the participants able to choose to take part in the group (market). The buyers are dependent on the sellers to sell. Similarly, the sellers depend on the buyers to buy, reflecting the collaboration necessary by the participants for the group's effective functioning. Although the market participants' motivations may differ, a consensus outcome can be achieved only through the group members' interactions. The information communicated through the market price informs the participants' deliberations while more fundamentally being the ongoing aggregated consensus outcome of the participating group members. A collective decision-making process relies on the group members interacting collaboratively based on their dispersed information to achieve a consensus outcome. For the group members on a jury, a project, or negotiating team, to achieve a consensus outcome, the participants have to interact and deliberate on the shared opinions of the group members. In these specific group situations, the consensus outcome is more prone to a negotiated compromise than an aggregation of views.

Collective decision-making can be distilled down to the two elements of participants' interactions and aggregating their choices to reach a consensus outcome. For making decisions collectively, how these elements are structured has a bearing on the approach taken and the involvement of the participants in achieving a consensus outcome. For the users of the consensus results, how the two elements are structured represents a balance between the resources and time available and the reliability of the consensus outcome achieved (Armstrong, 2008). A distinction is drawn between the collective decision-making participants' role and the consensus outcome's determination. When the group participants have a passive role in determining the aggregated consensus outcome process, the collective decision-making approach is classified as deliberation, reflecting their input in surfacing but not determining the aggregated consensus outcome. Deliberation broadly represents the participants' preferences among predetermined choices (Ackerman and Fishkin, 2005; Aristotle and Everson, 1988; Galton, 1907a; Hong and Page, 2012; Surowiecki, 2004; Sunstein, 2006). In some instances, the aggregated consensus outcome cannot be achieved without the involvement of the group participants or could be seen as explicitly emanating from their mutual interactions. The difference versus deliberation is the active involvement of the group members in the aggregation mechanism, reflecting the need for the group members to collaborate to achieve a consensus outcome (Brabham, 2013; Hayek, 1945; Rijshouwer, 2019; Sunstein, 2006; Tetlock and Gardner, 2016). *Figure 2-1* illustrates the two elements across different settings for making collective decisions and how the participants' involvement differs between deliberation and collaboration.

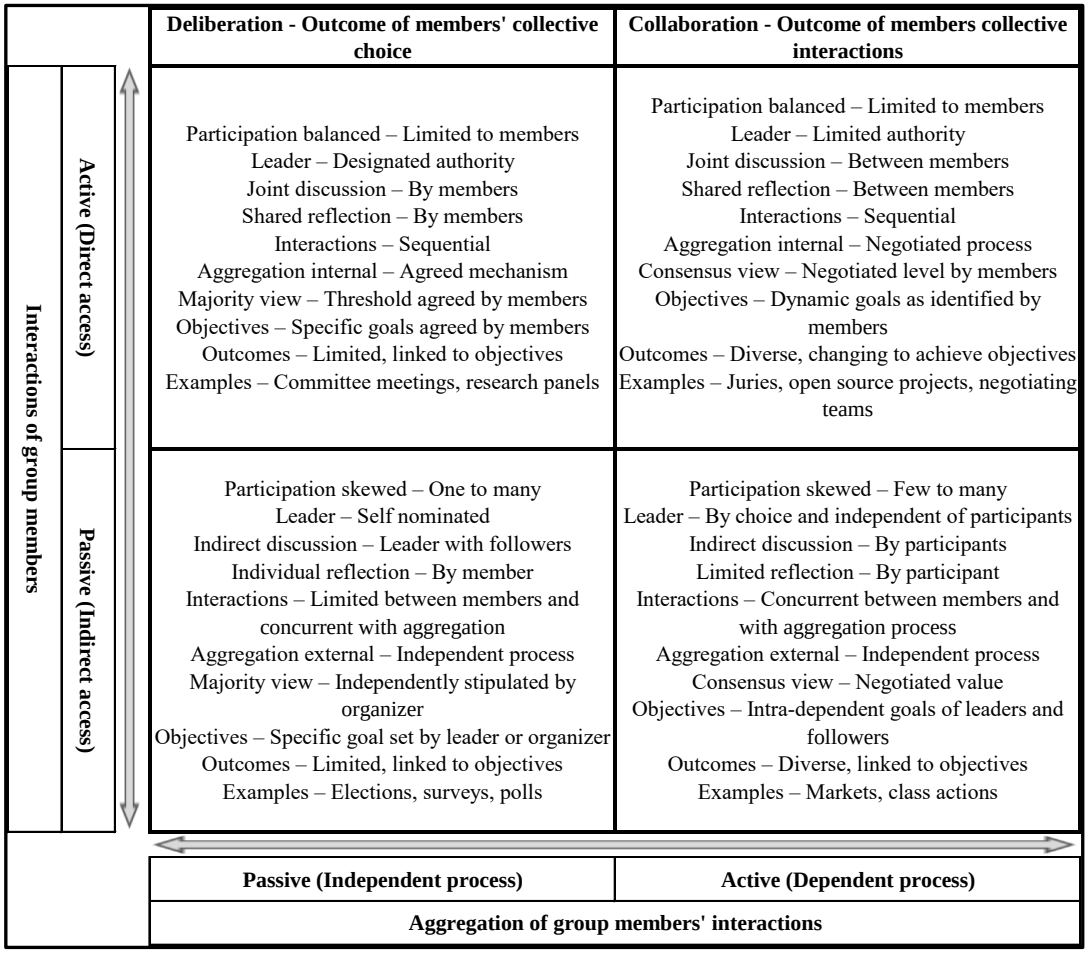


Figure 2-1 Collective Decision-Making Classification

The literature review explored the contrast between deliberation and collaboration and the aggregation process to achieve a consensus outcome. The Delphi Method and crowdsourcing are reviewed as collective decision-making approaches that involve more significant levels of collaboration to reach a consensus outcome, epitomizing the idea of obtaining greater access to "*the collective wisdom of the crowd*" (Brabham, 2013; Rowe and Wright, 2001; Sunstein, 2006; Surowiecki, 2004; Tetlock and Gardner, 2016). Groups involving broad participation in joint decision-making reflect the anticipated collaboration of the various action research modalities (Coghlan, 2007; Coghlan and Brannick, 2014).

2.2. **Deliberation – The Foundation for Collective Decision-Making**

The concept of deliberation by many participants to achieve an outcome considered better than that of individuals deciding alone is associated with the ancient Greeks, such as Aristotle. The following quote from Aristotle epitomizes the rationale

for collective wisdom and is widely seen as a justification for joint decision-making by groups:

"*For each individual among the many has a share of excellence and practical wisdom, and when they meet together, just as they become in a manner one man, who has many feet, and hands, and senses, so too with regard to their character and thought. Hence the many are better judges than a single man of music and poetry; for some understand one part, and some another, and among them they understand the whole.*" (Aristotle and Everson, 1988, pp.66).

The argument for the superiority of collective wisdom has been called the *" summation argument"* (Keyt and Miller, 1991, pp.270). Waldron (1995, pp.569) criticizes Keyt and Miller's "*summation argument*" as the aggregation of "*random and unordered collection of experiences"* that compares individual capabilities rather than being their collective wisdom. The superiority of a group collaboration involves deliberation and aggregation of their collective wisdom, instead of everyone only reporting their insights without any "*summation*" process. For Cammack (2013), the "*rule of the multitud*e" where individual virtue is aggregated is "*amplified when they act collectively*," as they are less easily swayed than individuals or small groups. Aristotle's argument of crowd superiority is that individuals acting alone are less virtuous (*arête*), giving rise to biased judgment, inferring a balancing mechanism exists within a collective decision-making process. Cammack (2013) explains the collective synergistic benefit as joint deliberation brings out the best in people because of emotional factors such as rivalry and social conformity. They wish to be admired rather than gain additional knowledge.

Waldron (1995) recognizes that collective wisdom may occur when individuals act out of self-interest. Still, when their views are aggregated, he sees the collective deliberation as the *"doctrine of the wisdom of the multitude" (DWM),* a more encompassing idea than Keyt and Miller's (1991) *"summation argument."* For Waldron, *DWM* comes from accessing the individual capabilities and aggregating their expertise as a collective group view. *"The people acting as a body are capable of making better decisions, by pooling their knowledge, experience, and insight, than any individual member of the body, however excellent, is capable of making on his own"* Waldron (1995, pp.564). The possibility of questioning during the shared deliberation

makes the collective aggregation of knowledge compared to the individual's capabilities significant to Waldron's DWM.

Explaining the possible limitation of unique knowledge or experience compared to collective wisdom, Aristotle uses the analogy of a feast to highlight the value of diversity versus individual capability. *"And as a feast to which all the guests contribute is better than a banquet furnished by a single man, so a multitude is a better judge of many things than any individual."* (Aristotle and Everson, 1988, pp.76). Under DWM, the value of diversity comes from the collective sharing and deliberation of the information, recognizing "*that we have something distinctive to learn from one another.*" (Waldron, 1995, pp.577). In a contrasting interpretation, Cammack (2013) sees Aristotle's reference in the banquet analogy as not to the pooling of knowledge but to virtue (*arête*). The understanding of Cammack that Aristotle refers to virtue rather than knowledge implies it is the actions of the individuals and not their wisdom that leads to the benefit of acting collectively. The inference that Aristotle was referring to virtue (*arête*) by Cammack links the value of collective participation to the positive emotional interactions of the group, citing Susemihl and Hicks (1976, pp. 398) that "*crowd emotion would inhibit rather than support virtue."*

The group influences a person's contribution they are a member of, opening the possibility that the ability of the collective is greater than that of the individuals acting alone (Waldron, 1995). To prove the point, Waldron uses the example of an organizational department that functions better with one individual than another because of the difference in group dynamics surrounding everyone's group interactions. Group participation requires contributing, and a willingness to deliberate with others their views and accept alternative proposals, or the benefit of collaboration will be lost. The possible differences in individuals' approaches to group interactions exemplify the Waldron point.

Aristotle describes *"endoxa"* as increasing knowledge levels arising through increasing expertise. *Endoxa* progresses from the opinion of many to the majority's opinion, to the experts' opinion, and finally to the opinion of the most knowledgeable (Wikipedia, 2019). Waldron sees Aristotle's philosophy of "*endoxa"* as the progression of understanding alternative views and, through collective deliberation, new knowledge appears. The emergence of knowledge requires a person or method to combine the collective views, for which Waldron refers to Mill's (1956) "*On Liberty"*

and how DWM depends on dialectic interactions. Waldron (1995, pp.577) concludes with his understanding of Aristotle's philosophy of collective deliberation, *"people do better in their practical thinking when they work in groups rather than when they rely, one by one, on their individual excellence.*"

The philosophy of Aristotle about collective deliberation has not materialized as expected, despite broad support. Amongst the issues found by practitioners and researchers are the failure to achieve an inclusive outcome and the possibility of herding or mob rule (Ackerman and Fishkin, 2005; Rheingold, 2002; Sunstein, 2006). An experiment conducted in Colorado in 2005 (Hastie, Schkade, and Sunstein, 2007) saw that rather than allowing opposing views to be deliberated, the group interactions caused a coalescence around previously held positions, resulting in greater extremism in the outcome. In the Colorado experiment, the bipartisan views became more entrenched within the subgroups, limiting the acceptance of opposing views and minimizing any benefit from the group's collective insights. Sunstein (2006) likens the positional coalescing seen to the tendency of groupthink identified by Janis (1973).

For Rawls (1971), deliberation is essential to legislators' functioning, as it would not be possible for all members to know and understand everything. By deliberating, they "*can make all the same inferences that they can draw in concert"* with the caveat (pp.315), *"in the ideal process the veil of ignorance means that the legislators are already impartial."* Rawls echoes the positive and negative aspects of collective deliberation, recognizing the possibility of a less-than-ideal outcome if the participants are not impartial in their acceptance of alternative opinions. Deliberation can boost confidence as group members become more confident of the shared views following their interactions, particularly if it corroborates their prior opinion. The positive connotations of deliberation could also be misplaced, and with pressures from the group, lead to conformity by the subjugation of diverse views (Baron et al., 1996; Brown, 2000; Heath and Richard Gonzales, 1995). In referring to Habermas and his *"forceless force of the better argument,"* Sunstein (2006) expresses caution that polarization and peer pressure may constrain the ideal outcome from deliberation.

The idea that the group can achieve a better collective outcome, based on the philosophy of Aristotle and mirrored by Waldron and Rawls, presupposes no specific experts are available to address the issue independently. The relevance of the group's capability compared to that of an expert is mentioned by Sunstein (2006), concluding

the best outcome can only be achieved if the group is willing to defer to the expert's opinion. A group's willingness to rely on an expert's opinion implies the group members agree to concede to the expert, and the expert accepts the responsibility to decide on behalf of the group. The significance of an expert is essential where the issue requires specific knowledge and differs from the more general situation where collective deliberation would allow diversity to achieve a better outcome, as posited by the proponents of the superiority of collective deliberation. The summation argument is a trade-off between deliberating with diverse views, conceding to the opinions of the more informed group members, and, crucially, knowing which alternative is best at which time (Cooper and Kagel, 2005; Sunstein, 2006).

The promise of deliberation is cautioned with concerns about peer pressure, groupthink, and mob mentality, opening the possibility that mistaken outcomes could occur. Sunstein (2006) defines three measures of the reliability of group deliberation:

- The group reached the best possible result.
- Alternatively, the group effectively aggregates the diverse information held by all members.
- The least ideal is the simple aggregation of group members' inputs to yield a statistical average.

To achieve the best possible outcome, the group members must harness their collective views and obtain sufficient support from within the group to make it the group's selected result. The ideal outcome assumes no hindrance to the deliberation process in the group as compared to preconceived beliefs dominating the deliberating group's interactions inhibiting the surfacing of pertinent information and skewing the aggregated inputs towards a mistaken outcome (Brown, 2000; Gigone and Hastie, 1997; Kerr et al., 1996; Sunstein; 2006). If deliberating groups cannot fare better than statistical averages, as an approach, it requires understanding why and, in the worst case, discarding the process in favor of survey panels (Gustafson et al., 1973; Sunstein, 2006).

For deliberating groups, the average group performance will not always exceed that of the best member's capability (Gigone and Hastie, 1997; Hastie, 1983; Stasser and Dietz-Uhler, 2001). When group members are expected to share their value judgments, which will be subjected to deliberation within the group, it may cause some

group members to become reticent about participating (Brown, 2000; Sunstein, 2006). Peer pressure matters less for deliberating groups dealing with definitive issues, and the problem is more if group members concede to the most informed member (Gigone and Hastie, 1997; Stasser and Dietz-Uhler, 2001; Sunstein, 2006). If a group participant knows the correct answer, it could inspire the other group members to support that position (Hastie, 1983; Kerr and Park, 2001). When deliberating groups are tasked with complex strategic problems, and the diversity of knowledge and expertise amongst group members is high, the group's performance is likely to outperform that of individual group members (Blinder and Morgan, 2000; Cooper and Kagel, 2005).

Sunstein (2006) summarizes the case for deliberation as it will outperform a statistical averaging approach if the correct outcome has initial backing within the group. The framework used by the group will allow convergence on the best result, mainly if the group is highly cohesive with a pronounced sense of identity (Cooper et al., 2001). The phenomenon of deliberating groups of members not contributing their insights has two underlying causes, fear of social stigmatization for not conforming to majority group norms or if a predominant expert in the group occupies the position of gatekeeper (Krech et al., 1962; Loury, 1994; Marques et al., 2002).

2.2.1. **Conclusion: Deliberation – The Foundation for Collective Decision-Making**

Deliberation as a collective decision-making approach can have varying levels of interaction combined with a passive aggregation process and is well suited to determining the preference of groups. Regarding access, participation is restricted to group members permitted to deliberate on the issue under consideration. For those situations that involve many group members, the level of interaction, the influence exerted on other group members' decisions, and sharing of insights are likely to be limited. The role of a leader in the collective decision-making process would be aimed at achieving their preferred outcome. When the number of group members is smaller, based on more restrictive participation criteria, the level of interaction is expected to be more significant, opening the possibility of greater sharing of insights and the potential to influence other group members. Given the smaller select group of participants, the leader's preferences are probably less influential in deciding the outcome.

Regardless of the level of interactions among the group participants, the aggregation process is expected to be passive, involving limited input in its formulation by the group members. The aggregated outcome is decided by a majority vote or a similar consensus measure that includes all group members. The collective decision in these instances is more akin to the lowest common denominator than the broad consensus amongst all participants. Due to the participants' interactions, they could be influenced by more confident group members, succumb to peer pressure, or be unwilling to discard their prior beliefs, undermining the reliability of the collective decision reached. Although deliberation has the potential to surface the collective knowledge of group members, the potential constraining issues, such as limited interactions, group influences, and passive aggregating mechanisms, may limit the reliability of the outcome.

2.3. **Collaboration – Collective Participation and Crowdsourcing**

The internet has led to the ability to access contributions from numerous independent experts to achieve a collective result that exceeds the individual capability of the contributors (Brabham, 2013). Reflecting on the research of Jeppesen et al. (2007), Howe (2006, pp.4) observes, "*The most efficient networks are those that link to the broadest range of information, knowledge, and experience.*" Howe recognized the trend of organizations to increasingly reach outside their formal structures to source solutions for projects from outside experts, defining the approach as "*Crowdsourcing.*" Brabham (2013, pp.12) defines crowdsourcing as a "*distributed problem-solving and production model that leverages the collective intelligence of online communities to serve specific organizational goals.*" Underlying the functioning of crowdsourcing is the collective interaction of independent experts to collaboratively produce a result that exceeds that obtainable by relying on individual experts or other resources.

Part of the debate around crowdsourcing is how the approach can be defined to reflect the collective access to experts and how they collaborate to produce a result. Estellés-Arolas and González-Ladrón-de-Guevara (2012) identified the following elements forming the core of crowdsourcing:

- An organizational project.
- Several experts are willing to work on the project.

- A process that allows the experts to interact collectively on the project, such as an online participation system.
- The participants and users perceive a benefit for themselves in undertaking the project.

Of these defined elements, two aspects call for further consideration. The formulation of the definition of Estellés-Arolas and González-Ladrón-de-Guevara (2012) assumes the collaboration will only occur online and the impetus for the project will be driven by the convening organization (Brabham, 2013). Project teams within organizations can and do work collectively in person to collaborate on projects that call for multi-disciplinary expertise, not least the functioning of the board of directors of most organizations. Restricting the definition to only online limits the use of considering the approach in the practice of the action modalities in the workplace, where collaboration represents a key element (Coghlan, 2010; Coghlan and Brannick, 2014). The issue of who initiates and controls the project is more challenging to resolve. Although it is typical for a project to be undertaken by a user of the outcome, the possibility that the participating experts will collaborate to develop an independent solution is expected (Howe, 2006; Jeppesen et al., 2007). For Brabham (2013), Wikipedia fails to meet the criteria of a crowdsourcing project. Brabham sees the locus of control resting more with the contributors, questioning whether the same outcome could have occurred through a different approach. Suppose the control issue is not as crucial as Brabham indicated. In that case, the Wikipedia experts' collective effort to produce an outcome that exceeds that of the individual contributing members in terms of the volume and content could represent a crowdsourced outcome.

Wikipedia provides articles covering numerous topics, with over forty-eight million articles available online as of January 2020 (Wikipedia, 2020). A challenge in delivering online articles is ensuring the trustworthiness of the information provided. The reliability issue stems from trying to gain access to the relevant experts, who may not be willing to contribute their knowledge willingly or may be inaccessible because other people have gained access to them first. The experts could be seen as a "*common*" online resource available to all for contributing and collaborating on a wide selection of issues and topics (Hardin, 1968). Ostrom (1973) cautions about the challenges of accessing a common resource, such as the internet experts, as a "*tragedy unless the structure of decision-making arrangements can be modified to enable persons to act*

jointly in relation to those resources as a common property" Ostrom (1973, pp.210-211).

In his controversial paper, "*The tragedy of the commons,*" Hardin (1968) set out how unchecked population growth was a problem with "*no technical solution*" because of the individual pursuit of self-interest and because of no control deterrents or specific prohibiting regulations. In formulating his view of access to and use of common resources, Hardin conflated the use of the resource and the control thereof into an indistinguishable whole (Frischmann et al., 2019). The conclusion of Hardin that a resource could be depleted faster than it could be replenished did not consider the possibility of regulating controls by those with a vested interest, as has been more commonly seen in practice (Frischmann et al., 2019; Ostrom, 2000). As a "*common*" resource with limited control and regulations, the internet poses challenges concerning accessibility and trustworthiness. Researchers have seen the internet as a "*common*" source of data to investigate the phenomena they are interested in, without always considering the circumstances that gave rise to the existence of the data.

Ostrom (1998) questioned the behavioral model that collective decision-making is primarily based on self-interest, noting that communication occurs between participants when faced with a conflict. Face-to-face interactions that allow for verbal and visual observation and feedback lead to better outcomes, with indirect communication having a similar benefit, albeit to a lesser extent (Frischmann et al., 2019; Ostrom, 1998). For Ostrom, communication within a group led to more favorable outcomes when those interactions were based on trust, reciprocity, and reputation (Frischmann et al., 2019). Trust refers to our and others' expectations about themselves, reciprocity relates to the fairness of our interactions with others, and reputation is the record of past interactions with others (Ostrom, 1998). The factors of trust, reciprocity, and reputation within the Ostrom behavioral model are difficult to conceptualize within the indirect communication that occurs on the internet, challenging the potential benefit from group interactions in such an environment.

Wikipedia exemplifies a "*common*" resource functioning according to the precepts of trust, reciprocity, and reputation of the Ostrom behavioral model. The emergence of Wikipedia was made possible by developing communication channels, which opened participation to information providers and users without needing them to interact directly within a self-governing system of trustworthiness (Rijshouwer, 2019).

The growth of Wikipedia is ascribed to the actions of earlier participants drawing in newer participants and allowing them to make further additions to the information pool (Heylighen, 2016). The founder of Wikipedia, Wales, compares the collaboration of Wikipedia participants to the aggregation of information performed by Hayek's price system. *"If information is widely dispersed, and if no single "planner" has access to what is known, then Wikipedia's method of operations has the same general justification as the price system"* (Sunstein, 2006, pp.156-157).

Wikipedia's founders envisaged broad collaboration, controlled through the participants' interactions, to provide a reliable, neutral information source mediated through a collective spirit of goodwill *(*Rijshouwer, 2019; Wikipedia, 2001*)*. Wikipedia founder, Wales, aimed to reach a consensus through deliberative decision-making by acknowledging other participants' contributions in an environment of "*love*" for their collective contributions (Rijshouwer, 2019; Wales, 2004). Möller (2003) called for "*WikiLove*," the "*common goal*" for the "*love of knowledge*" achievable through keeping an "*NPOV*" [neutral point of view] despite any differences, enabling a positive collaboration amongst Wikipedia contributors (Rijshouwer, 2019, Wikipedia, 2001). Wikipedia's ideology mirrors the elements Ostrom found for positive group interactions without the restrictive regulations recommended by Hardin (1968). As a specific internet project, Wikipedia has the possibility of aiming for trustworthiness and supplies a structural model for using the internet as a forum for group collaboration.

Wikipedia uses an open-source approach for hosting content verified by editors. Volunteers who receive no monetary reward perform many tasks based on a consensus ideology (Rijshouwer, 2019). Fundamental to the functioning of Wikipedia is the assumption of collaboration in good faith and neutrality in the actions of participants (Reagle, 2010; Rijshouwer, 2019). Wikipedia has evolved a set of Policies and Guidelines that are expected to be followed by organizational contributors, consisting of "*policies*" requiring strict compliance, "*guidelines*" requiring community compliance, and "*essays*" requiring consideration (Rijshouwer, 2019; Wikimedia, 2017). The internet functions as an "*open source*" of information available to any interested person without any cost. Like Wikipedia, many internet forums have "ground rules" that must be adhered to by participants, mirroring the expectations of Hardin (1968). The "*common*" resources function best when controlled and regulated.

Raymond offers a charitable explanation for the willingness to contribute without monetary reward (2001, pp.110), *"participants compete for prestige by giving time, energy, and creativity away"* as a "*gift*" rather than a commercial exchange. Over time, the demographics of Wikipedia volunteer contributors have become more predominantly Western-educated white males (Rijshouwer, 2019). Rijshouwer sees the demographic changes at Wikipedia as detrimental to the organization's diversity aims. Still, the consequence is linked to the notion of Raymond that contributors look to serve with their peers to whom they can relate emotionally when purely financial considerations do not drive their exchanges. An implication of using the internet as a platform for collaboration and encouraging diverse participation is to reward participants by recognizing their differences rather than looking to control their contribution through consensus and conformity.

2.3.1. **Conclusion: Collaboration – Collective Participation and Crowdsourcing**

Collaboration as a collective decision-making approach anticipates significant interactions between group members, and the aggregation processes involve the direct input of the participants. Regarding access, participation concerns participants who agree to join the group because of a vested interest in the outcome. The level of interaction is high, and the consensus outcome is the culmination of the collective contribution of several but not necessarily all, group participants. Because of the extensive interaction between participants, it allows intragroup persuasion, moderated by the group's norms and rules as has evolved from their past interactions. If some group members exercise greater control and influence over the composition of the participating group members, those participating will progressively conform more closely to a defined group norm. The result could be the group composition becoming more homogenous and the outcome aligned with the dominant members of the group. Potentially not all participants are regarded equally in the collaboration process. A dominant leadership position can be ascribed to a group member or assumed if a group member can exert more influence within the group than other participants. Typically, the group members should have positive reputational regard for their cohorts for their collective interactions to capture the potential contribution of all participants.

The aggregation process involves the collective outcome of the group members' interactions to result in a consensus acceptable to the majority of the participants. The significant collaboration between the group members could result in a consensus

outcome that surpasses the participants' capabilities. The aggregation process is likely more sophisticated than the standard statistical methods used for deliberation, as the consensus outcome may reflect different permutations than some predefined choices. The potential of achieving enhanced results from group interactions is more likely if trust and reciprocity in the group are balanced. The risk of group members enforcing conformity amongst participants could diminish the diversity of the group's collaboration efforts, influencing the motivation of new participants to join, contribute their insights and remain vested in the group's collaboration efforts.

2.4. **Aggregation - The Law of Averages and Condorcet's Jury Theorem**

When acting collectively, people need a method to aggregate their opinions to represent their shared consensus view. In 1785 Condorcet set out an approach to predict the probability of the outcome of voting in political events, "*Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix" (*Translation: *"Essay on the Application of Mathematics to Decision-Making,"* Baker, 1976). Condorcet's approach has become widely used in many settings to predict the outcome of group voting and in many academic studies (Dietrich and Spiekermann, 2016). Although the method has restrictive assumptions about the participants' behavior, applying the approach and its results are still valuable in practice (List and Goodin, 2001; List, 2008). In contrast to the collective interactions between participants in the deliberation approach based on Aristotle's philosophy, the underlying premise of Condorcet's approach assumes independence in the actions of the participants.

In an experiment to assess the statistical accuracy of collective group actions, in 1906, Sir Galton calculated the median of participants' estimates at a fair in an ox-weighing competition, with the computed result almost matching the actual ox weight. In a more recent review of the collective wisdom of crowds, Surowiecki (2004) cites the example of finding a sunken submarine, with the group's average location forecast proving surprisingly correct in finding the debris from the sunken submarine. A fundamental difference between merely aggregating the views of a group and prior deliberation is the degree to which the interaction can influence the decision-making process and, so, the outcome.

Aggregating group opinions using Condorcet's Jury Theorem requires understanding the detail and considering how as an approach, it differs from aggregation methods used in deliberation approaches. An expectation exists that taking the average of participants' views when information sharing is limited or non-existent could yield a more accurate outcome than from deliberating groups, be it because of significant prior knowledge or less undue influence on the opinions of other participants (Armstrong, 2001a; Batchelor and Dua, 1995; Drummond, 2001). Using statistical aggregation to collate the opinions of a group includes the diversity of views more effectively than in deliberating groups, where peer pressure may impede diverse opinions from being discussed and considered (Rowe and Wright, 2001; Stewart and Lusk, 1994; Sunstein, 2006). Despite the concerns about the restrictive assumptions about participants' independence, the Condorcet Jury Theorem will reflect the majority opinion, regardless of whether it is the right or best outcome (List, 2008).

Using Condorcet's Jury Theorem to calculate the probability of achieving a clear majority approaches one hundred percent as the group size increases significantly, with a similar trend for smaller groups if a predominant preference exists (Mueller, 2003). Mathematically, the Condorcet Jury Theorem aggregates the participants' opinions without making a value judgment of their actions, reflecting the probability based on the majority's preference. The assumptions made by Condorcet are as follows (Sunstein, 2006):

- Participants are not concerned if their vote is decisive in the outcome – the Independence of the participation.
- Participants are not concerned with how other participants voted – the Independence of participants' votes.
- Participants are not concerned with other participants' views – the Independence of the participant's actions.

The aims of limitations assumptions are explained by List (2008) as the "*coherence challenge*," which seeks to minimize the consequence of inconsistencies that could influence the outcome, to make the results biased and unusable.

The assumption of participation independence holds, as, in most voting scenarios, the vote-counting only occurs after everyone has voted rather than progressively while the voting is ongoing. The position could be different under committee settings where

votes are visible to other members; for instance, a chairperson has a deciding vote. The second and third assumptions are more difficult to ignore, as peer pressure or similar backgrounds could influence participants' vote independence and preferences. Prior research has proven that the Condorcet Jury Theorem yields reliable results even with some violation of these assumptions (Bottom et al., 2002). An example of the application Condorcet Jury Theorem in practice is opinion polls. Participants express a preference about an issue instead of a random guess, and the outcome is decided by the majority preferences (Sunstein, 2006).

To better understand Condorcet's Jury Theorem and the importance of participants' preferences and independence, it is helpful to consider the calculation without going into specifics. For a group with [n] participants voting on two alternatives, each participant has a probability [p] of choosing a particular choice. The likelihood of the group selecting a specific option can be calculated as follows:

$$Pn = \sum_{h=(n+1)/2} \left[\frac{n!}{h!(n-h)!}\right] p\text{^}n(1-p)\text{^}(n-h)$$

(Mueller, 2003).

From the above formulation of the Condorcet Jury Theorem, as the number of participants increases, so will the rate at which the calculated outcome probability approaches one hundred percent. As the probability value [p] increases, so will the cumulative probability of achieving a majority supporting the choice. Suppose there is no majority preference amongst group members for an outcome. In that case, the failure to have some uniformity can result in the probability declining to zero as opposing votes offset each other.

An outcome of the Condorcet Jury Theorem probability calculation is the possibility of rapidly approaching extremes if the number of participants becomes large and the shared beliefs of the participants are skewed above fifty percent. The consequence can be accepting or rejecting a choice based on an overwhelming or underwhelming estimate of the probability when the reality might be different. The Condorcet Jury Theorem derivation can be influenced by the shared biased beliefs of the participants, which may not be directly accounted for in the formulation of the voting assessment, such as another exogenous factor (Baker, 1976).

The Condorcet Jury Theorem assumes that participants have a similar belief in an alternative's preference, the probability [p]. If, for example, a core of expert participants votes for an option and the rest of the participants vote randomly but equally distributed across the alternatives, the best outcome could still be achieved. The result could be less accurate when the less-informed participants mostly oppose the better-informed participants (Sunstein, 2006). Expecting to have a group of informed experts determines the outcome, while the less knowledgeable participants' votes are more or less offset, placing reliance on the chance to ensure a reliable result. The possibility that all participants may be biased towards a particular outcome, such as herding seen in stock markets occasionally, may result in a poor outcome regardless of the experts present (Shiller, 2005).

In casting their votes in polls, participants are subject to similar influences to those found in deliberating groups, such as group biases, peer pressure, and conformity expectations. As with any decision-making process, undue pressures will affect the outcome of the polling process (Sunstein, 2006). In a similar vein, should the participants be uninformed about the issue being questioned and the votes are random, despite the assumptions of Condorcet Jury Theorem holding, the resulting outcome is unlikely to be correct or reliable, reflecting a wide dispersity of [p] or probability values (Mueller, 2003). Condorcet Jury Theorem could also be ineffective when the issue is unfamiliar to the participants, causing the participant's contributions to be randomly distributed and leading to inconsistent errors, reflecting a low [p] value (Lorge et al., 1958; Mueller, 2003). The Condorcet Jury Theorem limitation with unfamiliar issues highlights a relative benefit of deliberation, where the participants' interactions could surface a less random outcome.

Condorcet was aware of the concerns around participants' limitations, calling *"that voters be enlightened; and that they are the more enlightened, the more complicated the question upon which they decide." (*Baker, 1976). To overcome the possibility of poor decisions being made by uninformed participants, Condorcet favored the delegation of responsibility for voting to expert participants who would be less likely to make poor decisions because of the higher probability [p] that they knew the correct answer (Baker, 1976). Condorcet's idea of having better-informed representatives vote on behalf of the populace is questionable, as it goes against the grain of independence of voting and actions, reducing the reliability of the approach. When the majority vote,

those considered to have less knowledge of the issues involved would result in an outcome more reflective of the popular preference and less biased based solely on the experts' opinions (Sunstein, 2006).

If an expert can persuade the group of a more accurate outcome, it would make sense for a panel of experts to outperform a single expert if the interaction amongst the group of experts is constructive (Sunstein, 2006). A panel of experts is best when the issue under consideration involves a high degree of judgment rather than a mere factual question (Armstrong, 2001a; Batchelor and Dua, 1995). Armstrong (2001a) cites a group of experts outperforming individual experts in diverse issues by twelve percent, even more significantly in specialists' areas where experts have unique insights. For organizations striving to make more reliable forecasts, relying on a panel of experts rather than just a sole expert would seem advisable (Armstrong, 2001a; Sunstein, 2006). Numerous services exist to supply consensus expert opinions on various economic and related industry trends (Consensus Economics, 2020; The Economist Intelligence Unit, 2020; Thomson Reuters GFMS, 2020; Wood Mackenzie, 2020).

The Condorcet Jury Theorem aggregates the inputs for questions with a limited plurality of outcomes, most effectively if some participants know the correct answer and the distribution of the other participants' errors is random. Also, if the participants' probability [p] favors the best outcome, the aggregating approach of the Condorcet Jury Theorem functions well. However, the Condorcet Jury Theorem has no guarantee of surfacing the best result if there is a systematic bias in the participants' judgments. The conclusion points to balancing the use of many participants with fewer better-informed experts if the aim is to achieve a more reliable forecast in the workplace while being mindful of groupthink (Armstrong, 2001a; Batchelor and Dua, 1995; Janis, 1973; Mueller, 2003; Rowe and Wright, 1999).

Surowiecki (2004) motivates the concept of the "*wisdom of the crowd*" by discussing an experiment conducted by Sir Galton in 1906 at an English country agricultural fair, helping to popularize the application of aggregating groups' views. Sir Galton tested the phenomenon of collective wisdom of many participants, referring to the experiment as an *"investigation into the trustworthiness and peculiarities of popular judgments"* (1907a, pp.450).

The 1906 fair competition structure can be compared to Condorcet Jury Theorem's assumptions. The comparison is relevant beyond the experiment of Sir Galton, as the structure is comparable to the format found in many forecasting events.

- No speechmaking was used to influence the judgment of the participants.
- A participation fee was charged to encourage reliable estimates and minimize random guessing.
- Numerous participants were competent in the competition issue and included some with the ability to have a fair idea of the final result.
- The group included a wide diversity of participants.

Comparing the competition conditions to the assumptions underpinning Condorcet Jury Theorem, none of the assumptions could be completely satisfied. Given the payment to take part, it could be assumed each participant believed they had the winning estimate. Galton's comment contradicts the assumption of independence of voters, "*others were probably guided by such information as they might pick up"* (Galton, 1907b, pp.450). Although no speechmaking was observed, the participants interacted in formulating their estimates. Galton (1907b, pp.451) mentions "*the use of a small variety of different methods, or formula,"* showing participants' experience guessing ox weights.

In reply to Sir Galton's article, Perri-Coste (1907) comments that many of the participants were more capable of making their estimates than a general collection of people:

- Many participants were experts in estimating the correct answer in the competition.
- For the experts, the nature of the competition was akin to their usual business practices and the field in which they had extensive knowledge.
- The participants often competed in similar competitions, gaining experience in the field.

Galton (1908, pp.281) later acknowledged that "*the proportion of the voters who were practised in judging weights undoubtedly surpassed that of the voters in ordinary elections who are versed in politics."*

Galton (1907a, pp.450) reported the results of the ox weighing competition, "*The distribution of the estimates about their middlemost value was of the usual type, so far that they clustered closely in its neighbourhood and became rapidly more sparse as*

the distance from it increased." The outcome matched the Condorcet Jury Theorem expectation, given the knowledge and experience of the experts participating in the competition. The actual result of the competition was that the slaughtered ox weighed 1198 pounds (Wallis, 2014). Galton advocated taking the median as the most representative aggregated value, "*one vote one value, the middlemost estimate expresses the vox populi, every other estimate being condemned as too low or too high by a majority of the voters*" (1907a, pp.414). Galton's calculated median of the 787 participants was 1207 pounds. Questioning Sir Galton's use of the median, Hooker estimates the mean value to be 1196 pounds. Galton (1907b) confirmed the average was 1197 pounds, only one pound less than the actual weight.

In defense of his use of the median, Galton said, *"The best interpretation of their collective view is to my mind certainly not the average, because the wider the deviation of an individual member from the average of the rest, the more largely would it affect the results. In short, unwisdom is given greater weight than wisdom"* (Galton, 1908, pp.281). In support of Galton's approach Levy and Peart (2002, pp.358) point out giving an equal vote to *"cranks in proportion to their crankiness"* would unduly skew the results of the average for extreme estimates made by participants and instead supports the use of the median. Mirroring Galton, Levy, and Peart (2002, pp.358-359) state, "*The best interpretation of their collective view is to my mind certainly not the average, because the wider the deviation of an individual member from the average, of the rest, the more largely would it affect the result. In short, unwisdom is given greater weight than wisdom."* The debate around the best method of aggregating results is informative but not conclusive as, in practice, both medians and means are used.

2.4.1. **Conclusion: Aggregation - The Law of Averages and Condorcet's Jury Theorem**

The Law of Averages as an aggregation approach requires minimal interaction from the group participants. It is well suited for determining the majority consensus result for large groups choosing from limited options or pooling their factual judgments. The aggregation process could yield a probability measure of the result or another consensus value, such as the average or median of participants' estimates. The most common outcome measure is average, but the median is preferred in some circumstances to avoid outliers' undue influence. As the Law of Averages is suited to

aggregating the views of many participants with minimal interactions, the possibility of excessive influence by other participants is expected to be limited. However, influential external sources can play a significant role. If the process is duly hampered by external interference, the consensus outcome could similarly be influenced relative to the participants' preferred choice. Given the low level of interaction and opportunity for persuasion, surfacing new insights will be limited. In smaller group settings, the possibility of undue influence can be more pronounced. The influence could arise from the authority of the group leader or a deemed expert within the group. Peer pressure can also be a more significant factor in smaller groups. Because of the passive nature of the aggregation, it should be less prone to bias or silence within a larger participating group, although still possible from external sources.

2.5. **Aggregation: The Coordination Role of Market Prices**

Under a barter system, value derivation is difficult but still necessary to arrange an exchange. With the use of currencies, the exchange is less complex, although it still requires the value of the purchased items to be mutually agreed upon. For the vast diversity of goods and services transacted continuously, determining a value is necessary for the exchange process to proceed efficiently. How the complexity of the value determination is achieved is best described by Hayek (1945, pp.519), "*the data from which the economic calculus starts are never for the whole society given to a single mind which could work out the implications, and can never be so given."* Hayek (1945) understood that bits of economic information are dispersed among the many market participants, with no single person or entity having access to or understanding all the available information.

For Hayek, the challenge of economic activity is complex planning, requiring many people to coordinate diverse and widely distributed information they have and their need to collaborate to achieve the market outcome. Economic activity is seen as taking advantage of unique information within a specific time and place to make the most informed value judgments for the market participants involved. Despite having unique information, market participants must consider the broader economic consequences of their decisions to make the best choice for each of them (Hayek, 1945). For Hayek, the scope and dispersion of necessary information to achieve the market outcome could not be captured and processed by a central planner. Hayek understood some of the information would only be implicit to those with specific

detailed knowledge and become explicit to other market participants through the exchange process.

The participants' exchange actions are coordinated through the setting of the mutually agreed market price in Hayek's analysis, *"to coordinate the separate actions of different people in the same way as subjective values help the individual to coordinate the parts of his plan."* (Hayek, 1945, pp.526). The functioning of the market system is perceived to work by Hayek through the interactions of market participants with "*sufficient overlapping*" interests and information without any single person knowing everything. For Hayek, prices supply the coordination role for the practical problem of dispersing information to those requiring it, while the system functions based on participants' interactions.

Sunstein (2006) sees Hayek's price mechanism as an effective aggregator of dispersed unique information held by many independent participants, including judgments about value, without coercion or control by any third party. For Sunstein, the market price mechanism achieves the requirement for deliberation across dispersed participants, which dynamically signals to other market participants the evolving value judgments of all other market participants. Compared to the assumptions underlying Condorcet's Jury Theorem, market prices aggregate information because the participants have a vested interest in the outcome based on their interactions by relying on their insights. For Sunstein, the structure of Hayek's market information aggregation surpasses the probability approach of the Condorcet Jury Theorem.

An example of using market prices to forecast outcomes is prediction markets, e.g., Iowa Electronic Markets (IEM, 2019). In several practical applications, prediction markets have proven more reliable than deliberating groups (Abramowicz, 2004; Hahn and Tetlock, 2006; Hanson, 1999; Levmore, 2002; Wolfers and Zitzewitz, 2004). Prediction markets address one of the shortcomings of deliberations groups by incentivizing participants to share their unique information (Sunstein, 2006). For prediction markets to be effective, it has been found that the incentive need not be financial, for example, recognition from fellow participants (Figlewski, 1979; Roll, 1984; Sauer, 1998; Servan-Schreiber et al., 2004; Surowiecki, 2004; Wolfers and Zitzewitz, 2004). Another contributing factor to the reliability of the prediction market is the focus on the expected outcome of all votes, compared to polls, such as in the

Condorcet Jury settings, that consider the actions of the individual participant that are to be aggregated (Sunstein, 2006).

In an organization study at Hewlett Packard, several participants were asked to predict equipment sales in conjunction with the California Institute of Technology. The results from the internal prediction market proved to be more reliable than estimates made by deliberating groups (Chen and Plott, 2002). In 2005 an internal prediction market was set up by Google to forecast aspects of the organization's performance (Cowgill, 2005). Google found that the internal prediction market worked well for its project, especially as time passed and uncertainty surrounding scheduled events diminished. The following comparison by Cowgill offers an insightful understanding of the dynamics of prediction markets. "*Our search engine works well because it aggregates information dispersed across the web, and our internal predictive markets are based on the same principle: Googlers from across the company contribute knowledge and opinions which are aggregated into a forecast by the market"* (Cowgill, 2005).

The Efficient Market Hypothesis of Fama (1970) makes assumptions about the efficiency of information dispersion in determining stock prices of publicly-traded organizations. The degree of information available to participants is comparable to Hayek's philosophy that prices are decided through the actions of a myriad of participants who have unique information which becomes known to other participants through their interactions. As a qualification of the perfection of stock prices as an example of Hayek's price formation expectations, Sunstein (2006) cites the inefficiencies noted by behavioral economists about undue influence in collective deliberation (Shleifer, 2000; Thaler, 2005). Shiller (2005) describes occasional significant stock market price adjustments as a psychological re-evaluation of future expectations made when prior pervasive sentiment becomes less optimistic, a case of substantial changes in the expectations of many markets' participants. The implication is prediction markets offer an alternative to other group aggregation processes but can similarly change their predictions if the "*mood*" of participants changes simultaneously.

For the Iowa Electronic Markets (IEM, 2019), Klarreich (2003) saw that the frequent traders, accounting for about fifteen percent of all trades, are the price-makers, acting as arbitrageurs to reprice predictions for changing expectations. The

activity of the price-makers drives the prediction market to its outcome, as opposed to deliberating groups that seek consensus across all participants and statistical groups that measure consensus based on all votes cast. The reward for aggregating the dispersed information makes the outcome of prediction markets more reliable than for the other approaches, partly because the process is a continually evolving outcome over time (Sunstein, 2006). The risk for organizations using internal prediction markets is the possibility of undue influence and groupthink, while divergent predictions might attract the ire of management, despite being more accurate (Abramowicz, 2004; Chen and Plott, 2002; Sunstein, 2006). As an example of groupthink within an organizational setting and the undue influence of senior management, the systematic error of the United States of America's assessment of Iraq's weapon capabilities is cited as an example (Snowberg et al., 2005). The failure of the prediction markets concerning Iraq's weapons capabilities highlights that access to dispersed information is necessary for the market to function effectively (Sunstein 2006).

As with traditional markets, prediction markets can be manipulated by traders looking to extract an undue advantage, although the actions could also be linked to biased beliefs leading to actions in hindsight seen as mistaken (Abramowicz, 2004; Hanson, 1999; Shefrin, 2001; Shiller, 2005; Thaler and De Bondt, 1993). The prediction markets for political elections show biases amongst participants for the party or candidate they support, influencing their actions in setting prices, mirroring similar investor trading patterns seen in traditional markets (Forsythe et al., 1999; Jolls, 1998; Lord et al., 1979). As with conventional markets, prediction markets can be prone to herding or information cascades when participants are influenced by preconceived beliefs or following the lead of influential figures (Sunstein, 2006). The ongoing prevalence of bias goes contrary to the perception that the arbitrage actions of the price-makers will not price the inconsistency out of the prediction market. Concerns that prediction markets could be seen as betting platforms and suffer from long-shot bias have not been seen (Forrest and McHale, 2005; Manski, 2006; Thaler and Ziemba, 1988; Wolfers and Zitzewitz, 2004). As with traditional markets, some prediction market participants may trade expected outcomes at significantly different prices to the other market participants. Their trades could represent greater insight or an opportunistic trading strategy, which will only be known when the future event

occurs while still registering as a factor in the overall prediction price determination (Posner, 2004).

2.5.1. Conclusion: Aggregation – The Coordination Role of Market Prices

The functioning of markets epitomizes an aggregation approach that depends on limited direct access between the participating group members. By collaborating, the market participants determine a consensus outcome that reflects their collective decisions. By collaborating, the participants' value judgments are systematically aggregated without their intervention, which is then reflected in the market price. In prediction markets, the "price" can more correctly be considered a collective group probability of a specific outcome occurring rather than a value in the traditional market sense. Due to participants' limited direct interactions, intragroup persuasion can be expected to be minimal and instead conveyed through the changing price. Still, external influences on the market can play a significant role. The limited direct interaction could constrain the surfacing of new knowledge for specific group members, although the consequence of the knowledge is likely to be observable to all participants. Through the actions of arbitrageurs, such information is more probable to emerge, exerting an effect on the market outcome. In an active market, the prevalence of bias should be eliminated by arbitrage traders, assuming the full availability of relevant information. Participation targets self-selected participants interested in the specific item or issue and are well suited for making value judgments.

2.6. Delphi Method – Anonymous Deliberation and Collaboration

An approach used to collect group opinions and reach a consensus from interacting experts is the Delphi Method developed by the Rand Corporation in the 1950s (Dalkey and Helmer, 1963). The Delphi Method incorporates deliberation, collaboration, and averaging within a framework structured to mitigate the possible consequence of undue influence between group members and the group convener and users of the group's outcome. The Delphi Method approach aims to achieve an unbiased result from a group of experts. It tries to minimize the negative consequences of the alternatives of averaging, deliberation, and collaboration approaches to see if the method can achieve a consensus result (Bolger et al., 2011; Bolger and Wright, 2011).

The Rand Corporation developed the Delphi Method to undertake strategic military intelligence evaluations for the United States Airforce. The primary aim was

to access experts' confidential opinions in developing a consensus on the expected outcome of various possible military actions (Brown, 1968; Dalkey and Helmer, 1963; Rowe and Wright, 1999). Owing to the military environment within which the approach was developed, the structure of the Delphi Method had specific protocols for its use, which may not be as essential in other workplace environments. The Delphi Method is structured around four requirements (Dalkey and Helmer, 1963; Rowe and Wright, 1999):

- Anonymity: The participants are not informed who the other group members are and are discouraged from discussing the group's activities with others. An essential consideration in making participation anonymous is to minimize the possibility of undue pressure from other participants or interested outsiders to conform to peer predictions. Participants are encouraged to be forthright in making their estimates through anonymous participation.
- Iteration: Multiple rounds of questioning and feedback until the moderator or end-user of the outcome accepts the consensus outcome. The first round can be unstructured to surface relevant issues for further consideration in later rounds. The total number of iterations usually does not exceed more than four rounds. The motivation for multiple iterations is to allow participants to revise their estimates based on the feedback from the moderator of the other group members' estimates, aiming to result in a consensus estimate that is considered the best.
- Feedback: The moderator collates the participants' responses and gives feedback, either only a consensus statistic, such as the median, or the reasons provided by other participants, on an anonymous basis, for their estimates. The feedback encourages consideration of different perspectives and achieves a consensus acceptable to most participants. The possibility exists that by giving feedback to participants, either other group members could guess their identities, or pressure could occur to conform to the consensus opinion.
- Aggregation: Depending on the nature of the questions posed to the participants, the usual aggregation would be either the average or the median of the group's responses. By aggregating participants' responses, the feedback could influence the later responses of participants, causing them to change their estimates closer to the prevailing group consensus measure. As the moderator undertakes the aggregation,

it opens the possibility of biased reporting of the participants' rationale for their estimates.

Underlying the Delphi Method as an approach for surfacing a group consensus estimate for an identified problem anonymously from selected participants requires questioning some aspects further (Bolger et al., 2011; Bolger and Wright, 2011):

- The role of the moderator is fundamental to the functioning of the Delphi Method. The moderator controls the questions asked, the feedback provided, and the compilation of the aggregate consensus feedback. Any bias on the part of the moderator could influence the outcome, as the interactions between the moderator and the participants are not observable to the individual group members during the inquiry. The moderator decides after each iteration what feedback is given to the participants, together with the aggregated results.
- Participation is at the election of the group's convener and moderator, with the expectation that the selected participants will be willing to share their expertise on the issue during the inquiry. If some participants are less knowledgeable or unwilling to share all their insights, their contributions may adversely affect the feedback shared, and the consensus outcome realized.
- The success of the Delphi Method is predicated on participants questioning their assumptions and revising their estimates based on the feedback provided by the moderator. The inclination for participants to change their estimates is more likely among those participants who are less sure of their original estimates. The possibility also exists that the more dogmatic participants will be less willing to change their opinions, regardless of the correctness of their views. By contrast, the outlier participants may feel more compelled to change their estimates which may be more accurate, to conform to the consensus opinion, opening the risk of groupthink (Janis, 1973).

Using the *Figure 2-1 Collective Decision-Making Classification*, the Delphi Method could fall between deliberation and collaboration. It has moderate levels of interaction and value judgments. Participants are selected to join the group because of their ability in the project area. Group participation is at the moderator's election, with the participants co-opted into the group. The level of interaction is moderate, although indirect through the moderator. The outcome is the culmination of participants' successive rounds of contribution, with filtered feedback provided by the moderator.

Aggregation is an iterative process through a moderator to achieve an outcome over successive rounds of discussion. The approach aims to reach an acceptable consensus outcome as the moderator and the outside project convener decide. Owing to the lack of direct interaction, the Delphi Method limits the degree of collaboration, with any shared deliberation limited to the interactions with the moderator. The control exercised by the moderator over the feedback and aggregation of the participants' inputs can make the determination of the consensus outcome obscure, limiting the possible benefit of insights held by group members but not necessarily shared with other participants.

From an AR perspective, the lack of transparency in the Delphi Method goes contrary to the idea of collegial fellowship, despite the cycles of action, interaction, reflection, and action (Pedler, 2012). The necessity of anonymity, which is core to the Delphi Method but runs counter to the idea of open collaboration and joint deliberation identified as necessary, must be questioned as essential in environments where autocratic leadership and confidentiality are less significant.

2.7. **Reciprocity**

When a conflict occurs between interested participants, resolving the impasse may require a compromise based on mutually recognizing each participant’s interest in the shared outcome (Fisher, 1989; Ury, 2013). Collaborating to solve a problem, which could be under instruction or by choice to cooperate, is still prone to the agendas each participant brings to the interaction (Fisher et al., 2012). The possibility of fostering cooperation amongst participants based on shared contribution and benefit underpins social interactions dating back to the emergence of communal living (Alvin, 1960). When the participants' interactions are by choice, the dynamic is challenging, as no clear lines of authority may exist to ensure a positive outcome (Cohen and Bradford, 1989). If participants are compelled to interact, reluctance towards cooperation may manifest in opposing the aims of their desired interactions (Fisher, 1989; Ury, 2013). Group interactions are an act of reciprocity that rest on the giving and receiving value. The value of the reciprocity exchanged need not be monetary if it satisfies the interest of each participant involved (Fisher, 1989).

Reciprocity implies that interactions between participants are transactional, with the expectation that comparable value is expected for the value given, if not at the time

of the exchange, then at some future time. Alvin (1960, pp.162), referencing Simmel (1950, pp.387), explains the transactional nature of interactions as "*the reciprocity of service and return service*," for which the "*contacts among men rest on the schema of giving and returning the equivalence*." Fisher (1989) explains the need for understanding participants' interests and sharing the benefit in a mutually beneficial manner, such that each receives value, an act of reciprocity (Fisher et al., 2012, Ury, 2013). The need for reciprocity between participants involved in achieving a shared outcome is identified by Rijshouwer (2019) as necessary for the effective functioning of Wikipedia.

Reciprocity is central to the exchange between participants as the transactional requirement that gains their commitment and has implications for structuring group interactions. In terms of reciprocity and its relevance to group interactions, several factors for improving the interactions can be identified (Alvin, 1960; Badaracco, 1992; Cennamo et al., 2009; Cohen and Bradford, 1989; Fisher, 1989; Fisher et al., 2012; Goodpaster, 1991):

- The exchange between the participants needs to be voluntary, without any undue compulsion that compels interaction to the disadvantage of some participants.
- The sharing of value between participants must be balanced over time.
- There must be a shared interest in the outcome of the group by the interacting participants, or the participants will not be willing to work toward the best collective result.
- To get the participants to join the group and continue participating in its interactions, they must feel they will benefit from the time committed and insights gained.
- If the primary value for contributing to the group is not a direct tangible benefit to the participant, the participant would want recognition for their contribution, regardless of when received. The need for explicit recognition implies a group structured with anonymous participation and limited peer recognition for participants' contribution to the outcome is less likely to be effective.

The colloquial phrase, "*what is in it for me*," captures the essence of reciprocity and has implications when structuring group interactions. At the start of forming the group, participants must understand the value they will receive for being part of the process. As group interactions occur, obtaining the value they expect to receive is

necessary. The trend can be seen in many online internet forums with enormous subscription bases, but the top influencers are only a tiny fraction of the total membership. For Wikipedia (2022), the active editors (December 2021 39,052) accounted for about 10.5% of all registered editors (December 2021 370’865), with the top 100 editors accounting for around 1.3 million (~26%) of the 4.97 edits made in the month. Reciprocity was relevant to the AR project consensus panel composed of participants expected to interact to achieve consensus metal price forecasts.

2.8. **Diversity and bias in collective group forecasts**

An expectation of the different group approaches for achieving a reliable consensus outcome is the need for diverse participants to contribute. Aristotle explains the need for diversity: "*some understand one part, and some another, and among them they understand the whole”* (Aristotle and Everson, 1988, pp.66). For Condorcet, the expectation was that participants should be “*more enlightened, the more complicated the question*” to avoid outcomes swayed by the ignorance of the masses (Baker, 1976, pp.156-157). Galton (1908, pp.281) similarly recognized the relevance of ability in achieving the outcome in his experiment, for those “*who were practised in judging weights undoubtedly surpassed that of the voters in ordinary elections.”* For Hayek (1945, pp.520), the importance of the collective “*utilization of knowledge not given to anyone in its totality”* reflected the necessity of diversity to surface an agreed price. For Wikipedia, diversity is reflected in its strategy to “*reach out to developers that are not reached yet to work on a greater diversity of tools in order to get a greater diversity of perspectives*” (Rijshouwer, 2019, p.151).

Batchelor and Dua (1995) proposed aggregating forecasts made using different methodologies, seeing a reduction in the dispersion around the average. The decreased range of outcomes was linked to including diverse participants that used alternative forecasting methods. Armstrong (2001a) supports the importance of diversity and finds that using forecasts based on different methods and data sources can significantly improve forecasting accuracy by reducing the inherent biases of individual participants. For Hong and Page (2012), the “*wisdom of crowds*” is not the fortunate outcome of averaging many estimates offsetting incorrect predictions. Instead, the benefit occurs because of the diversity of predictive models used by participants.

Hong and Page (2012) see diversity as occurring because of the participants' different "*interpreted signals*" received, then analyzed using their unique predictive models. The inference from their dichotomy is that the perspectives held will differ within a crowd. Combining the diversity of insights creates the possibility of achieving an outcome that surpasses that of the individual group members. Page (2007) explains diversity can "*trump*" individual ability because "*collective ability equals individual ability plus diversity*." An assumption made by Hong and Page is that the predictive models of participants are likely to be negatively correlated. The dispersion of the errors around the average is offsetting, resulting in a more reliable consensus outcome. The offsetting distribution of incorrect estimates around the average or median is comparable to Galton's (1907a) and Levy and Peart's (2002) expectations. The assumption of offsetting errors distributed around a central measure ignores the possibility of an inherent bias amongst many participants that may be anchored around a similar value (Bazerman and Moore, 2008; Drummond, 2001).

A common concern in the different group consensus approaches was that the participants' bias could influence the outcome. Discussing Aristotle's understanding of virtue, Susemihl and Hicks (1976, pp. 398) explain that "*crowd emotion would inhibit rather than support virtue,"* causing a biased outcome. Galton (1907a, pp.450) recognizes the possibility of bias, "*The judgments were unbiassed by passion and uninfluenced by oratory and the like."* Condorcet (Baker, 1976, pp.62) considered the possibility of bias, "*provided a society possesses a large number of enlightened men, … an adequate assurance of decisions conformable to truth and reason can be attained*." Hayek's pricing model recognized the existence of individual subjectivity "*the same way as subjective values help the individual to coordinate the parts of his plan."* (1945, p.526). Wikipedia, in its policies, recognizes the possibility of bias by requiring "*articles should be unbiased or written from a neutral point of view*" (Rijshouwer, 2019, pp.172).

The possibility of biases amongst participants in group interactions is likely, with some biases considered more probable than others. The foremost bias because the process involves collective interactions is the risk of groupthink (Janis, 1973), which heightens the possibility of not including all the information held by all group participants (Sunstein, 2006). Another significant bias is "*anchoring*," as participants focus on what has occurred more recently or has been most meaningful to them and

influences their value estimates (Drummond, 2001). Anchoring can also happen because of more exposure to some opinions, which can cause fixation on that view without considering other alternatives. Within groups as within society, more credence is given to the opinion of more outspoken individuals or leaders, resulting in a disregard for the views of others and a fixation on the referential source (Bazerman and Moore, 2008).

Occasionally, the outcome from a collective process can be meaningless, with no definitive preference identifiable. Landemore (2012b) refers to the lack of an ideal outcome as "*the mass of "noise" represented by other people's random opinions*" that obscures the insights held by the better informed, although a smaller number of participants. The phenomenon of "*noise*" mentioned by Landemore can cause a failed outcome from a group process while not a bias. Aristotle wished to limit the decision-making only to the best informed to ensure the emotions of the "*crowd*" did not decide the outcome. Condorcet recognized the possibility of the result being swayed by the preference of the majority, which might not necessarily be the best outcome for the "*enlightened*" voters (Baker, 1976). The recommended statistical solution to uncertain outcomes is to increase the sample size or retest with another group (Armstrong, 2001a). An alternative perspective could be that the central consensus values, such as the average and median, do not consider the environment in which the group interacts. Other consensus measures may more accurately reflect the best estimate.

2.9. **Conclusion: Literature Review**

Collective decision-making involves reaching a consensus through the cooperation of the eligible participants, be it the select few, as in Aristotle's assembly, or through many pooling their choices, for instance, in political elections (Aristotle and Everson, 1988; Galton, 1907a). Anticipating the probable outcome using Condorcet's Jury Theorem relies on the underlying most prevalent preference of the participants (Baker, 1976). The median favored by Galton reflects the central estimate splitting the participants below and above the midpoint (Galton, 1907b; Levy and Peart, 2002). By comparison, the average as a central measure equates to the value with a zero mean square error, which captures the diversity of views (Page, 2007). Hayek (1945) sees market prices as coordinating the interactions of interested participants and their dispersed specific insights. The internet has made it easier for many people to collaborate and achieve a collective outcome, commonly called crowdsourcing

(Brabham, 2013; Howe, 2006). The collective decision-making evolutionary process has retained the elements of deliberation and collaboration in the different approaches, depending on the situational setting (Sunstein, 2006).

From the idealized deliberating possibilities espoused by Aristotle, the "*better argument*" hopes of Habermas (1998), the constitutional underpinning of Western democratic political processes, and modern management practices of teamwork and collaboration, collective decision-making has had mixed successes (Mueller, 2003; Sunstein, 2006). Recognition of the impact group members can have on each other and the existence of disparate power amongst members within the group have been identified partly for the shortcomings of achieving reliable group consensus results (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973). Unfettered group interactions can allow unique knowledge of interacting group members to surface. For collective decision-making to succeed, formulating the rules of interactions that encourage the surfacing of new knowledge is best achieved if the playing field among participants is level and the aggregation of the disparate insights of the participants is impartially done (Rowe and Wright, 2011).

From Condorcet's Jury Theorem and the probability of the outcome of political events to Galton's testing of the general population's ability, the use of averages to decide the result in many collective settings has become ubiquitous. Their methodologies have gained broad support, despite potentially predicting mistaken outcomes (Armstrong, 2001a; Mueller, 2003; Sunstein, 2006). The lesson from group aggregation approaches is that biases can unduly influence the result (Batchelor and Dua, 1995; Shiller, 2005). Depending on the group interactions, the participants can sway the actions of others (Gigone and Hastie, 1997; Schkade et al., 2007; Sunstein, 2006). One of the perils of undue influence and observed remedy in practice is the stratification of participants into subgroups (Armstrong, 2001a). The challenge lies in knowing which subgroup a participant should be in beforehand, and such stratification is often impossible.

Hayek's explanation of how collecting and assimilating dispersed information forms the foundation of market prices makes the possibility of considering using prices to collect and distill information for collective decision-making workable (Berg et al., 2008; Sunstein, 2006). The advantage of using prices to predict outcomes is that the participants' interest is observable and automatically aggregated to reflect the

consensus opinion (Abramowicz, 2004; Hahn and Tetlock, 2006; Hanson, 1999; Levmore, 2002; Wolfers and Zitzewitz, 2004). The use of prices, such as in prediction markets, does have its challenges. As in other markets, an information "bubble" is possible, or even rogue traders, but the expectation would be that the arbitrageurs would ensure the "right" price appears (Klarreich, 2003; Shiller, 2005; Sunstein, 2006). Prediction markets measure the probability of the outcomes without the participants considering the consequences of the possible result (Berg et al., 2008; Sunstein, 2006).

The egalitarian underpinnings of open-source projects undertaken by volunteers, contributing collectively to achieve an outcome, premises broad equality among the contributors (Brabham, 2013; Rijshouwer, 2019; Wales, 2004). Orwell's (2013) "*Animal Farm*" and open-source projects may share some characteristics. For an open-source project, some participants are elevated to a more senior status, setting the rules to be observed and the contributors' consensus outcome reflecting the result of least resistance (Hardin, 1968). Open-source collaboration and its collective participation function within structures with regulations to control the processes while allowing many participants to contribute to the extent they are willing (Reagle, 2010; Rijshouwer, 2019). An important consideration in obtaining the participants' commitment is the need for reciprocity amongst the group members (Möller, 2003; Raymond, 2001).

Considering the interaction and aggregation framework (*Figure 2-1 Collective Decision-Making Classification*), *Table 2-1 Collective Decision-Making Application* summarizes the application of collective decision-making, and the situations best suit the different approaches.

Table 2-1 Collective Decision-Making Application

Collective decision-making application	Circumstances	Aggregation	Participation	Influence (Bias)
Deliberation / Law of Averages	General public Frequent event	Average Median	General Concurrent	Intragroup low External high
Deliberation / Majority opinion	Unique objective Specific setting	Majority vote Consensus[1]	Selective Sequential	Intragroup high External low

Collective decision-making application	Circumstances	Aggregation	Participation	Influence (Bias)
Collaboration / Most favored	Elected objective Unique setting	Consensus[2] Strongest preference	By Choice Sequential	Intragroup high External low
Collaboration / Markets price	Unique time Specific purpose	Price Probability	By choice Concurrent	Intragroup high External low

1. *Lowest common denominator*.
2. *Route of least resistance.*

The aspects of access and aggregation pertaining to collective decision-making provide a framework for considering the application of the approach for surfacing a consensus outcome that captures the diversity of the participants involved and meets with the users' approval of the group result.

CHAPTER THREE – RESEARCH METHODOLOGY

3. Research Methodology

3.1. Introduction

When a workplace task is judgmental, if it is undertaken in collaboration with other team members, the theories and past research would show that the outcome may be more reliable if a group deliberates the problem rather than relying only on a single expert. The AR project aimed to evaluate mining organizations' possible use of collaborative metal price forecasting, which requires a high degree of judgment. A key consideration was to see if the collective outcome would exceed the capabilities of the individual participating group members. Coghlan et al. (2012) 's interactive cycles were reflected in the AR project's design. It allowed the participants to deliberate on their metal price forecasts relative to that of the other group members and how consensus metal price forecasts collectively surfaced could be considered a reliable alternative for future workplace use.

The cooperative inquiry structure, as a case study, used a consensus survey panel with diverse participants with relevant expertise to collaborate in surfacing consensus metal price forecasts (Hong et al., 2016; Landemore, 2012a; Malla, 2016). The AR project fostered collaboration by sharing participants' metal price expectations interactively for reflection by other group members rather than sequentially, similar to a Delphi Method consensus-building approach (Rowe and Wright, 2001). A key aspect of collective decision-making is the recognition that consensus estimate depends on accessing the dispersed insights of diverse participants and using a method to share and aggregate their collective opinions (Armstrong, 2001a; Page, 2007; Surowiecki, 2004; Tetlock and Gardner, 2016).

A central aspect of the action inquiry was examining the significance of collaboration in conjunction with diversity for surfacing collective metal price estimates compared to individual forecasts as substantiation for using a cooperative inquiry approach for tackling the issue within the workplace (Coghlan et al., 2012; Heron and Reason, 2001).

3.2. Action Research

As a collaborative management research approach, the AR project was premised on the active engagement of participants jointly with the researchers to surface a

beneficial management practice and new knowledge. It falls within the field of practical knowing, which aims to influence future actions (Coghlan et al., 2012). The foundations of AR are based on methods that incorporate different elements of collaboration, action, and scientific theory. Across the various modalities of AR, differing expectations about the nature of knowledge, the theory of knowledge, and approaches for uncovering the bases of learning are used, considering the "*nature of how we know and how we inquire into how and what others come to know*" (Coghlan et al., 2012, pp.47). In acknowledging the coexistence of various modalities of AR, the differences are seen as alternative approaches for collaboration, each focusing on different aspects through their questioning and actions.

Practical knowing is context-specific, based on the setting and understandings of those involved, imposing a need to appreciate situational differences (Coghlan et al., 2012). Each situation is an emergent construction of the participants looking to understand the meaning of the situation. AR occurs in the present, referring to the past to understand the origin while looking to change the future through cycles of action and reflection. The general empirical method is realized in the present tense of AR by:

- Experiencing the occurrence consciously and attentively within the situation.
- Understanding the significance of the insights and knowledge within the situation.
- Judging the reasonableness of explanations rationally within the situation.
- Taking action responsibly, considering the consequences of the situation.

When undertaking the human cognitions of experience, understanding, judgment, and action, these occur within the realm of the AR inquiry as a community of practice between the participants, as well as with the researcher, as they collectively inquire into the situation and reflect on the joint actions to take (Coghlan et al., 2012).

Action research as a collaborative management research approach involves participants in interactions to surface their insights within the situation, considering the expectations of the interested stakeholders and their need to engage in joint action to address the challenge (Coghlan et al., 2012). As an approach to practical knowing, AR needs to adapt to each situation and be capable of reacting to the interventions chosen and undertaken. The group interactions aim to surface participants' experiences, consider their insights, develop judgments, and uncover actions to be taken jointly.

Shani and Pasmore (1985, p. 439) give the following definition of AR to exemplify its main characteristics:

"Action research may be defined as an emergent inquiry process in which applied behavioural science knowledge is integrated with existing organizational knowledge and applied to solve real organisational problems. It is simultaneously concerned with bringing about change in organisations, in developing self-help competencies in organisational members and in adding to scientific knowledge. Finally it is an evolving process that is undertaken in a spirit of collaboration and co-inquiry."

An ancillary and important interlinked element of AR generates actionable knowledge through the emergent collaborative inquiry undertaken jointly by participants and researchers into real problems (Coghlan et al., 2012). Tied to the issue of generating actionable knowledge is the point of judging the quality of AR. Considering the definition of Shani and Pasmore (1985), Coghlan et al. (2012) name five elements as necessary:

- The AR inquiry surfaced an understanding of the situational context.
- A high degree of collaboration between participants and the researcher occurred within the AR inquiry.
- During the AR inquiry, active cycles of action and reflection surfaced new meanings for the participants and researcher.
- The AR resulted in sustainable outcomes beneficial to the participants.
- The AR inquiry was able to generate new actionable knowledge.

Action research is situated within a specific organizational context and expects collaboration between practitioners and researchers as a collaborative management research process. It uses accepted methodologies to address existing crucial problems and add to the management knowledge, with collaboration being the cornerstone of the approach (Coghlan et al., 2012). As an essential element, collaboration implies the sharing in the framing, evaluation, assessing, finding, and taking actions. However, such sharing is not necessarily equal for each aspect. Fundamental to collaboration is the collective inquiry into real issues through dialogue between participants and researchers, looking to understand through the joint application of scientific and practical knowledge. Action research aims to transform management practices. The focus can be an individual, a group, an organization, or subsections of the latter two,

seeking ways to address aspects of actions, processes, performance, or coordination. During the AR inquiry, the action of experiencing within the system will lead to the emergence of an understanding. The understanding will, in turn, lead to assessing the judgments made and how taking different actions could eventuate in more effective outcomes, with the cycles of AR captured as new managerial knowledge.

Process management research is similar if the inquiry starts from a collaborative proposal. Still, it is more accurately considered a process of collaboration and intervention, with the action element occurring within the later action inquiry (Coghlan et al., 2012). When the AR inquiry is undertaken with a researcher, the researcher becomes a co-actor, contributing to the management learning that appears. In the action inquiry, the initiation of the research may occur internally as a change initiative, externally as a research proposal, or jointly as an area of mutual interest between the organization and the researcher. The origin of the action inquiry will influence the AR design and focus formulation.

Shani and Pasmore (1985) find four factors to consider when choosing an AR approach:

- The context in which the problem is situated includes both external and internal stakeholders.
- The nature of the relationship between the participants and the researcher's role will influence the quality and effort exerted in the inquiry.
- Structuring how the AR will be undertaken, from its origination, whether the researcher is an insider or outsider, collecting data, formulating the action cycles, and concluding the undertaking.
- How will a practical, workable outcome be achieved for the participants within their workplace while simultaneously creating new management knowledge to be communicated by the researcher?

Action research as an approach to practical knowing with others occurs in the outer world. The recursive AR cycles involve experiencing, understanding, judging, and taking action, manifesting a collaborative, reflective process of surfacing and learning new actionable management knowledge (Coghlan et al., 2012). At the same time, Coghlan et al. (2012, pp.63) propose the notion of interiority that "*involves shifting from what we know to how we know, a process of intellectual self-awareness*." Interiority supplies a system to assess the value of

practical knowing and theory generation within the ambit of AR, recognizing the purpose of the research and the roles performed in the inquiry.

3.3. Action Research Modalities

Shani et al. (2008) see collaborative management research as the "*umbrella term*" for collaboration between practitioners and researchers as well as insiders and outsiders, together forming a community of practice to generate new knowledge and actionable practices (Coghlan, 2010). Raelin (2009) shows several similarities across the action modalities:

- The focus is on the context and collaboration of participants.
- Learning and change occur by participants and within the system processes.
- Reflection-in-action through facilitated reflection.
- Double-loop learning leads to meta-competence.
- Uncertain outcomes surfaced through dialectic interactions.
- Basing the inquiry process on relevant theory and generating functional theory in practice.

Action modalities can be differentiated according to the following:

- The reality of the action.
- The extensiveness of the collaboration between the participants.
- The extent of the participants' interactions.
- The sustainability of the outcome.
- The generation of new managerial knowledge.

For each of these elements, how they are surfaced and included influences the quality of the inquiry. Coghlan (2010, p.151) suggests "*how we know provides a general method*" to understand how we undertake an action inquiry from within the alternative action modalities. In the action modality categorization of Coghlan is the recognition that the knower is "*in here,*" reflecting on contextual issues while also constructing mediated meaning, as experiences trigger questioning, which leads to judging, and informs actions. Raelin (2009) sees AR as the internal processes of learning and knowing occurring within the inquiry, integrating the participants' meanings into what is being inquired about. Coghlan (2010, pp.159) clarifies the distinction between various AR modalities based on their generative insights.

Table 3-1 Action Modalities - Generative Insight and General Empirical Method

Action Modality	Generative Insight	General Empirical Method
Action learning	There can be no learning without action and no (sober and deliberate) action without learning. Those unable to change themselves cannot change what goes on around them.	L = P + Q. Subjecting experience to questioning insight in the company of peers and taking action.
Action science	People are unaware of their theories in use. Systemically analyzing reasoning and behavior to find causal links can produce actionable knowledge.	Testing privately held inferences and attributions in action.
Appreciative inquiry	When people focus on what is valuable in what they do and try to work on how this may be built on, it uses metaphors and conversation to help transformational action.	Attending to insights from the power of positive questioning leads to action.
Clinical inquiry/research	When researchers gain access to organizations at the organization's invitation to be helpful and intervene to enable change, this is the most fruitful way of understanding and changing organizations.	Helping clients attend to their experience, have insights into that experience, make judgments about whether the insights fit the evidence, and then take action.
Cooperative inquiry	Each person is a co-subject in the experience phases by participating in the research activities and a co-researcher in the reflection phases.	Continuing cycles of shared experiences, questions asked, insights generated, meanings articulated and tested in action, further questions and insights tested, understood, and acted on.
Developmental action inquiry	Learning to inquire and act promptly has central and implicit frames that each person acts out of in given periods.	Inquiring-in-action through attending to the four territories of experience.
Learning history	Capturing what individuals and groups have learned and presenting it through the jointly told tale enables readers to know about organizational dynamics.	Attending to and discussing one's questions and insights through reading the experience and insights of others leads to shared understandings.

For the participants involved in an AR inquiry, achieving an outcome based on cycles of action and reflection is a collaborative process. An added perspective of the

AR is to see it as a theory verification and knowledge generation process by the researcher, secondary to the primary aims of the participants. The AR process, when considered in a broader context, involves three perspectives:

- Intentionality: When undertaking an AR inquiry, it is an activity purposefully undertaken to achieve an expected outcome through inviting participants, questioning the problem, taking action, and reflecting on the consequences.
- Expectation: The cognitive behavior during the AR process of questioning, understanding, making judgments, and taking actions, aims to achieve an outcome expected to alleviate the consequences of a problem. Actions taken are to achieve the desired result.
- Divergence: At the start of the AR inquiry, the process is about finding a suboptimal outcome that could be improved. Core to the first AR cycle of questioning, the later cycles of reflection, questioning, and action are concerned with a divergence of the outcomes being experienced from the result judged to be ideal.

Using Coghlan's (2010) perspectives of intentionality, expectation, and divergence, practitioners can group the seven approaches of action modalities.

Table 3-2 Action Modalities

Action Modality	Action Intentionality	Expected Outcome	Reflection on Divergence
Action learning (Revans, 1971; 1998)	Within a select group, question a problem to explore courses of action to address best the issue perceived by the group, fostering learning within the group.	Implement the group's recommended action to test its effectiveness, followed by reflection before retrying another round of questioning and action if unsuccessful.	Resolving the problem depends on the group's interaction effectiveness and their perceived understanding of the problem.
Action science (Argyris, Putnam, & Smith, 1985)	Questioning the theory-in-use instead of the espoused theory in addressing a problem to surface the disconnect between our actions' actual and expected outcomes.	Using double-loop learning, the disparity between theory-in-use and espoused theory is uncovered as the basis for our actions and generates actionable knowledge for	By continuing to experience a disconnect between an expected and actual outcome for a phenomenon, the cause is linked to inadequately questioning the

Action Modality	Action Intentionality	Expected Outcome	Reflection on Divergence
		participants and researchers.	progression from insights and participants' judgment.
Appreciative inquiry (Srivastva & Cooperrider, 1987)	The occurrence of positive phenomena within an organization can be used across other possibilities through metaphors and conversations. Positive questioning can lead to new insights and approaches within the same context.	Based on positive questioning and reflection cycles, moving from experience to insights, then judgments, and leading to new actions based on a positive perspective.	Failure to extrapolate a positive perspective across multiple arenas is because of structural impediments or an inability to have considered the total ambit of the two settings or not fully understanding the differences.
Clinical inquiry/research (Schein, 1987; 1995; 2008)	The client decides on the need for the intervention, the participants to be involved, the desired outcome, and the process to follow. The researcher is "consulted" as an expert to help question the problem jointly with others and find a solution acceptable to the client.	With the help of the researcher/consultant, the organizational issue will be resolved to the client's satisfaction, with the expectation that a workable solution can be found by engaging the proper consultant.	The formulation of the problem, the selection of the participants and the researcher have not been well matched with the inquiry. The crux of the problem is correctly identified, but biases and defensive routines may have hampered the process.
Cooperative inquiry (Heron, 1996; Heron and Reason, 2008)	Collectively participants experiencing the same phenomena question the problem, discuss alternatives, and take action. The activities within the group are shared, with all participants having equal opportunities to contribute to engaging with other group members.	By the participants engaging together to explore a common problem, the expectation would be that their actions would achieve an outcome that captures their shared knowledge and is the most effective resolution to the issue known to the group.	Despite the intentions of the collective participation and actions of the group participants, no assurance can be held that it would include the most effective solution or that the makeup of the group has the ideal composition, or that a shared bias has not blindsided the outcome.

Action Modality	Action Intentionality	Expected Outcome	Reflection on Divergence
Developmental action inquiry (Torbert, 1991; Torbert & Taylor, 2008)	A prompt inquiry that questions beliefs and actions to develop insight and understanding of the individual and organization's actions as they occur, including reflection loops to surface biased assumptions.	The inquiry aims to reduce uncertainty by framing the problem, advocating actions, developing clarity through illustration, and inquiring if others share the same understanding.	The failure to reduce the uncertainty around a problem can be caused by not acting timeously when the data is pertinent, not including all those to whom the issue is relevant, and not questioning actions and assumptions because of biases and firmly held beliefs.
Learning history (Bradbury & Mainmelis, 2001)	Questioning the events from the perspective of those with direct experience of the phenomenon in co-junction with facilitators. The narratives can be juxtaposed to surface-specific and broader understanding among participants.	By reaching out to those directly involved and allowing them to reflect on their experiences and understandings of events, aspects not usually considered will be surfaced, giving a voice to those less able to express their concerns usually.	As events unfold, failure to delve into the impact on those less visible in the context will, by default, present a biased understanding and open the future to unexpected consequences that may cause outcomes different from that intended by the primary action instigators.

The categorization of the AR modalities, the situation, the circumstances, and the expected outcome of the AR intervention will guide which AR modality is best suited to meet the participants' overall aims. Based on the influencing factors, the AR project was most suited to the expectations of a cooperative inquiry.

3.4. Cooperative Inquiry

For the AR project, the focus was centered on the workplace requirement of participants making metal price forecasts often for planning purposes. When forecasting metal prices, the participants could experience a similar phenomenon, expecting that they have correctly considered all the possible influencing metal price factors but fundamentally knowing this could not be possible. Collaborating with other

participants in a comparable position opens the possibility of finding new insights that may positively contribute to future metal price forecasts' reliability. The benefit of transparently and collectively collaborating could allow all the participants to gain other insights and aid in making their future metal price forecasts. A potential shortcoming in using a collective approach with participants having a similar workplace environment is the risk of having common beliefs about future metal prices, manifested in anchoring around the same forecasted metal price levels.

"Cooperative inquiry is a way of working with other people who have similar concerns and interests to yourself, in order to: (1) understand your world, make sense of your life and develop new and creative ways of looking at things; and (2) learn how to act to change things you may want to change and find out how to do things better" (Heron and Reason, 2001, pp.179). As an AR modality, a cooperative inquiry mirrored the concept, spirit, and outcome the AR project hoped to achieve by working with people for their shared benefit in exploring an alternative approach for a challenging workplace task.

Forecasting future metal prices is prone to error. The final metal prices occur based on the subsequent interactions of thousands of market participants, making it improbable that a single or group of individuals could correctly predict all future events and estimate the expected future outcome in advance. Within their workplace practices, those trying to make reliable metal price forecasts share a "*similar concern*" of trying to "*understand the world,*" either acting alone or relying on third-party independent research to make their metal price projections. A possible practical way of sharing insights is for practitioners to find a "*way of working with other people*" with similar workplace requirements. The collaboration would be the "*development of new and creative ways of looking at things,*" opening the possibility of participants "*learning how to act to change things*" and, in the process, "*finding out how to do things better*" together.

Using the AR modality of cooperative inquiry for the AR project was "*research conducted with people rather than on people*" (Heron and Reason, 2001, pp.179). The crux of the AR project was setting up a consensus survey panel of industry experts to work together to surface consensus metal price forecasts. Expanding on the idea of "*with people*," in a cooperative inquiry, the project plan is shared between all the participants and the researcher, so all share in experiencing, understanding, judging,

taking action, and reflection, allowing everyone to influence the process. Action research involves enacting change relevant to a real problem and generating actionable outcomes. Cooperative inquiry achieves its aims because people choose to act together on a problem experienced by them and, in the process, revise their understanding of their world and practices.

For Heron and Reason (2001), a cooperative inquiry has multiple cycles of reflection and action, starting with formulating the focus of the investigations. During the cooperative inquiry, the participants as co-researchers take action as planned, seeing the outcome and reflecting on any divergence from their prior expectations. Based on the observed results and reflections on deviations from expectations, the participants focus on the insights from the experiences within the cooperative inquiry. The culmination of the cooperative inquiry results in the reflection on the insights learned and formulation of how the outcome could inform the future practices of the participants. As an AR modality, a cooperative inquiry is based on repetitive cycles of reflection and action. The distinct phases envisaged within a cooperative inquiry could be seen in a broader context as consisting of successive rounds of cooperative inquiries to expand the reliability of the outcomes, the insights gained, and the workplace practices adopted.

The interested participants are expected to jointly start the cooperative inquiry (Heron and Reason, 2001). The possibility that a few members or single-member co-opt others to join the inquiry process is more likely. Alternatively, the impetus for the cooperative inquiry could come from an outside researcher who looks to gain entry through a shared area of interest. Heron and Reason distinguish if the cooperative inquiry involves the participants' shared experiences and reflections and whether the interactions between participants are limited to those participating in the cooperative inquiry. The AR modalities have recursive cycles of reflection and action that structurally may range from rational and methodical to impassioned and random (Heron and Reason, 2001). A further differentiation shown by Heron and Reason is whether the inquiry is primarily informative about the context and experience of the participants or transformative, focusing on how the participants' experiences changed due to the collaboration.

The researcher's interest in consensus metal price forecasts for the AR project served as the nexus for constituting the cooperative inquiry. Participants with a similar

interest in metal price forecasting were co-opted to join the cooperative inquiry, which excluded other participants who did not wish to engage on the consensus survey panel. The researcher structured how and which metal price forecasts were to be investigated, using an asynchronous online website to allow diverse participants to collaborate. At the commencement of the AR project, the cooperative inquiry participants were informed how they could interact with the other participants, share and deliberate their insights, and take action by revising their metal price forecasts after reflecting on the emerging consensus and market metal prices. The AR project was aimed at collaboration in making consensus metal price forecasts, which were decided by the asynchronous interactions of the participants considering the estimates of their cohorts and the emerging market average metal prices, which required an open boundary approach for the cooperative inquiry. For the AR project, the metal price forecasts needed to fit in with the comparable market metal prices, which required a structured approach to the participants' interactions. As the motivation for the AR project was to consider the workplace practice of forecasting metal prices and explore the possibility of setting up a consensus metal price forecasting panel, the cooperative inquiry was informative with the intention of becoming transformative in the future.

3.5. Research Philosophy – Ontological Position

At a point in time, the actual metal price, e.g., daily close or monthly average, as reported by international commodity exchanges and accepted by all interested stakeholders as the market reference price, is known. At any prior time, the certainty of the future metal price cannot be definitively known, as change can still occur until the final reference point. It is standard business practice for stakeholders involved in the natural resources industry to make estimates of what the future metal prices are going to be for planning purposes. In making their metal price forecasts, the estimators believe, or hope, that the outcome will correspond with their projection, inferring a reality to their outlook that may not eventuate. The relevant metal price forecasts are made within organizations by individuals or groups, who may consult outside sources to inform their estimates. Of all the metal price predictions made in practice, few will be absolutely accurate, with others being reasonably accurate sometimes. The expectation of the metal price forecasters that the actual future market metal prices would be comparable to their prior estimates infers an anchoring of their internal expectations on the expected

future external reality (Bazerman and Moore, 2008; Drummond, 2001). Over the last decade, the impairment for the top forty mining organizations amounted to around a quarter of a trillion United States dollars (PwC, 2021), partly reflecting the divergence between the expected and realized metal prices of mining organizations.

The personal worldview of the metal price forecasters could ontologically be considered realism, as they each believe in their "real" future market metal price estimate (Sankey, 2000). The metal price forecasters infer a realism of their predicted metal prices, hoping they have considered the possibility of all the intervening factors that could influence the eventual outcome (Johnson and Duberley, 2003). In retrospect, some rationalization of why their metal price estimates proved wrong could be given by the metal price forecasters as they try to conceptualize where their paradigms of the expected future reality differed (Gioia and Chittipeddi, 1991; Weick, 1988). Their perceived reality is extrapolated into the future to create an expected outcome that is considered probable, inferring a sense that the conceived result is discernible by the metal price forecasters without an assurance that it will be (Easterby-Smith et al., 2012).

From a relativist ontological perspective, metal price forecasters have differing views about future metal prices. By contrast, it could a particular forecaster would have a dominant expectation (Easterby-Smith et al., 2012). The metal market interactions reflect the unique information held by the many metal market participants, which they have interpreted differently to form their worldviews (Hayek, 1945). The factors influencing metal prices are not likely to be the same, and their significance is uncertain over time. Consequently, individual market participants can perceive the expected interaction of all the factors differently, creating many expectations of future metal price levels (Collins, 1983). The market participants' diverse future metal price expectations imply a relativist ontological perspective. It would only be over time that a convergence of views would occur, resulting in the final market metal prices observable to all. Under a relativist ontological assumption, metal prices emerge from market participants' interaction. However, some participants would have more financial influence than others in deciding the eventual outcome (Easterby-Smith et al., 2012; Knorr-Cetina, 1983).

The cooperative inquiry aimed to surface consensus metal price forecasts from several metal forecasting experts on a consensus survey panel. It was possible and

probable that the participants held different metal price expectations, creating a diversity of ontological perspectives on the consensus survey panel. As a cooperative inquiry participant, the researcher needed a relativist ontological perspective when considering the other participants' metal price forecasts. The ontological perspective reflected the AR project expectations of surfacing a collective consensus metal price from the different participants' perspectives. For the AR project to succeed, it was crucial to allow the participants of the cooperative inquiry to share their insights and surface a consensus metal price forecast. The cooperative inquiry was premised on the individual participants' dichotomous ontological perspectives of realism and the relativist perspective when considering all survey panel participants' consensus metal price forecasts (Easterby-Smith et al., 2012).

3.6. Research Philosophy – Epistemological Position

The currently observable metal price could be perceived as the interactions of thousands of market participants, being the ongoing balancing of the market forces of supply and demand. If the market is in equilibrium, it could be assumed that the current metal price level will continue until some factor changes the market equilibrium, a positivist empirical assessment of market metal prices. Extrapolating the past reflects a positivist interpretation that assumes the past will continue influencing future events similarly (Easterby-Smith et al., 2012). The dynamic of future metal prices could be understood as the interaction of market participants based on their paradigms, which are not entirely known to other market participants, resulting in evolving metal prices determined jointly through the interactions of the market participants (Hayek, 1945). The ongoing interactions of the many market participants cause metal prices to appear as a joint social construction rather than being predetermined by any process, group, or prior understanding of metal price factors (Easterby-Smith et al., 2012).

An ontological perspective of realism coupled with a positivist epistemological paradigm would assume future metal prices are forecastable. However, that would contradict observable outcomes that future metal prices cannot be forecasted consistently (Bek, 2013). The diversity of worldviews held by the metal market participants, including the CM panel members, would imply an ontological perspective of relativism and an epistemological paradigm of constructivism. The

multiple possible future metal price levels converge to a single discernible outcome through the interactions of all the market participants' results and their collective social construction of the market metal prices (Easterby-Smith et al., 2012). The CM panel participants' world views were their individual beliefs that did not represent the shared reality of the consensus metal prices that surfaced during the cooperative inquiry. By imposing a consensus metal price construct, the researcher created an operative paradigm for the research, drawing together the "realities" of the CM panel participants into another worldview (Sankey, 2000). "[W]*hen paradigms change, the world itself changes with them*" (Kuhn, 1970, pp.111).

Examining the structure and context of the AR project, it occurred primarily within a social constructionist paradigm, as described below.

Table 3-3 Structure and Context of the AR Project

Role of the researcher	Participant on the CM panel and acted as the moderator. (Coghlan and Brannick, 2014; Coghlan, 2007; Rowe and Wright, 1999).
Element of interest	Would it be possible to orchestrate a more sustainable metal forecasting practice for organizations in the natural resource industry? (Armstrong, 2001a; Hong and Page, 2012; Sunstein, 2006).
Questioning, action, and reflection	Are the past practices for forecasting metal prices effective compared to a consensus forecast based on the shared insights of the CM panel participants? (Coghlan, 2007; Coghlan and Brannick, 2014).
The research aimed to access	Sharing expectations on the CM panel about future metal prices through sharing quantitative and qualitative insights to collectively surfacing consensus metal price forecasts (Creswell, 2012; Heron and Reason, 2001).
Perspectives accessed	Contrary to the prevailing practices of relying on organizational forecasts, encourage participants to share their metal price expectations and construct a joint consensus forecast as an alternative estimate (Heron and Reason, 2001; Rowe and Wright, 1999).
Unit of analysis	The individual participants and their future metal price predictions, and as the consensus panel deliberations unfolded, focused on the reliability of the alternative consensus estimates compared to the individual participants' forecasts (Armstrong, 2001a; Batchelor and Dua, 1995; Creswell, 2012).
Generalization of the research	Applying the principles of collective decision-making to a workplace problem and creating a forum to share insights and alternative future metal price expectations, explore the possibility of an alternative workplace practice (Heron and Reason, 2001; Hong and Page, 2012; Surowiecki, 2004; Tetlock and Gardner, 2016).

Research sample	An exploratory AR project intended to examine the concept of a consensus metal price forecasting panel focused initially on two metals (Copper & Gold) and a limited number of participants to allow for the completion of the pilot study in a reasonable time (Creswell, 2012).

Under a positivist paradigm, the research focus would have been quantitative, focused on the reliability of historical practices in metal price forecasts made by mining organizations and how these forecasts compared to later market prices (Easterby-Smith et al., 2012). Some industry participants could have considered metal price forecasting indeterminate under a positivist approach. While under a relativist social construction, other industry experts' aggregated metal price forecasts could collectively be regarded as a single forecast (Heron and Reason, 2001).

3.7. Research Methodology

Different approaches are used in the natural resource industry to forecast metal prices, whether for budgeting, longer-term planning, or other strategic purposes. Amongst the more common methods used in practice to forecast metal prices are:

- A historical multiple-year average or cyclical trend such as a moving average (Cuddington, 1992; Genre et al., 2013; Roberts, 2009; Tapia Cortez et al., 2018; Went et al., 2009).
- Specialized research organizations rely on fundamental analysis based on the assumption that the forecasted metal prices will equate to the market equilibrium price level (Hayek, 1945), e.g., The Economist Intelligence Unit (2020), Thomson Reuters GFMS (2020), and Wood Mackenzie (2020). Some larger organizations with the necessary resources could also undertake a similar analysis internally.
- An estimate based on the prevailing Futures metal prices, which are considered by some to represent anticipated future market conditions (Aggarwal et al., 2014; Aruga and Managi, 2011; Chen et al., 2013; Cheng and Xiong, 2013; Chinn and Coibion, 2010; Choi et al., 2015; Millard Fernández, 2016; Frankel, 2014; Gorton and Rouwenhorst, 2004; Goss and Avsar, 2013; Hu and Xiong, 2013; Huchet and Fam, 2016).
- Consensus metal price predictions are prepared by organizations that supply the service based on forecasts obtained from third parties, primarily financial

institutions and researchers, e.g., Consensus Economics (2020) and Oxford Economics (2020). The results are presented as a consensus prediction, and the practice could be considered a statistical group average or median of the survey panelists.

Alternative metal price forecasting approaches have been used to predict future metal price levels in the mining industry. However, over the last decade, the impairments incurred by the top forty mining organizations was a quarter of a trillion United States dollars (PwC, 2021), raising questions about the reliability of the historical metal price forecasting practices used in the industry.

Forecasters face a challenge in making reliable metal price forecasts related to the emergent nature of metal prices. The influencing exogenous factors affecting future metal price expectations are emergent.

- Changing consumption patterns, for example, decreasing or increasing manufactured product usage.
- Changing economic conditions include growth, exchange rates, and trade tariffs.
- Changing resource supply and demand conditions caused, for instance, changing political and economic policies.
- Changing market perceptions in reaction to financial markets contagion (Shiller, 2005) or the influence of speculators due to the financialization of commodities (Adams and Glück, 2015; Cheng and Xiong, 2013; Demirer et al., 2015; Mazur, 2015; Yin et al., 2017).

Some of the considerations used to inform expectations about future metal prices levels are:

- Metal prices track global economic trends (Borin et al., 2012; Crowson, 2018; Fernandez, 2015; Haque et al., 2015; Issler et al., 2014; Stuermer, 2015; Stuermer, 2017).
- Metal supplies are finite at a particular time, and longer-term metal prices should reflect the marginal cost of future production (Deaton and Laroque, 2003, Dobra and Dobra, 2014; Gaugler, 2015; Golding and Campbell, 2014; Humphreys, 2019; Zhang et al., 2015).

- Metal prices are determined by the fundamental factors of supply and demand (Arbatli and Vasishtha, 2012; Beber and Pianay, 2016; Bredin and Poti, 2018; Boudoukh et al., 2007; Brooks et al., 2015; Cifarelli and Paesani, 2013; Gauvin and Rebillard, 2015; Gilbert, 1995; Guzmán and Silva, 2018; Hayek, 1945; Lescaroux, 2009; Tilton, 2018; Wellmer and Scholz, 2017).

Using the same approaches to forecast metal prices could show a preferential information selection bias (Bazerman and Moore, 2008). Explicitly choosing not to consider the added insights of some market participants to the detriment of the reliability of the forecast may be a groupthink ('industry') bias (Janis, 1973). These biases may partially account for the periodic occurrence of significant impairments amongst mining organizations, e.g., from 2013 to 2015 (Drummond, 2001; PwC 2021). The AR project was structured to evaluate how a CM panel with participants from diverse organizations, including the natural resource industry, could collaborate in making metal price forecasts over six months for copper and gold (Creswell, 2012; Heron and Reason, 2001; Rowe and Wright, 1999). By focusing on two metals and seven forecasting periods, the respective outcomes for each metal and forecast period could be evaluated, allowing triangulation of the results within the AR project (Stake, 2003).

A case study was used to evaluate the viability of using a CM panel approach to collectively forecast metal prices as a possible alternative to existing practices in the natural resource industry (Creswell, 2012; Stake, 1995). While the CM panel participants were familiar with existing forecasting approaches, the AR project's uniqueness was the participants' observable collective interactions during the forecasting process. The collaboration allowed the CM panel participants to share their insights and metal price expectations while simultaneously seeing the other CM panel participants' insights and metal price forecasts. The AR project expected that the consensus estimates of the metal prices forecasted would be more reliable than the CM panel participants' estimates. The AR project questioned the assertion that the collective actions of many would outperform the actions of the individual participants (Hong and Page, 2012; Page, 2007). It reflects Surowiecki's (2004) assertion of "*the wisdom **of** crowds*" [Emphasize added], a view commonly advocated for using pooled medians or averages. It contrasts with Perri-Coste's (1907) inference that Sir Galton's (1907a) experiment's accuracy was the possible

outcome of "*the wisdom **in** crowds*" [Emphasize added] rather than being purely the result of good dispersion around the central statistical measures.

The CM panel aimed to evaluate if the participants' expectations about future metal prices were influenced by the other participants' metal price forecasts and insights before making their subsequent predictions (Argyris, 1977; Creswell, 2012; Raelin, 2009; Rowe and Wright, 1999). Specifically, would their interactions on the CM panel cause them to revise their last metal price estimates, improving the consensus metal price forecasts' reliability and helping all the participants (Bolger et al., 2011; Bolger and Wright, 2011; Rowe and Wright, 1999)? The process could have been seen as being cycles of action (forecasting), reflection (on shared forecasts and insights), and action (forecasting) (Coghlan, 2007; Coghlan and Brannick, 2014).

The CM panel approach was a variation of the Delphi Method, with the role of the moderator eliminated by the observable sharing of all participants' interactions, albeit anonymously, in the AR project (Rowe and Wright, 1999). The AR project considered assessing whether the CM panel observable interactions caused the participants to revise their subsequent metal price forecasts. Suppose the observable interactions on the CM panel had minimal influence on the participants' subsequent forecasts. In that case, the significance of collective interactions could be considered as the limited occurrence of group deliberation in surfacing a consensus outcome. Alternatively, the consensus outcome could be considered more appropriate as the group members' aggregated independent insights (Bolger et al., 2011; Bolger and Wright, 2011; Rowe and Wright, 1999).

The CM panel cooperative inquiry perceived the panelists' metal price forecasts as their relativist views while perceiving their subsequent interactions and predictions as their joint social construction. The CM panel reflected the social construction of an aggregated consensus estimate based on participants' interactions and emerging exogenous market events through the metal price forecasts submitted, reflecting an emergent worldview. (Creswell, 2012; Easterby-Smith et al., 2012). The recurrent CM panel's asynchronous cyclical interactions reflected the AR element of the project. The cooperative inquiry explored an alternative approach to existing industry forecasting practices by examining the possibility of shared interactions between participants when making their metal price forecasts using a case study approach (Blatter, 2008; Coghlan, 2007; Coghlan and Brannick, 2014). Examining the process of metal price

forecasting as a dynamic information discovery process allowed the CM panel participants to revise their future metal price expectations. The cooperative inquiry investigated the significance of enabling participants to share their metal price forecasts and insights with their cohorts to inform their joint expectations about future metal prices and the collaborative emergent consensus estimates (Bolger and Wright, 2011; Hong and Page, 2012; Page, 2007; Rowe and Wright, 1999; Surowiecki, 2004; Tetlock and Gardner, 2016). An expectation was that the cyclical metal price forecasts of the CM panel participants would converge on a reliable consensus metal price estimate as an emergent and socially constructed phenomenon through their interactions on the CM panel (Coghlan, 2007; Creswell, 2012; Easterby-Smith et al., 2012).

The cooperative inquiry focused on the workplace practice of forecasting metal prices, which allowed the collecting and collating of the CM panel information that tracked the emergent interactions of participants' metal price forecasts for copper and gold over six months (Blatter, 2008; Creswell, 2012). The AR project included both qualitative and quantitative elements, and it focused on the cyclical metal price forecasts and insights shared and observable by all the CM panel participants (Bolger and Wright, 2011; Rowe and Wright, 1999). The CM panel participants' insights observable by all other participants, albeit anonymously, were explored as a possible contributory factor in the emergence of an aggregated consensus metal price forecast. Could collaborating contribute to a more reliable approach for forecasting metal prices in the workplace (Heron and Reason, 2001)?

The case study allowed multiple participants to take part in making cyclical metal price forecasts collectively (Coghlan, 2007; Easterby-Smith et al., 2012). The elements explored in the case study would not have been possible to the same extent using other qualitative approaches that either focused on only one participant (narrative) or one aspect of forecasting (phenomenology) to achieve the same understanding (Blatter, 2008; Creswell, 2012). Using a case study method allowed a sense of the implications of collaborative metal price forecasting in the workplace and the consequences of cyclical observable interactions for surfacing consensus metal price forecasts.

From a constructionist perspective, using a case study approach enabled the observation of the shared phenomenon of observable cyclical metal price

forecasting and how the CM panel participants' interactions surfaced the aggregated consensus metal price forecasts (Coghlan, 2007; Creswell, 2012; Easterby-Smith et al., 2012). Based on the method adopted, the observations from the case study opened the possibility of examining the theoretical generalizations about the concept of aggregated consensus metal price forecasting (Armstrong, 2008; Bolger et al., 2011, Bolger and Wright, 2011, Rowe and Wright, 1999). Using a case study approach reflected the similarity with the workplace challenges experienced by CM panel participants. It opened the possibility of considering the viability of using an industry survey panel in the future to source collective aggregated consensus metal price forecasts.

Stake (2003) finds six elements that occur in conducting a case study in the context of an AR project.

Table 3-4 Six Elements of a Case Study

Element	Remark
Bounding the case	Making copper and gold prices forecasts over six months.
Research phenomena or issue	Is a collective observable approach for forecasting aggregated consensus metal prices workable in the natural resource industry?
Data to elucidate the issue	The participants' judgmental estimates and explanations submitted over the research period were collected using the CM panel website.
Triangulation to explore the issue	Compare the seven-year copper and gold forecasts to the aggregated consensus estimates and actual market outcomes.
Alternative interpretation	The observable interactions do not inform the participants' cyclical forecasts, and a systematic statistical aggregation method is more reliable.
Assertions and applicability	Information has value, so participants can achieve a more reliable outcome by sharing insights. The assertion rests on the assumption there is a willingness to share insights and consider the alternative perspectives offered.

To Stake (2003, pp.136), "*A case study is both a process of inquiry about the case and the product of that inquiry*." In exploring metal price forecasting as a cooperative inquiry, participants sought recognition for their participation on the CM panel, similar to the need for recognition observed at Wikipedia (Möller, 2003; Rijshouwer, 2019;

Wales, 2004), and reciprocity for the value contributed (Fisher, 1989; Fisher et al., 2012; Ury, 2013).

During the data collection phase of the AR project, from quarter four of 2020 to quarter one of 2021, the COVID-19 restrictions negatively affected the CM panel participation, reducing the number of participants. Invites to participate in the AR project were sent to around 450 potential participants found through LinkedIn (2020). In addition, invites were sent to 62 gold mining organizations and 24 copper mining organizations. Of all the LinkedIn (2020) invites, 115 accepted the invitation to connect. A follow-up email was sent to provide further information on the AR project to secure their participation. Apart from the potential participants, industry business associates were invited to participate in the AR project. The expectation was that this group of familiar colleagues would form a core around which the other participants could be encouraged to participate in the CM panel. Finally, 18 participants registered to forecast the copper price, and 17 participants registered to forecast the gold price. When the AR project was conceptualized, the prior expectation was to have between 15 and 20 active participants for each of the two metals.

Amongst the potential LinkedIn (2020) participants found to take part in the CM panel were several industry experts who regularly take part in the annual LBMA Precious Metals Forecast Survey (LBMA-PMS) (LBMA, 2020). When an invitee declined to participate, he recommended reviewing the LBMA-PMS historical metal price forecasts. On closer inspection of the LBMA-PMS, it had significant similarities to the AR project approach. The main difference is that the LBMA-PMS had no observable participant interactions before forecasting and no later cyclical revised forecasts. The LBMA-PMS data presented the opportunity to triangulate the results from the CM panel, as it covered a more extended period, twenty-one years (Stake, 2003).

LBMA-PMS (2020, pp.26) "*THE RULES OF THE COMPETITION The aim of the annual survey is to predict the average, high and low price range for the year ahead in each metal as accurately as possible. The prediction closest to the average price wins a 1oz gold bar. In the event of a tie, the forecast range is taken into account.*" The participating analysts also had to supply a brief comment for each precious metal they forecast, justifying their expected metal price estimates.

The LBMA-PMS analysts usually make their annual forecast following the first week of January in each year they participate. Before making their forecasts, the forecasters do not have sight of the forecasts of the other participants. After making their forecasts, the survey participants do not have the opportunity to revise or adjust their forecasts based on the predictions of other participants. The construction of the LBMA-PMS effectively yields the average of a statistical group, although the survey does not report the performance of the aggregated consensus estimates. For the LBMA-PMS, no guidance or restrictions are imposed on the forecasts made by the participants. A consideration in participating in the LBMA-PMS is the professional visibility it gives to the participants, "*Published annually, our forecast is the most prestigious survey in the precious metals calendar*," representing the reciprocity for participating.

Certain similarities and differences existed between the AR project CM panel, the Delphi Method, and the LBMA-PMS approaches. The differences stem partly from the application of these approaches in different settings.

Table 3-5 Approach Comparison of CM panel, Delphi Method, and LBMA-PMS

Description	CM panel	Delphi Method	LBMA-PMS
Participation	By invite of the researcher and complete discretion of the participant.	At the request of the arranging organization, with limited discretion of the participant.	The choice of the participant and with the prior consent of the LBMA.
Nature of participants	Participants with a background in the natural resource industry.	Experts with specific industry knowledge.	Experts with acknowledged industry and forecasting expertise.
Identity of participants	Anonymous relative to other participants but known to the researcher/moderator.	Anonymous to other participants but known to the moderator and the organization convener.	Pseudo-anonymous before the publication because prior-year participation is known. Known to the convener and, after publication, known to all.
Nature of participation	Private, specific to the AR project.	Private, within a particular organization.	Open to public participation, which only

Description	CM panel	Delphi Method	LBMA-PMS
			becomes known after publication.
Role of moderator	Facilitating information sharing between participants and observation by all participants' forecasts.	Controlling the distribution of information between participants and end-users.	Collation of forecast for widespread public dissemination.
Communication between participants	Indirectly, through the online portal with all other participants.	Indirect, through the moderator.	None, only possible outside of the LBMA forum after the event.
Number of rounds	Several, at the election of the participant.	Several, at the election of the moderator.	Single (Annual).
Nature of forecasts	Moderator decides on elements to be predicted and forecast parameters controlling inputs.	A moderator decides on elements to be forecasted but could include participant input in a first-round	LBMA specifies elements to be predicted, but participants have complete discretion on the range of forecasts
Visibility of other participants' forecasts	All forecasts and justifications of all participants.	Limited by the moderator to average and partial justifications.	None before publication, then full disclosure.
Aggregation	Mean and Absolute Percentage Error during the project.	Mean or Median at the discretion of the moderator during the project.	Focus on the most accurate individual forecast after the event.
Use of outcome	A research project aimed at informing a workplace practice.	Specific and unique to an organization.	Public relations and industry users.

The assumption of anonymity within the three alternative approaches is questionable. Participants shared opinions about the inquiry may reveal their identities, which the participants are likely to be aware of (Rowe and Wright, 1999). For the three alternative approaches, the role of the moderator is most influential for the Delphi Method in deciding the outcome by controlling the feedback shared with the participants (Dalkey and Helmer, 1963). Comparing the alternative methods of interactions between the participants opens the question of

how pivotal the role of the moderator is in deciding the outcome of the Delphi Method and whether it is necessary for all organizational settings.

Linked to the role of the moderator is the importance of sharing information between successive rounds and the value of having multiple rounds. The role of the moderator under the Delphi Method is to reach a consensus result. The less confident but more correct participants may feel compelled to change their positions to conform to the view shared by the moderator, possibly influenced by the more confident participants (Bolger et al., 2011; Bolger and Wright, 2011). A similar possibility of peer pressure towards conformity could materialize with multiple rounds, biasing the outcome towards the standard view, which may not necessarily be the best estimate (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973; Rowe and Wright, 2011).

Related to the issue of the number of rounds is the degree of ability of the participants. The idea that diversity contributes to the value of collective decision-making permeates the concept of consensus decision-making from Aristotle through to the more recent expectations of Sir Galton (1907a), Surowiecki (2004), Sunstein (2006), Page (2007), Hong and Page (2012), and Tetlock and Gardner (2016). The trade-off between the number of rounds, the ability of participants, and the number of participants go to the fundamental structure of using a collective approach for reaching a consensus outcome. Rowe and Wright (1999) see the ideal design of a Delphi Method inquiry as consisting of three to four rounds, with a possible first exploratory round and with six to eleven participants. The complexity of the inquiry can influence the best structure regarding the number of participants and rounds, as it does if a judgmental estimate or a determinable quantifiable quantity is the subject of the inquiry. If the issue is complex, judgmental, and the outcome is emergent, the possibility of using a consensus approach may be a functional approach, as it taps into the collective insights of several participants simultaneously (Armstrong, 2008; Batchelor and Dua, 1995; Hong and Page, 2012).

For an auction, the collective interaction of several participants decides the final price, with the identity of the other bidders known to the other auction participants. As the auction rounds progress, the number of participants usually tapers off. For an auction to be successful, it is helpful to have multiple participants initially, with the auctioneer (moderator) striving over successive rounds to drive the auction price higher. An auction does not aim to reach a consensus of the bidders but instead targets

the judgmental value of the highest bidder, which is hopefully the same as the seller's expectation. Taking the average of all auction participants' bids across all the auction rounds would not be a fair reflection of the best outcome. The price discovery is visible to all participants, with the actions of the participants driven by their unique judgmental values as envisaged by Hayek (1945).

Taking the average of the last few bidders would more accurately reflect the consensus value of all the auction participants. Supposedly, the bidders in the auction had a preconceived target price in mind when they decided to participate. If it could single out the lead bidders at the auction's beginning and access their price targets, it could supply a reasonable estimate of the auction outcome. Stratifying the auction participants and discarding the lower bidders contradicts a consensus outcome but typifies the auction result more accurately. In most other markets that require a judgmental decision to transact, such as commodity exchanges, the price level at which the transaction occurs is either at the upper or lower limit, depending on the market direction.

A consensus approach challenge is similar and dissimilar to the auction comparison. The auction participants have a common aim, the auctioneer (moderator) controls the process, and the actions of the auction participants culminate in achieving the auction outcome. Unlike some consensus inquiries, the participants in an auction are not anonymous. The auction bidders are both participants and potentially directly affected by the result of their actions, winning or losing the auction. Unlike an auction, the outcome could be positive, negative, or unchanged for consensus price inquiries. The possible range of outcomes makes predicting the expected direction of the result crucial, as potentially, the participants' forecasting errors will not offset to yield a reliable average or median (Levy and Peart, 2002).

The diminishing participation seen in auctions and mirrored in many survey panel approaches is accepted rather than questioned. The average is not the best estimate of the possible auction outcome. Those willing to reflect on the available insights and continue participating will decide the outcome. Commitment to ongoing participation is linked to the reciprocity participants receive from their fellow participants, achieving a result, and for themselves, the possible reward for being the last contributor. And just because a participant disengages from the

process does not necessarily imply they are no longer interested in the outcome, as is noticeable when bidders cease bidding in an auction. They continue to observe the auction to see the result, and survey panelists still consider the survey results after leaving.

3.8. Research Method

The AR project focused on evaluating the use of aggregated consensus metal price forecasts in the natural resource industry, focusing on the comparative reliability of the consensus outcome compared to an individual expert's estimates. As a workplace activity, metal price forecasting is undertaken in practice regularly. The AR project questioned if using an aggregated consensus forecasting approach in conjunction with or as an alternative to existing workplace practices merited consideration (Armstrong, 2001a; Batchelor and Dua, 1995; Hong and Page, 2012; Surowiecki, 2004; Tetlock and Gardner, 2016). As a researcher and participant in the AR project, a challenge in testing the viability of using an aggregated consensus approach in practice was assessing the reliability of aggregated consensus metal price forecasts for use in the workplace in the future. Another element of the AR project was to consider what impact observable interactions had on participants' metal price forecasts and the reliability of the collective consensus outcome (Rowe and Wright, 1999). During the AR project, the cycles of action, reflection, and action were assessed by reporting the aggregated consensus average to evaluate if the cyclical interactions led to reliable consensus metal price forecasts (Coghlan, 2007; Coghlan and Brannick, 2014). Fundamentally, the AR project looked to investigate if participants' interactions caused an improvement in consensus metal price forecasts or if the aggregation of the participants' first-round metal price forecasts could serve as reliable consensus estimates in practice.

A custom website collected quantitative and qualitative forecasting participation data from the CM panel participants. The CM panel website enabled control of participation, collection of the forecasting data, and allowing and encouraging interaction among the participants in making their cyclical metal price forecasts. Although participation was open to all interested industry participants, the focus was on those participants with relevant experience in making metal price forecasts, aiming to assess the importance of "*the wisdom **in** the crowd*" [Emphasize added] (Surowiecki, 2004). The researcher controlled the selection of participants through the

invites sent, which aimed to seek participation from industry experts. Per the University of Liverpool's ethical research requirements, registration on the CM panel website required accepting the ethical terms and conditions for participation in the AR research project before gaining access to the CM panel website. The registration terms and conditions mirrored the details in the invitation to participate in the CM panel (Annexure A: CM panel participant research project background, Annexure B: CM panel participant registration and consent form, and Annexure C: CM panel participant's additional research information).

Alternative approaches for collecting and collating multiple participants' forecasts of metal prices were reviewed in planning the AR project. None of the available internet service providers used a method that met the planned AR project's quantitative and qualitative data collection approaches. In considering the alternative service providers, the researcher registered on the Estimize (2020) website in 2018 and took part as a forecaster making around 500 estimates. Estimize's consensus data collection method informed the structuring of the CM panel website (Estimize 2020). Some significant differences were used in the approach taken for the AR project, using the CM panel website, compared to the Estimize (2020) process.

- Unlike the Estimize (2020) approach of allowing interested participants to register and submit forecasts, the CM panel was by invitation and focused on industry participants interested in metal price forecasts.
- The Estimize (2020) approach requires the participants to make a point estimate of the relevant economic measures. By contrast, the CM panel participants needed to submit probabilities for uniformly distributed metal prices centered around the prevailing market metal price at the inception of each forecasting period.
- The Estimize (2020) website displays the market consensus average, the range of estimates submitted previously, and the historical trend for the relevant estimate before the participant submits their point prediction. The daily appropriate prior month's copper and gold Future's metal prices were always graphically available on the website for the CM panelists. Before a participant had to submit their metal price forecast, the prevailing average probability distribution of metal price was graphically displayed on the forecasting input webpage. In addition, the most

recent justifications of participants were shown, allowing them to see other participants' rationale before making their metal price forecasts.

- A significant difference compared to the Estimize (2020) approach and a definitive aspect of the AR project was requesting all CM panel participants provide a brief explanation (maximum of 150 words) to justify their metal price forecast. In keeping with the aims of a cooperative inquiry, the CM panel participants could observe and reflect on the other participants' metal price forecasts and insights before making their cyclical metal price forecasts. Provision was made on the CM panel website to allow participants to comment on their fellow cohorts' submissions to encourage interactions amongst participants further.
- Like Estimize (2020), a ranking system encouraged the CM panel participants to make cyclical and accurate forecasts. The CM panel ranking methodology encouraged participants to submit cyclical predictions based on the CM panel interactions.
- The CM panel website made it possible for the participants to review their own and other participants forecasting history, including the interactions that had occurred previously, to encourage reflection and action. The CM panel transparency enabled cycles of reflection and action (Coghlan, 2007; Coghlan and Brannick, 2014).

While the CM panel approach had similarities to the Estimize (2020) participation process, it included transparency of CM panel participants' justifications and the ability for the participants to reflect on the emerging consensus metal price and resubmit their metal price forecasts. The cyclical cooperative inquiry element allowed the CM panel participants to reflect on the actions of their cohorts before acting and making their subsequent metal price forecasts (Coghlan, 2007; Coghlan and Brannick, 2014). To further encourage reflection and action, and create a neutral learning environment, anonymous participation on the CM panel was used by having participants register and take part using pseudonyms. The element of anonymity is like that used by Estimize (2020) and emulates the principle of anonymity recommended for the Delphi Method (Dalkey and Helmer, 1963; Rowe and Wright, 1999).

In striving for diversity on the CM panel, participants from different world geographic areas and professional backgrounds were invited to participate. Each participant could make asynchronous metal price forecasts, allowing other CM panel

participants to see their estimates and associated justifications before making their estimates and comments for reflection by the CM panel participants. The focus throughout was to allow sharing of insights and expectations and, through seeing and reflecting on the actions of the other CM panel participants, subsequently make revised metal price forecasts. The CM panel aimed to use collective observable interactions to surface a consensus metal price forecast that could be useful within the workplace of the CM panel participants and explore the possibility of using the approach in a broader industry setting in the future.

The cooperative inquiry investigated the viability of using consensus metal price forecasts in practice. The CM panel structure ensured total transparency of all participants' forecasts and justifications by all registered CM panel participants. The relevant metal price forecasts consensus averages for each forecasting period were accessible to all participants without the necessity of intervention by the moderator. The aggregation process was automated, as was the ability to access all previous metal price forecasts and their matching justifications, minimizing the role of the moderator. The automated aggregation of participants' metal price forecasts ensured that the consensus average metal prices were continuously available to all participants. The ability to constantly observe the interactions of the CM panel participants was to encourage reflection and action by allowing ongoing cycles of metal price forecasts. The structure was similar to the Estimize (2020) approach in its transparency while differing from the Delphi Method, where the role of the moderator is more substantial in controlling the flow of information.

Participants were requested to supply some demographic information when they accepted the invitation to participate in the CM panel without compromising their anonymity to assess CM panel members' diversity. The requested demographical information focused on the background of the CM panel participants.

Table 3-6 CM Requested Participant's Demographic Information

Age (Years)	25<	25-35	35-45	45-55	55-65	>65
Sex	Male	Female	Blank			

Geographic location	Africa	Asia	Australia	Europe	North America / Canada	South America
Profession background	Financial	Human Resources	Journalism	Legal	Technical	Other
Workplace	Academic	Consultancy	Financial institution	Mining organization	Technical services	Research organization
Forecasting approach	Econometric modeling	Fundamental analysis	Futures prices	Historical averages	Sourced from 3rd parties	Technical analysis

The motivation for requesting the demographic information of participants was to assess the significance of diversity, with the expectation that the difference in insights would encourage more cyclical forecasts and improve the reliability of the aggregated consensus metal price forecasts (Armstrong, 2008; Batchelor and Dua, 1995; Hong and Page, 2012; Page, 2007).

The CM panel justification requirement aimed to evaluate the factors perceived as informative by participants in forecasting metal prices (McKay and Marshall, 2001; Rowe and Wright, 1999). Allowing interactions as a part of the CM panel enabled the participants to interact, share their insights, and reflect on the opinions of other participants (Coghlan, 2007; Coghlan and Brannick, 2014). For the cooperative inquiry, it made it possible to assess how collective interactions influenced the emergence of the aggregated consensus metal price forecasts and inform the importance of interactions when considering using a similar approach as an alternative workplace practice in the future (Page, 2007; Rowe and Wright, 1999). Considering the CM panel forecasted metal prices reported in the relevant metal markets, the website included a daily update of the market metal prices for the applicable forecasting periods to encourage additional recursive forecasts (Creswell, 2012; Easterby-Smith et al., 2012).

Figure 3-1 CM panel Copper Forecast Input Web Page shows an example of the Copper Forecast Input page from the CM panel website.

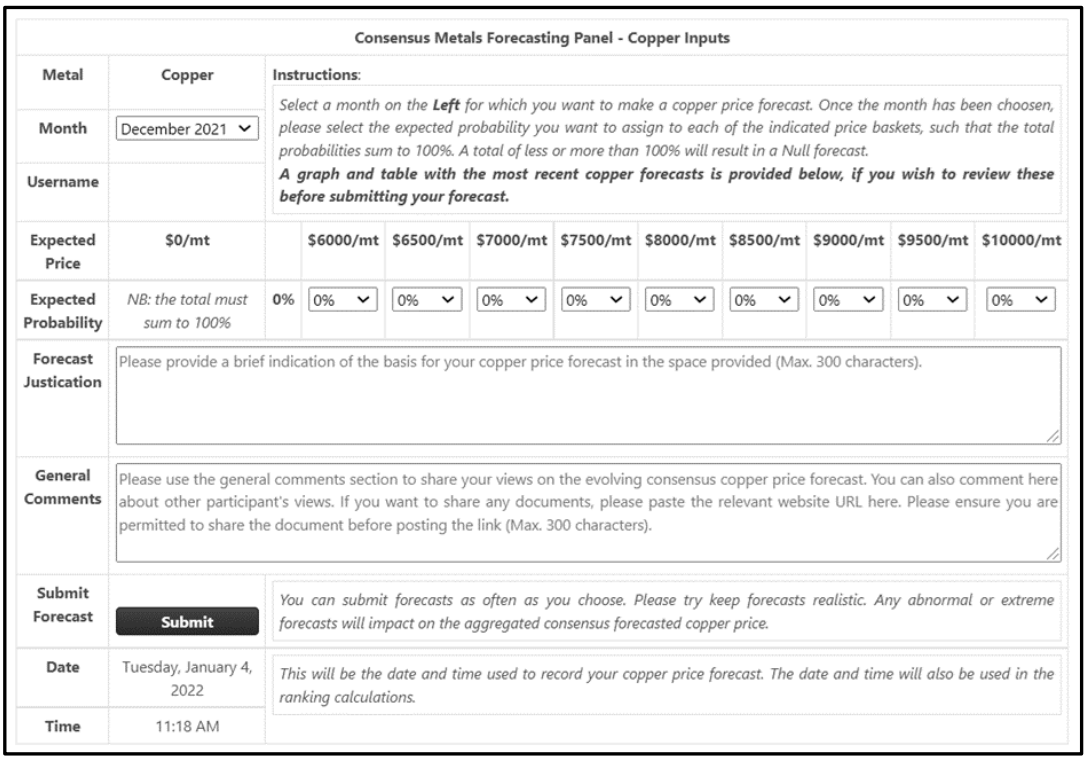


Figure 3-1 CM panel Copper Forecast Input Web Page

The CM panel Forecast Input web page included a table and graph that allowed participants to see the current aggregated consensus results before making their metal price forecasts.

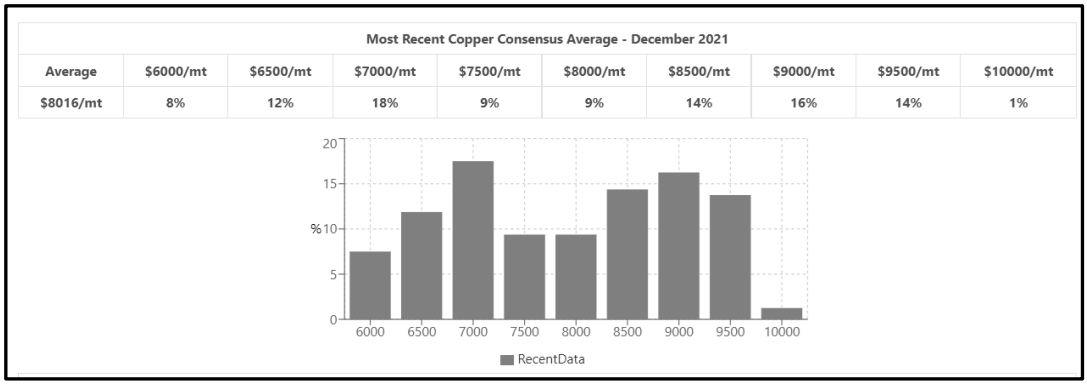


Figure 3-2 CM panel Copper Input Consensus Average Graph

And to allow for reflection and asynchronous interactions, the CM panel website included a summary of the most recent metal price forecasts and the matching justifications.

Most Recent Copper Forecasts - December 2021										
Expected Value	$6000/mt	$6500/mt	$7000/mt	$7500/mt	$8000/mt	$8500/mt	$9000/mt	$9500/mt	$10000/mt	Justifications
$6575/mt	20%	45%	35%	0%	0%	0%	0%	0%	0%	A slightly lower forcats - following on from last quarter as optimism on a global recovery continues to wane and another northern hemisphere winter of living with COVID-19 - even with a vaccine it will not be available in the numbers required for everyone
$7725/mt	0%	0%	20%	30%	35%	15%	0%	0%	0%	Run up in prices again in the latter part of the year, with buoyant economic outlooks across the globe.
$8845/mt	0%	0%	0%	0%	10%	30%	40%	20%	0%	
$7075/mt	0%	10%	65%	25%	0%	0%	0%	0%	0%	As the year draws to a close and COVID-19 is less of a daunting prospect, think the actual state of the world economy will be foremost in policymakers' decisions. Given the constraints faced in implementing fiscal stimulus measures, think the copper price will reflect the true weaker state of the global economy.
$8900/mt	0%	0%	0%	0%	0%	20%	80%	0%	0%	1st world basically vaccinated. Vaccine passports in operation. Travel industry starting up. Developing world starting to vaccinate. US economy stable, and improving. UK and Europe returning to normal. China continual push on Green energy rollouts, creating push for copper.

Figure 3-3 CM panel Summary of Participants' Copper Forecasts

The CM panel website included the daily Futures metal prices (CME Group, 2020) to inform participants' forecasts and encourage recursive metal price forecasts.

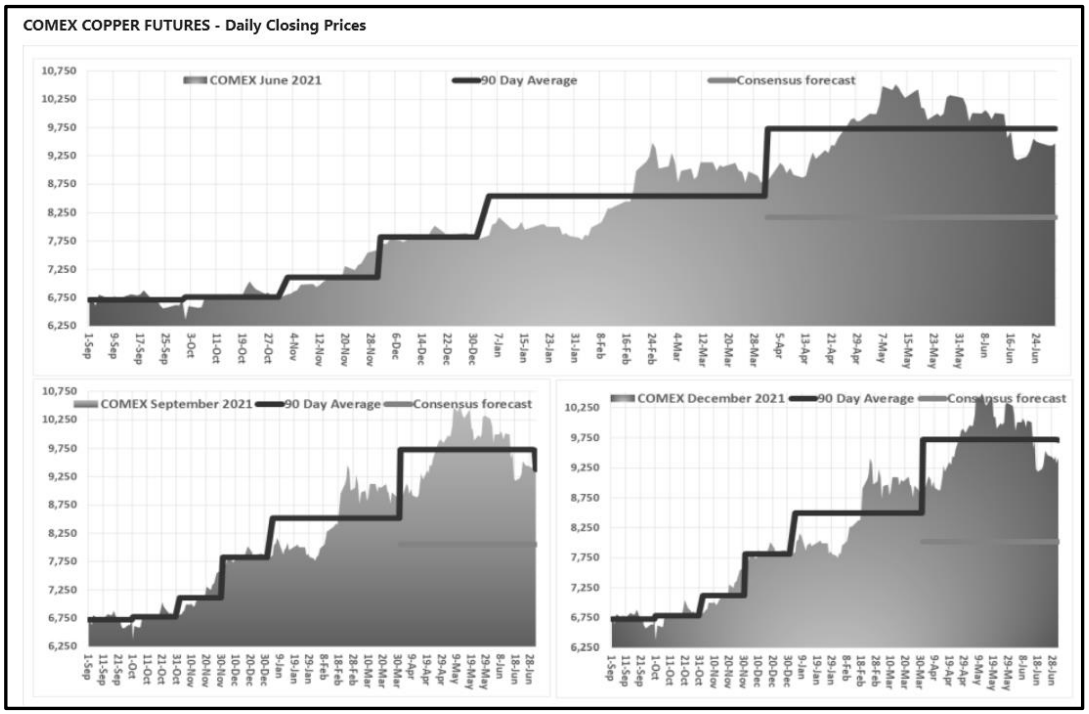
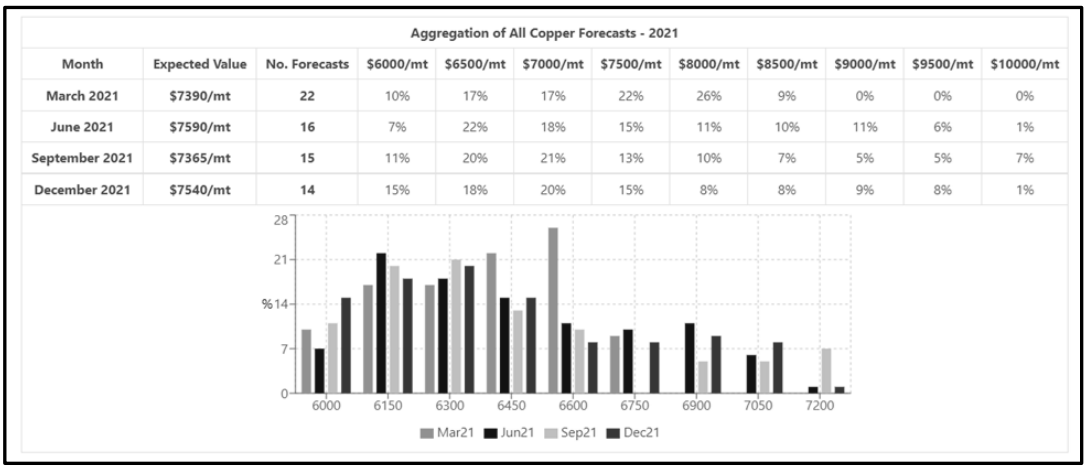


Figure 3-4 CM panel Copper Futures price graph

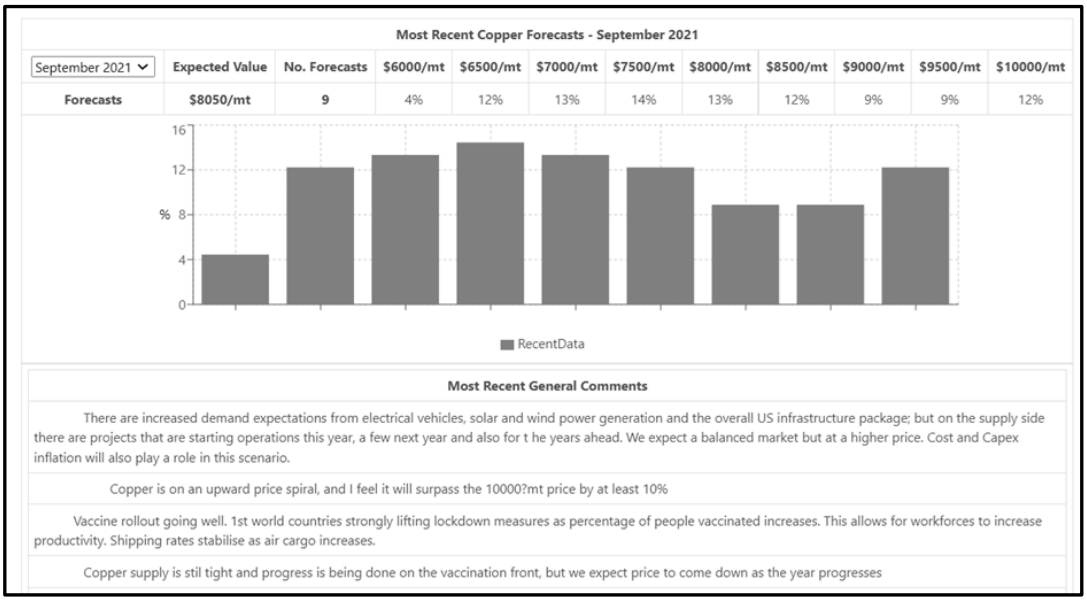
The CM panel website had several aggregation web pages that allowed participants to track the ongoing aggregation of the consensus forecasts to encourage reflection and recursive metal price forecasts.



Aggregation of All Copper Forecasts - 2021

Month	Expected Value	No. Forecasts	$6000/mt	$6500/mt	$7000/mt	$7500/mt	$8000/mt	$8500/mt	$9000/mt	$9500/mt	$10000/mt
March 2021	$7390/mt	22	10%	17%	17%	22%	26%	9%	0%	0%	0%
June 2021	$7590/mt	16	7%	22%	18%	15%	11%	10%	11%	6%	1%
September 2021	$7365/mt	15	11%	20%	21%	13%	10%	7%	5%	5%	7%
December 2021	$7540/mt	14	15%	18%	20%	15%	8%	8%	9%	8%	1%

Figure 3-5 CM panel Copper Aggregation Table and Graph



Most Recent Copper Forecasts - September 2021

September 2021	Expected Value	No. Forecasts	$6000/mt	$6500/mt	$7000/mt	$7500/mt	$8000/mt	$8500/mt	$9000/mt	$9500/mt	$10000/mt
Forecasts	$8050/mt	9	4%	12%	13%	14%	13%	12%	9%	9%	12%

Most Recent General Comments

There are increased demand expectations from electrical vehicles, solar and wind power generation and the overall US infrastructure package; but on the supply side there are projects that are starting operations this year, a few next year and also for t he years ahead. We expect a balanced market but at a higher price. Cost and Capex inflation will also play a role in this scenario.

Copper is on an upward price spiral, and I feel it will surpass the 10000?mt price by at least 10%

Vaccine rollout going well. 1st world countries strongly lifting lockdown measures as percentage of people vaccinated increases. This allows for workforces to increase productivity. Shipping rates stabilise as air cargo increases.

Copper supply is stil tight and progress is being done on the vaccination front, but we expect price to come down as the year progresses

Figure 3-6 CM panel September Copper Recent Forecasts Graph and Comments

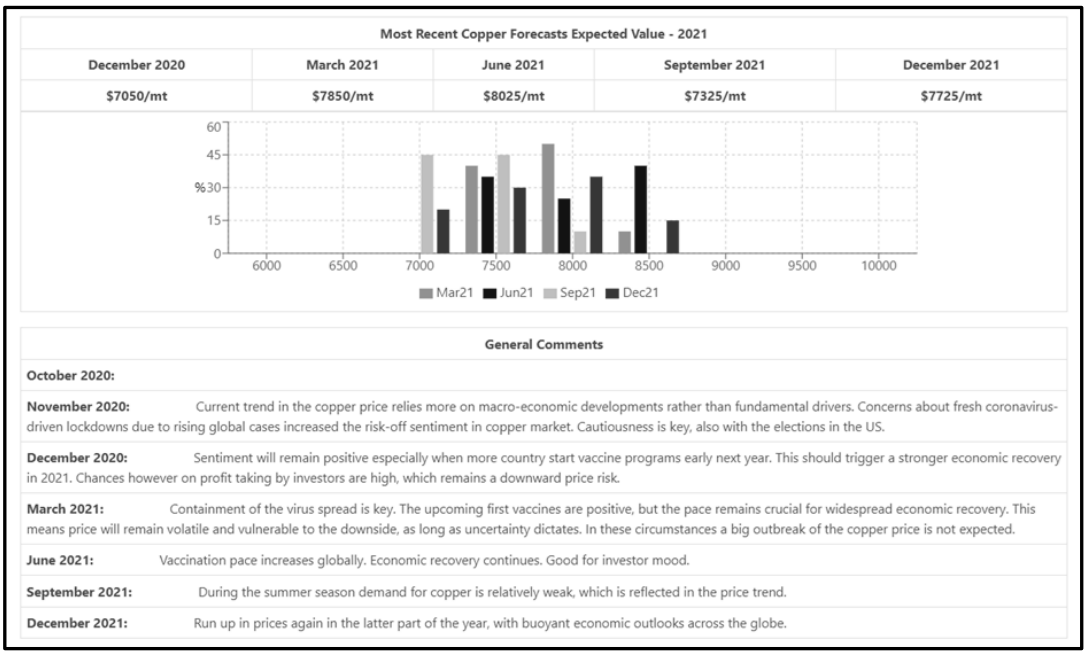


Figure 3-7 CM panel Copper 2021 Aggregated Expected Value Table and Graph

To encourage the CM panel participants to make revised cyclical metal price forecasts, the forecasts of all participants were available on the CM panel website for consideration by the participants. By making all prior CM panel participants' predictions observable and the latest aggregated consensus average, the cooperative inquiry encouraged adopting the aggregated consensus approach as an alternative workplace practice.

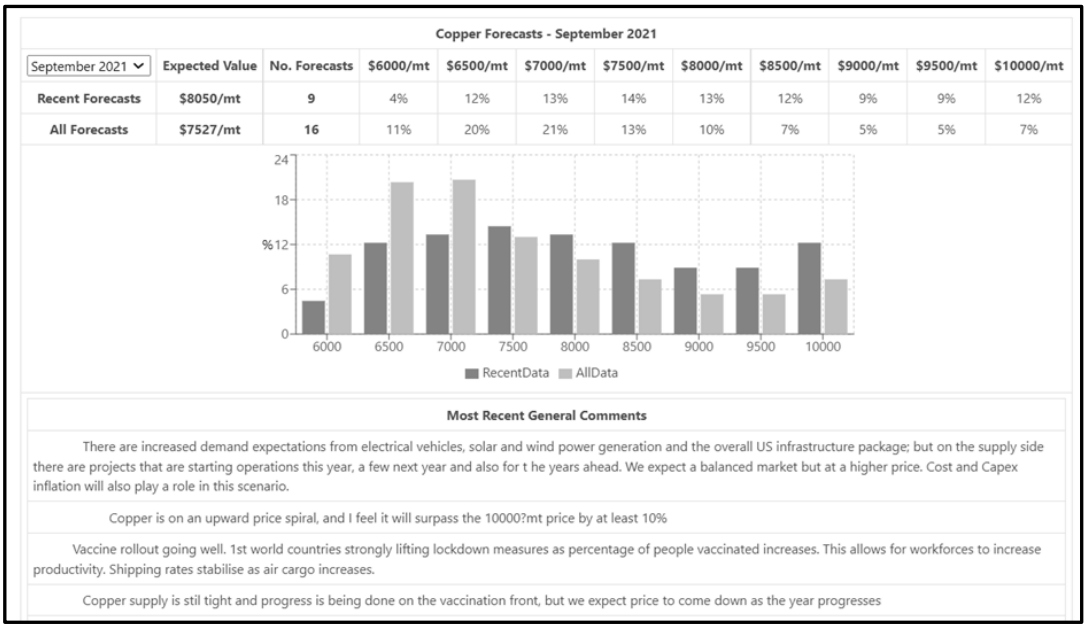


Figure 3-8 CM panel September 2021 Copper Graph and Participants' Forecasts

Most Recent Copper Forecasts - September 2021

Expected Value	$6000/mt	$6500/mt	$7000/mt	$7500/mt	$8000/mt	$8500/mt	$9000/mt	$9500/mt	$10000/mt	Justifications
$6725/mt Wed,21 Oct 2020 05:29 am	10%	45%	35%	10%	0%	0%	0%	0%	0%	Slightly more negative bias in this quarter as some of the optimism wanes and the realisation that economic recovery will be slow despite low interest rates and government spending.
$8850/mt Fri,09 Apr 2021 13:14 pm	0%	0%	0%	0%	0%	30%	70%	0%	0%	Vaccine rollout going well. 1st world countries strongly lifting lockdown measures as percentage of people vaccinated increases. This allows for workforces to increase productivity. Shipping rates stabilise as air cargo increases.
$6525/mt Mon,11 Jan 2021 19:14 pm	30%	45%	15%	10%	0%	0%	0%	0%	0%	
$8100/mt Mon,11 Jan 2021 03:11 am	0%	0%	0%	10%	60%	30%	0%	0%	0%	I believe that copper is going to continue to be in demand all year and supply is really going to struggle. At the end of summer and as people start to build stocks for winter, there may be real shortages.
$10000/mt Thu,27 May 2021 15:55 pm	0%	0%	0%	0%	0%	0%	0%	0%	100%	Copper is on an upward price spiral, and I feel it will surpass the 10000?mt price by at least 10%

Older Copper Forecasts - September 2021

Expected Value	$6000/mt	$6500/mt	$7000/mt	$7500/mt	$8000/mt	$8500/mt	$9000/mt	$9500/mt	$10000/mt	Justifications
$7175/mt Thu,01 Oct 2020 11:13 am	0%	10%	55%	25%	10%	0%	0%	0%	0%	Dr Copper starts to run!
$6925/mt Mon,05 Oct 2020 10:25 am	10%	25%	45%	10%	10%	0%	0%	0%	0%	World demand keeps steadily rising at a muted rate but pressure is upwards on prices.
$6400/mt Wed,28 Oct 2020 19:47 pm	40%	40%	20%	0%	0%	0%	0%	0%	0%	
$6450/mt Thu,29 Oct 2020 19:17 pm	30%	50%	20%	0%	0%	0%	0%	0%	0%	With the economic consequences of the COVID-induced slowdown becoming more widespread, the demand for commodities will decline while production remains steady. As a consequence expect downward pressure on the copper price.

Figure 3-9 CM panel September 2021 Copper Participants' Prior Forecasts

The final element on the CM panel website that aimed at encouraging recursive forecasts based on the information available to the CM panel participants was ranking the cyclical estimates at the end of each forecasting period.

Consensus Metals Copper - March 2021 Rankings

The average LME daily price for copper quarter one 2021 was USD8'504/mt. The average of all the consensus copper price forecasts submitted for the quarter was USD7'383/mt. For only the most recent predictions made for the quarter, the average consensus copper price forecasted was USD7'695/mt, reflecting the sharp upward movement in the copper price in 2021. The general narrative around copper was more positive as references to COVID vaccine rollouts became more frequent. However, there was a split amongst the forecasters on how the vaccine availability would translate into the global economy's recovery. Some participants were cautious in their expectations, anticipating a sluggish recovery. Conversely, other participants saw the likelihood of growing demand and fiscal stimulus as strong supporters of a higher copper price. What is evident from the forecasts, particularly those made in 2020, participants were caught unaware by the substantial rise in the copper price. It will be interesting to see how participants revise their forecasts for the remainder of 2021.

The total number of forecasts submitted was 19 out of the 20 current registered participants.

The copper rankings for March 2021 quarter were as follows, after taking into consideration the date when the forecasts were submitted:

Participant	Forecast date	Forecast price	Difference	Unadjusted Score	Adjusted Score	Rank
	31-Dec-20	$8175/mt	$329/mt	96.1%	96.1%	1
	23-Dec-20	$7875/mt	$629/mt	92.6%	92.6%	2
	21-Dec-20	$7850/mt	$654/mt	92.3%	92.3%	3
	11-Jan-21	$7950/mt	$554/mt	93.5%	91.7%	4
	13-Jan-21	$7950/mt	$554/mt	93.5%	91.3%	5
	11-Jan-21	$7775/mt	$729/mt	91.4%	89.6%	6
	27-Jan-21	$7950/mt	$554/mt	93.5%	89.0%	7
	20-Dec-20	$7525/mt	$979/mt	88.5%	88.5%	8
	31-Jan-21	$7700/mt	$804/mt	90.5%	88.4%	9
	2-Dec-20	$7425/mt	$1079/mt	87.3%	87.3%	10
	5-Mar-21	$8200/mt	$304/mt	96.4%	85.8%	11

Figure 3-10 CM panel Copper Rankings March 2021

Approximately 500 invites were sent in the middle of 2020 to potential participants to participate in the AR project. Initially, over 100 invitees had expressed an interest in the AR project. Later in 2020, many organizations implemented work-from-home arrangements due to COVID-19. Consequently, the later registration and participation on the CM panel were affected, with 26 registered participants and 12 active participants rather than the expected 15 to 20 active participants per metal (gold and copper). Sourcing the historical LBMA Precious Metals Survey (LBMA-PMS), data allowed triangulation with the metal price forecasts collected from the CM panel participants.

The main difference between the CM panel and LBMA-PMS forecasts was that they were made annually and involved no prior official interactions between the participants or the opportunity to revise their forecasts later. The annual gold LBMA-PMS forecast between 2000 and 2020 was sourced directly from the LBMA (2021), with most yearly surveys readily available on their website. Below is an example of the LBMA-PMS gold forecasts made in 2020. The annual LBMA-PMS requires each participant to forecast the estimated maximum, minimum, and yearly average of the daily precious metal price fixes reported by the LBMA (2021). In addition to the numerical estimates of the metal prices, the LBMA participants must briefly explain their metal price forecasts (Annexure E: LBMA-PMS 2001 forecasts). The

aggregation of the LBMA-PMS was tabulated and reported compared to the forecasted January gold price at the beginning of each year.

Table 3-7 LBMA-PMS 2001 Participants' Gold Forecasts

No	Name	Company	City	High	Low	Average
1	Arnold, Ted	*Prudential-Bache Int. Ltd.*	London	$283	$250	$268
2	Bianchini, Egizio	*Nesbitt Burns*	Toronto	$315	$260	$280
3	Christian, Jeffrey	*CPM Group*	New York	$314	$255	$294
4	Crisp, Kevin	*Credit Suisse First Boston*	London	$300	$250	$275
5	Cross, Jessica	*Virtual Metals*	London	$310	$250	$283
6	Fewings, Martin	*NM Rothschild & Sons*	London	$300	$250	$280
7	Goode, Keith	*Bell Securities*	Sydney	$290	$270	$277
8	Hawkes, Neil	*CRU International*	London	$290	$250	$265
9	Klapwijk, Philip	*GFMS*	London	$290	$252	$271
10	Laserre, Frédéric	*SG*	Paris	$300	$250	$270
11	Levine, Howard	*Bear Sterns & Co*	New York	$320	$260	$289
12	Mallalieu, David	*Scotia Capital Inc.*	Toronto	$300	$250	$286
13	McConvey, Daniel	*Goldman Sachs*	New York	$295	$250	$275
14	Murenbeeld, Martin	*M. Murenbeeld & Assoc.*	Vancouver	$324	$256	$297
15	Naqvi, Kamal	*Macquarie Bank*	London	$300	$240	$275
16	Norrish, Kevin	*Barclays Capital*	London	$295	$250	$267
17	O'Connell, Rhona	*Canaccord Capital (Europe)*	London	$325	$265	$285
18	Panizutti, Frédéric	*MKS Finance SA*	Geneva	$320	$260	$282
19	Reade, John	*UBS Warburg*	London	$310	$260	$282.5
20	Rhodes, Jeffrey	*Standard Bank London*	Dubai	$295.65	$253.50	$272.40
21	Rijnbeek, Hennie	*Rabobank International*	London	$295	$255	$268
22	Smith, Andy	*Mitsui & Co.*	London	$290	$210	$250
23	Takai, Bob	*Sumitomo Corp*	Tokyo	$285	$255	$265
24	Ward, Peter	*Lehman Bros.*	New York	$295	$240	$265
	AVERAGE			**$301.74**	**$251.31**	**$275.91**

3.9. Data Analysis

One of the expectations in structuring the AR project and collecting metal price forecasting data on the CM panel was that the participants' cyclical interactions would result in more reliable metal price forecasts (Armstrong, 2008; Coghlan, 2007; Coghlan and Brannick, 2014; Rowe and Wright, 1999). A complication in testing the emerging reliability of forecasts from October 2020 to June 2021 was the increased level of metal price volatility in response to the economic uncertainty resulting from the COVID-19 outbreak. The annual LBMA-PMS data collected did not have the recursive forecasts within a single year, making the market volatility less of a factor in assessing the reliability of these

forecasts. A factor that came to the fore in testing the reliability of both sets of metal price forecasts was correctly predicting the future market metal price direction. Suppose the forecasters incorrectly predict the metal price direction. In that case, the assumption of the participants' forecasts being evenly distributed around a central estimate, such as the median, becomes less applicable.

In assessing the quantitative reliability of the CM panel metal price forecasts, the AR project also evaluated the importance of the participants' forecast justifications. A section was included on the CM panel website for the participants to justify their metal price forecasts and comment on the justifications of other participants and their associated metal price forecasts. The expectation was that their cyclical interactions would lead to revised forecasts reflecting their insights from the observable interactions of the diversified group of participants. The moderator controls the participants' shared interactions compared to the Delphi Method. No such filter was applicable for the CM panel interactions, and the CM panel participants had the opportunity to comment on their cohorts' views.

Based on the theory of collective decision-making, spanning from Aristotle's 4th-century summation views to the 20th and 21st-century aspects of aggregation, deliberation, and collaboration, the metal price forecasters could reflect on the shared insights of their peers and use the information to inform their predictions (Armstrong, 2001a; Baker, 1976; Batchelor and Dua, 1995; Aristotle and Everson, 1988; Galton, 1907a; Hayek, 1945; Hong and Page, 2012; Rowe and Wright, 1999; Sunstein, 2006; Surowiecki, 2004; Tetlock and Gardner, 2016).

An element of the AR project was assessing the reliability of consensus metal price forecasts to assess if it was practical as an alternative workplace practice. Comparing forecasts across different periods and metals is challenging because the absolute forecasting errors are not comparable, as the scale can change. The average LME copper price for October 2020 was USD6'703/metric ton, and for March 2021 was USD8'504/metric ton, a 27% increase over six months. For the LBMA-PMS, the average consensus forecast gold price in 2000 was USD297.83/ounce. In 2020, the LBMA-PMS average gold forecast was USD1'558.82/ounce, the peak occurring in 2012 with an average of USD1'766.23/ounce. The fivefold increase in the gold price over the twenty-one years would make comparisons of measures based on the absolute unit of

measurement incomparable. The Mean Square Error (MSE), calculated as the squared difference between the forecast and the average market metal price, would increase as the gold price increased. At the same time, the relative forecasting error could be less. Using the Absolute Percentage Error (APE) eliminated the scale in measuring the forecaster's reliability error. The APE is the absolute difference, ignoring positive or negative, between the forecast and the actual average market metal price divided by the actual market metal price to yield a consistent scale-independent measure of forecasting accuracy.

In analyzing the data from LBMA-PMS to identify trends in the APE of individual forecasters, specifically the aggregated consensus average, no noticeable or discernible trend could be observed. The graph below shows the APE of the LBMA-PMS forecasters for the twenty-one years, and what is evident is the erratic trend for the individual participants and the aggregated consensus average (LBMA, 2020).

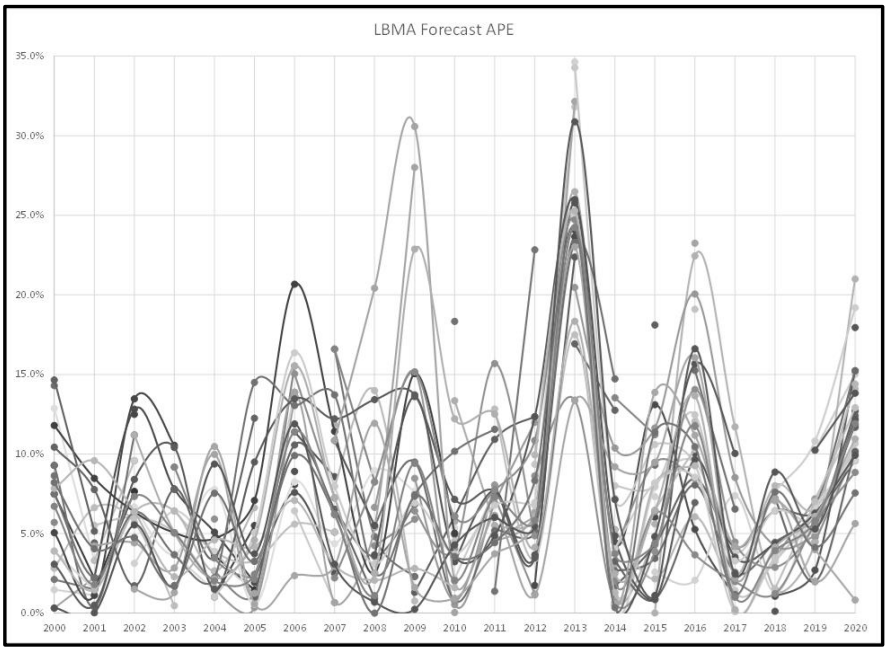


Figure 3-11 LBMA-PMS Gold Forecasters' APE

Reviewing the academic literature on comparing forecasts became necessary to address the complication in analyzing the quantitative data collected because of the inconsistent APE results for the LBMA-PMS forecasts. Numerous references to forecasting accuracy relate to weather prediction, which involves analytical modeling and expert judgment elements, similar to what occurs in metal price forecasting (Brier, 1950; Gilbert, 1987). A focus on assessing forecasting accuracy exists within the field

of econometrics, with the work of Theil (1961; 1966) attracting support and dissension (Ahlburg, 1984; Armstrong, 2001a; Bliemel, 1973; Koutsoyiannis, 1977). In framing the issue of forecasting accuracy, "*Principles of Forecasting*" by Armstrong (2001a) and "*Another look at measures of forecast accuracy*" by Hyndman and Koehler (2006) were reviewed. The following is a summary of the classification of the measures of forecast accuracy, detailed in the two sources mentioned, albeit not solely their derivations.

Four classifications of measures of forecasting accuracy can be identified, with the calculation of several of the measures possible based on alternatives such as the arithmetic mean, the median, and the geometric mean (Armstrong, 2001a; Hyndman and Koehler, 2006):

- Scale-dependent measures: Based on the forecast deviation from the actual reference value. The unit of measurement of the data makes comparison across different categories or periods difficult. Examples of these measures are the Mean Square Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).
- Percentage error measures: Based on the deviation of the forecast from the actual reference value divided by the actual reference value. The division by the actual reference value eliminates the scale issue, making a comparison across different categories and periods comparable. Examples of these measures are the Absolute Percentage Error (APE), Mean Absolute Percentage Error (MAPE), and Root Mean Square Percentage Error (RMSPE).
- Relative error measures: Calculated by taking the forecast deviation from the actual reference value and standardizing the result by dividing it by another relevant reference value, such as the prior period reference value. Because the measure is standardized through a similar scale reference value division, the result is not scale-dependent and can be compared across categories and periods. Examples of these measures are Mean Square Relative Error (MSRE), Mean Absolute Relative Error (MARE), and Mean Absolute Percentage Relative Error (MAPRE).
- Relative measures: Instead of standardizing the forecast deviation through division by a reference relevant value, the forecasted deviation is divided by a similar benchmark value, such as one obtained from an alternative forecasting approach.

The division by a comparable scaled value ensures the result is scale-independent and allows for comparison across categories and periods. Using the prior period value as the divisor implies a comparison to an expected static trend, referred to sometimes as the "naïve" forecast used as a measure of relative accuracy assuming no change. Examples of these measures are Relative Mean Square Error (RelMSE), Relative Mean Absolute Error (RelMAE), and Relative Mean Absolute Percentage Error (RelMAPE).

In deciding on the measure of forecast accuracy, Armstrong (2001a) recommends several areas for assessing the error measurements:

- Select error measurements independent of scale if comparisons are made across categories or periods.
- The error measures should be recognized and understood within the context of the forecasts.
- Use error measures that are understandable to the target audience without undue difficulty or computational complexity.
- Take care to use error measures that consistently report the results rather than measures that obscure outcomes, thereby biasing the results in a particular way.
- Consider the impact of outliers in calculating the results, as aggregated measures are influenced.
- Use care in using correlation as a relative accuracy measure across periods, as the underlying environment might be vastly different, obscuring or causing correlation outcomes that may not be indicative.
- Measures based on squared errors would be influenced by outliers, even if the measure is the square root of the calculated results.
- Use multiple error measures to report the results, even when using a predominant measure to discuss the results.
- In evaluating the forecast results, consider the alternative "naïve" benchmark or a similar option based on the ex-ante information as a comparison.
- Focus on "*practical significance*" instead of "*statistical significance*" in reporting the results, mainly if the results are not from similar events.
- To test the validity of the "*practical significance*" reported, conduct ex-post tests to measure the reliability of the results.

- Ensure the sufficiency of the sample size to conclude the results' reliability. Another aspect of sample size is replicating the sampling to make the successive samples comparable.
- Consider the trade-offs between costs and benefits, both in terms of monetary value and time, as the possibility of sampling the entire relevant population is unlikely to be workable or possible.

In reviewing the literature on forecast accuracy, it was noted limited attention was devoted to defining how the various measures are calculated. The issue surfaces in the discussion of Theil's U statistic on whether the calculation should be based on the forecast value change or the forecast value's absolute value (Armstrong, 2001a; Bliemel, 1973; Ahlburg, 1984). If the change in the forecasted value is small, the resulting calculation based on change can yield calculated accuracy measures that tend towards infinity. For the AR project, after considering both the change and absolute values in calculating the error measurements, the decision was taken to focus on the absolute values. The benefit was that it negated reporting substantial error measurement results unless the metal price changes were minor. It also ensured the basis of calculating the error measurements aligned with the base of the forecasts made. A secondary consideration in calculating the error measurements was assigning the forecast's deviation compared to the actual reference value to the numerator and the benchmark reference actual value to the denominator in the error measure calculations. The consequence was consistently comparing the results for the various error measurements.

Criticism of the Theil U(II) Index for reporting substantial error measurement when the denominator change is small is discussed by Bliemel (1973) and Ahlburg (1984). Despite the criticisms, the Theil U(II) Index proved useful in AR project quantitative analysis for comparing aggregated consensus estimates and individual participant forecasts. The Theil U(II) Index used the absolute rather than relative values change to address the calculation criticisms. The difference between the actual reference value and the forecast's median was used in the denominator to supply a referential benchmark index. Values greater than one showed less accurate forecasts; below one, more accurate forecasts. The calculated Theil U(II) Index was the RMSE based on the forecast and actual reference value divided by the RMSE based on the forecast's median and the actual reference value. Intuitively, The deviation of the Theil

U(II) Index from one for the forecasts and aggregated consensus estimates showed the extent they could be judged accordingly for their forecasting accuracy (Armstrong, 2001a; Theil, 1961; 1966).

Another consideration in choosing the Theil U(II) Index was the ability to decompose the Mean Square Error into its three constituent elements, as recommended by Koutsoyiannis (1977), Stewart and Lusk (1994), Page (2007), and Hong and Page (2012). The first element relates to the difference between the actual and forecasted means and is referred to as the bias element, being the difference between the actual and expected values. The second element relates to the difference between the actual and forecast variances and is called the variance element, which is the difference in variation between the actual and expected outcomes. The third element relates to the extent the actual and forecast results do not change in tandem and refers to the covariance element. The extent to which the actual and forecasted outcomes are not changing in sync. Conceptually, the bias and variance elements could be addressed by expanding the forecasting pool and reducing the negative error measures by including more estimates in the aggregated consensus estimates. The last element, the covariance element, is less easily reduced and depends on improving the forecasting skills of the forecasters to ensure better identification of the actual trends (Koutsoyiannis, 1977).

Measuring the relative accuracy of the forecasting participants and the aggregated consensus estimates was a crucial element of the AR project for finding a reliable forecasting approach. For the CM panel, participants were ranked according to their forecasting accuracy. For the LBMA-PMS, the participants were ranked based on the APE of their forecasts compared to the actual average metal price reported by The Independent Precious Metals Authority (2020). The expectation when the AR project was conceptualized, based on the "*wisdom of crowds*" literature, was that the aggregated consensus average would be the most reliable measure of forecasting accuracy. The CM panel data collected and the annual LBMA-PMS gold forecasts for the twenty-one years from 2000 to 2021 were ranked based on the APE for each participant as a reflection of forecasting accuracy. The following aggregated consensus estimates were used to quantitatively analyze the forecasts from the CM panel and the LBMA-PMS.

The term "aggregated consensus estimates" refers collectively to the following statistical measures when referenced in this AR report:

- Consensus Average: The arithmetic average of the forecasts made.
- Consensus Median: The forecast's median or mid-point with 50% above and below the median value.
- Lower Quartile: The lower quartile of the forecasts made with 25% of the forecast below and 75% above the lower quartile value.
- Upper Quartile: The upper quartile of the forecasts made with 75% of the forecast below and 25% above the upper quartile value.

Forecasters can be considered experts within their fields, and the logic would be to seek the advice of the most reliable forecaster when relying on forecasted metal prices. Metal price forecasting is judgmental, so no forecaster could be expected to predict future metal prices reliably and consistently. A graph of the LBMA-PMS APE was made to identify any discernible patterns of consistency, but no clear trends were observable (*Figure 3-11 LBMA-PMS Gold Forecasters' APE*). The Mann-Kendall statistic (Mann, 1945; Gilbert, 1987) and the Sign Difference (Moore and Wallis, 1943) were calculated for the LBMA-PMS forecasters and the consensus average to verify the lack of consistency in APE trends and the ranking of forecasters. In addition, to check for the consistency of forecasting accuracy across successive periods, the relative correlation between consecutive periods was computed as a correlation Test-Retest coefficient (Guttman, 1945; Weir, 2005).

A crucial part of forecasting reliability is correctly predicting the future direction of the market's metal prices. Any mistake in predicting the market direction would significantly affect the APE and account for some variation seen in the APE across successive periods. The probability of correctly predicting the direction of market metal price changes for the LBMA-PMS forecasters' was evaluated for reliability. The forecasted metal price direction and the actual market direction changes conditional probabilities were computed. The calculated conditional probabilities were differentiated according to the relative gold price in the first week of January of the forecasted year versus the prior year's average. From an AR project and a workplace planning perspective, finding the reliability of forecasting market metal price directional changes was a significant factor in improving the forecasting reliability of metal prices.

At the commencement of the AR project, the quantitative analysis focused on the APE for the respective forecasts made for the different periods and metals. The APE

is a simple calculation, and the collected quantitative data were tabulated in Microsoft Excel (Version Professional Plus 2021) (Microsoft Excel, 2021). The decision to use Microsoft Excel to perform all the error measurement calculations was also the benefit of observing data trends, patterns, and any discrepancies in the computational results. To verify the computations, the Microsoft Excel Add-in NumXL Pro: Time Series and Statistics Software (Version 1.67) (Spider Financial Corp, 2021) was used to perform the error measurements in tandem.

In addition to the quantitative element of forecasting metal prices, the cooperative inquiry also included a qualitative aspect that required participants to justify their forecasts. When called upon to supply a rationalization for our opinions, it may add greater certainty to the thought process. It could be because of the need to be more reflective than an impulsive response or because it encourages considering the biases that could influence the estimation more consciously (Argyris, 1977; Bazerman and Moore, 2008; Drummond, 2001). For the cooperative inquiry, the expectation was that the CM panel participants' interactions would positively influence their cyclical metal price forecasts, enhancing the reliability of the emerging aggregated consensus estimates. When conducting a Delphi Method project, a recommendation is that participants explain their judgments, which the moderator reviews and shares with the participants as deemed necessary. The motivation for sharing the information in a Delphi Method project is to surface added insights based on the collective knowledge of all participants (Dalkey and Helmer, 1963; Rowe and Wright, 1999). In the research conducted by Tetlock and Gardner (2016) in the Good Judgment Project, the benefit of participants' interactions was identified as contributing to the more reliable outcomes realized. There are two significant aspects underlying the inclusion of justifications in the AR project. Firstly, do participants exercise more consideration in making forecasts when explaining the basis for their estimates? Secondly, does the ability to see and interact with other participants prompt reflection and action?

The CM panel included a provision for the participants to justify their metal price forecasts. For the annual LBMA-PMS, the participants making forecasts had to explain the basis of their forecasts. A significant difference between the two data sets was that the CM panel forecaster’s participation was anonymous, private, and for a short duration, six months. For the LBMA-PMS forecasters, participants were known, the forecasts were public and widely seen within the industry, and the forecasts only

occurred annually. Because of the fundamental underlying differences between the two forecasting forums, it was expected the participants would exercise equal diligence in substantiating their forecasts.

The CM panel justifications were analyzed according to the word count, diversity of ideas presented, and frequency. The expectation was that a consistent pattern for each participant would be discernible, reflecting their thought processes when forecasting metal prices. An expectation was that no significant shift in outlook would occur over a short duration other than in response to external exogenous influences. The cooperative inquiry element of interest was whether participants reflected on the CM panel interactions before making their cyclical metal price forecasts based on the reference to their cohorts' forecasts and justifications. Lack of reference to the CM panel interactions and a greater focus on exogenous events could show an ego bias among the participants, with their reflections outwardly focused on the activities of the cooperative inquiry (Bolger et al., 2011; Bolger and Wright, 2011). From a research perspective, the focus was on the collective interactions of the participants. As an interacting group, would they be more capable of making reliable metal price forecasts than any individual participant? The qualitative analysis of the CM panel data was done using WordSmith Tools (Version 8) (Scott, 2020), and the output was tabulated and analyzed in Microsoft Excel (Microsoft Excel, 2021).

The forecast explanations sourced from the annual LBMA-PMS were analyzed differently from the CM panel data to avoid ethical issues based on the data source. Unlike the CM panel data, the participants in the annual LBMA-PMS were not approached and asked permission to analyze and discuss the findings of their explanations for the AR project. Although the LBMA-PMS participation was public, analyzing and discussing the participants' explanations individually and reporting on the comparisons and findings could harm the participants' professional standing within the industry. The AR project aimed to evaluate collective actions by testing the reliability of aggregated consensus metal price forecasts compared to the estimates of particular forecasters. Consequently, the LBMA-PMS participants' annual justifications were analyzed as a single annual collective justification.

Based on research conducted by the World Gold Council (Gold, 2021), they identified several factors that influence the future price of gold:

- Wealth and Economic Expansion: GDP growth drives the demand for gold as an investment and for industrial applications, particularly in emerging markets.
- Risk and Uncertainty: Gold is seen as a hedge against inflation, interest rate movements, and changes in money supply in key major economies.
- Opportunity Cost: The interest on sovereign risk-free investments compared to the holding cost of gold.
- Momentum and Positioning: The trend in future market open positions compared to the level of the gold price and the changing levels of the new gold supply.

Using the categorization of the factors identified by the World Gold Council (Gold, 2021) as a guide, the following formulation was used to group the occurrence of keywords in all the annual LBMA-PMS gold forecast explanations.

Table 3-8 World Gold Council Key Factors

World Gold Council Key Factors	Economy	Demand	Risk / Uncertainty	Opportunity cost	Trends / Supply
Economic	√				
USA	√				
World	√				
Oil	√				
Demand		√			
Market		√			
Central Banks		√			
Metal		√			
Production hedge		√			
Asset		√			
Jewellery		√			
ETF		√			
Reserve		√			
China		√			
India		√			
Asia		√			
US Dollar			√		
Currency			√		
EUR			√		

World Gold Council Key Factors	Economy	Demand	Risk / Uncertainty	Opportunity cost	Trends / Supply
Risk			√		
Inflation			√		
Policies			√		
Geopolitical			√		
Political			√		
Environment			√		
Safe-haven			√		
Interest Rates				√	
Investor				√	
Investment				√	
Federal Reserve Bank				√	
Equity Exchange				√	
Fund				√	
Supply					√
Producer					√
Mine					√
Production					√

After standardizing the terminology of the analysts' explanations, e.g., Dollar, US$, and USD, the keyword occurrences were measured for each of the twenty-one years of the LBMA-PMS. The AR project assessed the reliability of aggregated consensus estimate forecasts and the keywords' relative frequency. The keywords analysis evaluated whether a joint quantitative and qualitative approach could improve the reliability of aggregated consensus estimates. Based on the World Gold Council (Gold, 2021) keywords, the collocates of the keywords were cross-tabulated to evaluate if the concurrent occurrences were significant and indicated if they had a bearing on the reliability of the aggregated consensus estimates. The annual LBMA-PMS explanations were qualitatively analyzed using WordSmith Tools, and the output was tabulated and analyzed in Excel.

3.10. Research Hypotheses

"*Aristotle's definition of phronesis as "reason capable of action" means that phronetic research results ("reason") are results only to the extent they have an impact*

on practice ("action") (Flyvbjerg et al., 2012, pp.286). The AR project aimed to evaluate the task of making metal price forecasts ("*reason*") and consider the possibility of recommending alternative aggregated consensus metal price estimates for use in practice ("*action*") (Flyvbjerg, 2004; Flyvbjerg, 2006; Flyvbjerg et al., 2014; Schram et al., 2013). In undertaking the AR project as a pilot case study, the expectation was not to definitely validate the hypotheses considered but rather to understand the possibility of using collective metal price forecasts in practice (Creswell, 2012; Lincoln and Guba, 1985). Understanding relates to the "*interpretation*" and the "*accuracy*" of the findings by the relevant stakeholders (Creswell, 2012; Wolcott, 1990).

The AR project aimed to evaluate the reliability of using aggregated consensus metal price forecasts more widely in the natural resource industry. A cooperative inquiry approach was used as a case study to collect metal price forecasts made in observable interactive cycles by participants for the AR project (Coghlan, 2007; Heron and Reason, 2008; Stake, 2003). The CM panel participants were invited to submit their copper and gold metal price forecasts and supporting justifications over six months on the bespoke website designed for the AR project. As a pilot case study over a short time, the expectation was not to collect a substantial number of metal price forecasts. After the six-month data collection phase of the AR project, 191 metal price forecasts and 161 justifications were collected from the CM panel participants.

The CM panel data set was a small fraction of the total population of metal price forecasts made in practice, requiring caution in drawing general conclusions from the hypotheses evaluated (Armstrong, 2001a). To ensure the authenticity of the findings, Creswell (2012) recommends triangulating the results of data collected from case studies with comparable data from other sources, a view shared by Armstrong (2001b). In keeping with the recommendation of triangulating the results of a case study, the more extensive data set of metal price forecasts from the LBMA-PMS was sourced and analyzed in a manner comparable to the CM panel analysis. The combined data sets of metal price forecasts expanded the ability to draw conclusions from the AR project. As a phronetic social science AR project based on a limited sample of metal price forecasts, it could only provide a measure of "*adequation*" of what may work in practice (Flyvbjerg et al., 2012). The findings would best serve as a departure point for further related research by those interested in metal price forecasts.

The hypotheses and alternative null hypotheses evaluated for the AR project aimed to understand further the possibility of using aggregated consensus metal price forecasts in the natural resource industry (Creswell, 2012; Flyvbjerg et al., 2012; Lincoln and Guba, 1985; Wolcott, 1990). The CM panel data collected were used to evaluate the following hypotheses, and the results were triangulated with the additional LBMA-PMS data sourced:

Hypothesis 1

The aggregated consensus median will be the most reliable metal price forecast over multiple periods, capturing the diversity of estimates around the statistical midpoint.

Null Hypothesis 1 - Alternative 1

An individual expert can reliably forecast metal prices over multiple periods more than the aggregated consensus median.

Null Hypothesis 1 - Alternative 2

As alternative aggregated consensus estimates, the lower and upper quartiles can be used as more reliable metal price forecasts than the aggregated consensus median or specific experts' forecasts.

Hypothesis 2

Repetitive forecasting cycles will improve the reliability of metal price forecasts made by an individual expert.

Null Hypothesis 2 - Alternative 1

Using a collective decision-making approach to forecast metal prices will yield a reliable result.

Null Hypothesis 2 - Alternative 2

Over repetitive cycles of forecasting metal prices, alternative aggregated consensus estimates will prove more reliable than the median or a specific expert.

Hypothesis 3

In making metal price forecasts, forecasters have a consistent expectation and explanation of the future metal price outcome.

Null Hypothesis 3 - Alternative 1

Metal price forecasters focus more on the quantitative rather than the qualitative element of their predictions.

Null Hypothesis 3 - Alternative 2

Metal price forecasters focus on similar exogenous factors when making forecasts and supporting justifications.

The question is whether a cooperative approach for forecasting metal prices in the natural resource industry would be a reliable alternative workplace practice that warrants further consideration.

3.11. Ethical Issues

The cooperative inquiry that formed the basis of the AR project discussed in this thesis was in line with the University of Liverpool ethical approval granted. Invited participants anonymously made metal price forecasts on the CM panel website without any obligation to participate or remain involved in the AR project. The ethical challenges that did surface due to the AR project's data collection centered around three potential conflict areas. Despite our best endeavors, we cannot compel others to commit to something they are unwilling to do. Conversely, we should not let our actions impose an outcome on others without their agreement, especially if there is a risk of harm. For a researcher striving for greater generality of the research findings, gaining access to additional data should not be considered if it impacts the ethical foundations of the research project. In dealing with any ethical issues as a researcher, the AR project focused on avoiding any potential ethical problems to the maximum extent possible.

Participation on the CM panel website to make metal price forecasts were by invite, focusing on industry professionals who routinely within the workplace either make or rely on forecasted metal prices. By targeting potential participants with an understanding and interest in forecasted metal prices, the expectation was that they would benefit from the insights learned in exploring an alternative workplace practice. As industry experts, the participants would be familiar with making metal price forecasts, so agreeing to participate in the cooperative inquiry and make metal price forecasts should not involve any additional psychological pressure. Also, because of their familiarity with forecasting metal prices, the expectation was that it would only

require a minimal time commitment outside their normal work activities. The concepts underlying collective consensus forecasts were provided on the CM panel website to allow the CM panel participants to understand the background of the cooperative inquiry (Annexure A: CM panel participant research project background). The invite sent to potential participants stated that their participation on the CM panel website would be anonymous should they choose to participate. In addition, when the CM panel participants registered on the website, they had the choice of selecting a pseudonym to submit their metal price forecasts (Annexure A: CM panel participant research project background). The targeted metal prices for copper and gold price forecasts used in the cooperative inquiry were the London Metal Exchange (LME, 2019) and London Bullion Market Association (LBMA, 2020) prices, respectively. Both commodity exchanges' market prices serve as reference benchmark metal prices in the natural resource industry, and both use English as their primary language of operation. The decision to use English for the cooperative inquiry participation was based on the LBMA and LME practices and the assumption that industry professionals would be conversant in English.

The participants who accepted the invitation to submit metal price forecasts on the CM panel website were required to complete a registration process. Registration on the CM panel website required the participants to agree to the stipulated ethical conditions. The participants were expressly required to grant their consent that their metal price forecast data would be collected and analyzed for inclusion in this thesis (Annexure B: CM panel participant registration and consent form). Further clarification was provided on the CM panel website as to why the CM panel participants were invited to participate in the cooperative inquiry. The additional background information also explained the participants could cease participating at any time and have their forecasts excluded from inclusion in this thesis (Annexure C: CM panel participant additional research information). The additional information on the CM panel website also reaffirmed that no compensation was to be paid for participating. The benefit of participating was linked to the insights learned through being involved in the cooperative inquiry. The contact details for the University of Liverpool, the thesis supervisor, and the researcher were provided on the CM panel website if any participant wanted to raise any issues or concerns. The researcher and the thesis supervisor were anonymous participants on the CM panel website in making

metal price forecasts, giving them both visibility of the process and any issues or concerns raised by the other participants. During the seven months of collecting metal price forecasts on the CM panel website, no participants raised any ethical issues or concerns. Following the conclusion of the data collection phase, no CM panel participant requested to have their metal price forecasts excluded from the data analysis presented in this thesis.

The researcher held certain expectations about the number of forecasters taking part in making metal price forecasts on the CM panel website. As a pilot study, the expectation was to have between 20 to 30 participants registered on the CM panel website, with around a dozen actively making metal price forecasts. When about 100 of the approximate 500 invitees sent out to individuals and organizations active within the natural resource industry expressed some interest in the AR project, the hope was to exceed the original participation estimate. The added participation from the potential 100 invitees was limited, with some invites referring to the outbreak of COVID-19 and the decision of their organizations to adopt work-from-home arrangements affecting their choice of declining. After sending two further requests to encourage participation, the decision was made to respect the choice of those invitees who had declined to participate. Trying to compel participation based on the researcher's expectation of the value for participants in their workplace practices was insufficient justification for continuing to question their choice not to participate in the cooperative inquiry.

The second ethical challenge arose from the decision to source additional metal price forecasts that spanned a more extended period. That would allow triangulation with the findings based on the data collected from the CM panel cooperative inquiry. The annual LBMA-PMS supplied the opportunity to access historical metal price forecasts made by industry specialists spanning twenty-one years. A consideration in analyzing the LBMA-PMS forecast data set was the similarity with the data collected from the CM panel participants. For the annual LBMA-PMS, the participants know that their forecast accuracy, as measured by the Absolute Error relative to the market annual average gold price, would be used to rank their performance compared to their peers (LBMA, 2020). The annual LBMA-PMS participants' forecasts are published, as are the rankings of the top participants. The annual LBMA-PMS forecasting competition occurs in the public domain, so comparing the participants' performance

to their peers or other reasonable benchmarks did not pose any added ethical issues for the participants other than those already existing.

The cooperative inquiry included a requirement for a justification by the CM panel participants in making their forecasts. They consented to have their explanations seen by their cohorts and analyzed for the AR project. The annual LBMA-PMS had a similar requirement for the participants to justify their forecasts. However, the LBMA-PMS competition includes no review or ranking of the individual explanations in their competition results. Analyzing and ranking the LBMA-PMS participants' justifications relative to a set of keywords, such as the World Gold Council (Gold, 2021), was not explicitly considered by the LBMA-PMS forecasters. The possibility when analyzing and ranking the LBMA-PMS individual justifications for some professional harm to the participants without their express consent exists. Even if the justification analysis was anonymously done, cross-referencing the original annual LBMA-PMS explanations could still make it possible to identify the forecaster cited. To avoid any ethical issue in the researcher's analysis and published thesis, the LBMA-PMS explanations were aggregated annually as one composite explanation to prevent the possibility of identifying any specific individuals in the qualitative analysis and causing harm to them.

During the search for more consensus metal price forecasts to triangulate with the results from the data collected from the CM panel cooperative inquiry and the LBMA-PMS, Consensus Economics (2020) data set was identified as a possible additional data source. The Consensus Economics (2020) data set includes metal price forecasts from industry experts for a more extended period than the CM panel or the LBMA-PMS. Unlike the LBMA-PMS metal price forecast, the Consensus Economics (2020) forecasts are updated more regularly during the year, either bi-monthly or quarterly. The Consensus Economics (2020) metal price forecasts include no justifications by the forecaster and are reported by the institution rather than an analyst. Despite the differences, Consensus Economics (2020) was approached to gain access to their data set to triangulate the aggregated consensus metal price estimates with forecast data from the LBMA-PMS. The data analysis approach used for the CM panel forecasts and the LBMA-PMS was the same, allowing triangulation of the findings. The hope was to expand the generality of the findings by accessing the Consensus Economics (2020) data. Unfortunately, Consensus Economics (2020) was only willing to share

their data set for their standard fee of several thousand Pounds Sterling. In addition, Consensus Economics (2020) wanted prior consent on publishing any report that included their data. Comparing the Consensus Economics (2020) and LBMA-PMS data for the twenty-one years analyzed would have offered an opportunity to evaluate AR project findings regarding aggregated consensus forecasted metal price estimates and, specifically, the importance of requiring justifications. The Consensus Economics (2020) requirement of prior consent on publishing the researcher's thesis was considered unacceptable. Rather than pay the Consensus Economics (2020) subscription fee and face the ethical dilemma of publication permission, the opportunity to explore the triangulation of the AR project findings was foregone.

3.12. Summary

If you are unsure of the answer to a question, ask some people with relevant expertise, and use a method to aggregate their views to reach a consensus opinion, combing their "*collective wisdom*" (Aristotle and Everson, 1988; Galton, 1907a; Page, 2007; Sunstein, 2006; Surowiecki, 2004; Tetlock and Gardner, 2016). The concept of collective decision-making is ubiquitous across time, cultures, and situations. How the collaborative approach is structured has been considered as it pertains to the access of participants and the aggregation of their opinions (Armstrong, 2001a; Batchelor and Dua, 1995; Hong and Page, 2012). Using a cooperative inquiry, structured similarly to groups' decision-making approaches, to better understand the judgmental nature of metal price forecasting by allowing participants on the CM panel to engage in collaborative interactions (Heron and Reason, 2008; Rowe and Wright, 1999). The interaction between participants as they shared their views and could reflect on the insights of other group members had the potential to surface a more reliable outcome than that of the individuals involved. The group interactions develop from the individual positivist perspective to a more encompassing social construction premised on all the group members' insights and actions to surface a consensus outcome (Easterby-Smith et al., 2012). Forecasting metal prices is a judgmental activity often undertaken in isolation. Through structuring a cooperative inquiry, the evaluation of collaborative approach was explored in the AR project to assess its reliability as an alternative workplace practice.

CHAPTER FOUR – RESULTS AND FINDINGS

4. Results and Findings

4.1. Introduction

When making metal price forecasts, the anticipation of both the forecaster and the users is how accurate the predictions will be compared to the eventual outcome. In the natural resource industry, the importance of making reliable metal price forecasts is epitomized by the impairments incurred by the top forty mining organizations of circa US dollar two hundred and fifty million in the last decade, or equivalent to one-third of their net profit for the same period (PwC, 2021). An implication of the impairments could be that the industry did not reliably forecast future metal prices during the decade, particularly during downturns in metal prices. Without access to how the metal price forecasts of the relevant organizations were done during this period, it is impossible to say if the metal price forecasts reflected a general industry misjudgment of future metal price changes or related to specific forecasters. The substantial impairments occurred around significant metal price declines in 2012, 2013, and 2015, which accounted for approximately sixty percent of the total impairments incurred in the last decade (Annexure D: PwC Top 40 Mines Report 2011-2020).

Typical approaches for forecasting metal prices are fundamental analysis, extrapolating historical trends, referencing external sources and market futures, or combining these approaches. Amongst forecasters, knowledge of and preferences for the different approaches exist, reflecting how the individual forecasters make their estimates based on their insights (Armstrong, 2001a; Batchelor and Dua, 1995). Given the importance of the mining organizations' sustainability, a reasonable assumption would be that the metal price forecasting process involves collaboration between those involved in making the relevant business decisions that rely on the forecasted metal prices. The advantage of a collaborative approach would be to tap into the unique insights of those involved, potentially resulting in a more reliable outcome. The AR project focused on using a cooperative inquiry approach to collect observable metal price forecasts from participants on the CM panel. The research aimed to evaluate the reliability of aggregated consensus estimates and the importance of requiring justifications for metal price forecasts made. The AR project findings showed that a reliable alternative to relying on a specific forecaster is possible, although not based

on the aggregated consensus average or median, as is widely used in practice. A crucial element influencing the reliability of using a consensus approach for forecasting metal prices was correctly anticipating the future direction when aggregating the participants' metal price predictions.

4.2. **Consensus Metals Panel Forecasts Results**

The CM panel website was constructed to collect average copper and gold price forecasts from registered participants from October 2020 to December 2021. The participants were requested to make average metal price forecasts monthly for October 2020, November 2020, and December 2020 and quarterly for March 2021, June 2021, September 2021, and December 2021. The reference prices for the copper forecasts were the London Metal Exchange daily cash average for the relevant periods (LME, 2019). The reference prices for the gold forecasts were the London Bullion Market Association (LBMA) daily PM fix average for the relevant periods (LBMA, 2020). The CM panel participants were also asked to supply a brief justification (Maximum 300 characters) for their forecasts. The participants made their forecasts on an anonymous basis. However, all participants' forecasts and explanations were visible to the other CM panel participants, and the progressive aggregated consensus average for each metal and period. Even if the registered participants did not make any forecasts, they could still view the other participants' metal price forecasts, the consensus average, and the justifications for the forecasts submitted. As a cooperative inquiry, participants were allowed and encouraged to make cyclical forecasts based on the insights gained, either because of the CM panel interactions or current metal market events.

Table 4-1 tabulates the demographic details of the CM panel participants registered to participate in the cooperative inquiry.

Table 4-1 CM panel Registered Participants' Demographic Details

No. of participants		**Sex**		**Age (Years)**		**Geographic location**	
Total	23	Male	19	25-35	5	Africa	6
Copper	20	Female	2	35-45	3	Asia	4
Gold	18	Other	1	45-55	7	Australia	4
				55-65	5	Europe	6
				>65	3	North America	1
						South America	1
						Other	1

Background		Workplace		Forecasting approach	
Economics	8	Academic	2	Fundamental analysis	6
Engineering	4	Consultancy	5	Futures prices	1
Financial	7	Financial institution	3	“Gut feel”	6
Technical	1	Government agency	2	Historical averages	1
Other	3	Mining organization	8	Mineral economics	4
		Technical services	2	Other	5
		Other	1		

The original request at the start of the cooperative inquiry in October 2020 was for the CM panel participants to make metal price forecasts for six months. Post-March 2021, only eight copper and one gold forecasts were made, reflecting the winding down of the data collection phase of the cooperative inquiry as expected. At the end of June 2021, the CM panel participants stopped making further metal price forecasts.

4.3. **Consensus Metals Forecasts Quantitative Results**

Table 4-2 CM panel Forecasts - Copper Consensus Estimates Summary summarizes the copper forecasts made from October 2020 to December 2021 by the CM panel participants. The AR project's quantitative element focused on analyzing the collective forecasts to evaluate the reliability of the aggregated consensus estimates compared to the participating forecaster's forecasts for each period. A significant relevant indicator in analyzing the CM panel forecasts was correctly predicting the future market copper price direction. The actual market copper price percentage changes highlight the actual market copper price direction compared to the anticipated market direction of the forecasters (*Table 4-2*; Column 2; Row a). The copper price aggregated consensus estimates forecasts errors are reported as the APE compared to the actual market copper price for the respective periods (*Table 4-2*; Columns 3-6; Row a). The smallest APE for the seven periods was consistently the consensus upper quartile (*Table 4-2*; Column 6; Rows a & b). The Theil U(II) Index is reported compared to the consensus median, with a value of less than 100% reflecting a more reliable consensus estimate than the median. Only the consensus upper quartile yielded a Theil U(II) Index less than 100%, 61.3% (*Table 4-2*; Column 6; Row c). The ReLMAE for each aggregated consensus estimate compared to the actual market copper price prior period was calculated. Based on the ReLMAE (*Table 4-2*; Row e), the consensus upper quartile (123.0%) was the most reliable estimate for the seven forecasting periods compared to 144.6% for the consensus median.

Table 4-2 CM panel Forecasts - Copper Consensus Estimates Summary

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LME average price for the prior period	LME average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	October 2020	$6,712/mt	$6,703/mt	$6,463/mt	$6,588/mt	$6,613/mt	$6,706/mt
A	Actual % Δ / APE %	**Neutral**	-0.1%	3.6%	1.7%	1.4%	**0.1%**
	November 2020	$6,703/mt	$7,063/mt	$6,613/mt	$6,829/mt	$6,738/mt	$6,931/mt
A	Actual % Δ / APE %	**Up**	5.1%	6.4%	3.3%	4.6%	**1.9%**
	December 2020	$7,063/mt	$7,755/mt	$6,713/mt	$7,082/mt	$6,900/mt	$7,725/mt
A	Actual % Δ / APE %	**Up**	8.9%	13.5%	8.7%	11.0%	**0.4%**
	March 2011	$7,755/mt	$8,504/mt	$6,756/mt	$7,314/mt	$7,475/mt	$7,869/mt
A	Actual % Δ / APE %	**Up**	8.8%	20.6%	14.0%	12.1%	**7.5%**
	June 2021	$8,504/mt	$9,700/mt	$6,755/mt	$7,534/mt	$7,300/mt	$8,138/mt
A	Actual % Δ / APE %	**Up**	12.3%	30.2%	22.3%	24.7%	**16.1%**
	September 2021	$9,700/mt	$9,372/mt	$6,625/mt	$7,527/mt	$7,200/mt	$8,125/mt
A	Actual % Δ / APE %	**Down**	-3.5%	29.3%	19.7%	23.2%	**13.3%**
	December 2021	$9,372/mt	$9,698/mt	$6,444/mt	$7,427/mt	$7,133/mt	$8,044/mt
A	Actual % Δ / APE %	**Up**	3.4%	33.6%	23.4%	26.7%	**17.1%**
	Oct '20 – Dec '21	$7,973/mt	$8,399/mt	$6,627/mt	$7,186/mt	$7,048/mt	$7,648/mt
B	Actual % Δ / APE %	**Up**	5.1%	21.1%	14.5%	16.1%	**8.9%**
C	Theil U(II) – Median			127.7%	89.3%	100.0%	**61.3%**
D	Theil U(II) – Prior			332.4%	232.4%	260.3%	**159.5%**
E	ReLMAE - Prior			131.7%	146.5%	144.6%	**123.0%**

The quantitative evaluation was informative from a cooperative inquiry perspective, where the focus was on evaluating the possibility of using aggregated consensus metal price estimates in the workplace. The quantitative analysis showed that aggregated consensus estimates other than the average or median could represent a reliable workplace forecasting estimate. Reviewing the APE of the consensus average, median, lower quartile, and upper quartile, the upper quartile was the most reliable forecast for the copper price forecasted by the CM panel participants. The APE for the aggregated copper consensus estimates over the seven forecasting periods were 14.5%, 16.1%, 21.1%, and 8.9%, respectively, for the consensus average, median, lower quartile, and upper quartile compared to the actual market change of 5.1% (*Table 4-2*; Row b).

Between October 2020 and December 2021, the average market copper price increased by around 40% (LME, 2019). The CM participants had not anticipated the rising market copper price. Participants cited the United States elections at the end of 2020 and the possible impact of policy changes on the relative value of the United States Dollar. There were also several references by participants to COVID-19. They expressed concern about its effects on world economic growth, especially for industrial activities that used copper as an input. The expectation was that the general economic environment would be negative for copper. Of the 106 forecasts made by the CM panel copper participants, 82 (77.4%) predicted the future copper price to be less than 1% higher than the market prior period copper price, and only 24 (22.6%) anticipated a

higher future copper price. With many CM panel copper participants incorrectly predicting the market copper price direction, the longer the lead time for the forecasts made, the less accurate they were. The more recent CM panel copper price forecasts reflected the rising market copper price, helping to improve the aggregated consensus copper price estimates compared to the relevant average market copper prices.

Table 4-3 CM panel - Copper Consensus Estimates Reliability

Cm panel copper participant	No. periods forecasts made	No. period most accurate	No. forecasts submitted	APE measure	APE (%)	Average Rank	Rank z-Score
Part. 1	7	0	19	APE	22.7%	8	(1.16)
				APE Med	16.1%	6	(0.39)
				APE Q3	**9.0%**	3	0.78
Part. 2	7	0	16	APE	12.1%	4	0.39
				APE Med	16.1%	6	(0.39)
				APE Q3	**9.0%**	3	0.78
Part. 3	7	2	13	APE	12.1%	4	0.39
				APE Med	16.1%	6	(0.39)
				APE Q3	**9.0%**	3	0.78
Part. 4	7	0	12	APE	17.9%	5	0.00
				APE Med	16.1%	6	(0.39)
				APE Q3	**9.0%**	3	0.78
Part. 5	6	0	10	APE	16.0%	6	(0.39)
				APE Med	18.0%	7	(0.78)
				APE Q3	**10.1%**	3	0.78
Part. 6	4	3	9	**APE**	**11.3%**	2	1.16
				APE Med	22.0%	7	(0.78)
				APE Q3	13.7%	3	0.78
Part. 7	3	1	6	APE	1.3%	1	1.55
				APE Med	5.9%	6	(0.39)
				APE Q3	**0.8%**	2	1.16
Part. 8	6	0	6	APE	22.0%	7	(0.78)
				APE Med	18.0%	7	(0.78)
				APE Q3	**10.1%**	3	0.78
Part. 9	4	2	4	**APE**	**13.3%**	3	0.78
				APE Med	22.0%	7	(0.78)
				APE Q3	13.7%	3	0.78
Part. 10	3	1	3	APE	6.7%	4	0.52
				APE Med	9.5%	7	(0.52)
				APE Q3	**3.4%**	2	1.22
Part. 11	3	2	3	**APE**	**7.5%**	2	1.16
				APE Med	20.3%	7	(0.78)
				APE Q3	12.5%	3	0.78
Part. 12	2	0	2	APE	4.5%	7	(0.78)
				APE Med	3.0%	6	(0.39)

Cm panel copper participant	No. periods forecasts made	No. period most accurate	No. forecasts submitted	APE measure	APE (%)	Average Rank	Rank z-Score
				APE Q3	**1.0%**	2	1.16
Part. 13	1	0	1	APE	7.5%	8	(1.53)
				APE Med	1.4%	4	0.22
				APE Q3	**0.1%**	1	1.53
Part. 14	1	0	1	APE	5.3%	7	(1.09)
				APE Med	1.4%	4	0.22
				APE Q3	**0.1%**	1	1.53
Part. 15	1	0	1	APE	15.2%	9	(1.22)
				APE Med	11.0%	7	(0.52)
				APE Q3	**0.4%**	1	1.57
Part. 16	1	0	1	APE	8.8%	5	0.00
				APE Med	11.0%	7	(0.78)
				APE Q3	**0.4%**	1	1.55
16	**63**	**11**	**107**	**Total**			
Average APE / Average Rank / z-Score <1.4> (CL 90%)				**APE**	**11.5%**	**5**	**2**
				APE Med	**13.0%**	**6**	**0**
				APE Q3	**6.4%**	**2**	**4**

Table 4-2 CM panel Forecasts - Copper Consensus Estimates Summary examined the CM panel copper forecasts from the perspective of the reliability of the aggregated consensus estimates. *Table 4-3 CM panel - Copper Consensus Estimates* considered the reliability of the individual participants. Around half, 9 (56.3%) of the 16 CM panel copper participants achieved a lower APE than the equivalent aggregated median APE. By comparison, only 3 (18.8%) of the 16 CM panel copper participants had a lower APE than the equivalent APE for the Upper Quartile aggregated estimate. The 3 participants with the lowest APE (*Table 4-3*, Participants 7,12 & 14) only submitted copper price forecasts in 3, 2, and 1 of the seven forecasting periods, respectively. Using the CM panel copper participants' APE in each forecasting period, they submitted forecasts; a ranking was assigned from most to least accurate. The participant's rankings achieved for each period they participated were averaged to obtain an overall ranking. Considering the comparable average number of participants, a z-Score of the ranking was calculated for each participant. At a 90% confidence limit, only 1 participant (*Table 4-3*, Participant 7) had a positive z-Score exceeding the 1.4 confidence limit threshold. Significantly, the average APE (1.3%) for the participant with the lowest APE (*Table 4-3*, Participant 7) was higher than the equivalent APE for the Upper Quartile (0.8%).

The CM panel copper aggregated consensus metal price estimates from October 2020 to December 2021 were evaluated for accuracy using the Theil U(II) Index based on the consensus median (*Table 4-2;* Row d). The aggregated consensus upper quartile had the lowest Theil U(II) Index value, 61.3%, representing a more reliable forecast than the median (100%). The RMSE forecast error for the actual versus forecast (Nominator) and actual versus median (Denominator) was decomposed into their respective bias (Um), variance (Us), and covariance (Uc) error components to highlight the relevant impact of each in the forecast deviations. The formulation of Theil's U(II) Index is comparable to the RelMAE when the prior period is used as the reference comparison. The Theil U(II) Index based on the preceding period and the RelMAE were computed to evaluate the forecasting error (*Table 4-2*; Row d & e).

Table 4-4 CM panel - Copper Theil U(II) Index RMSE Decompositions

Theil U(II) Error Elements	RMSE	Um	Us	Uc
Nominator: Average (Actual versus Forecast)	1,473	67.9%	31.0%	1.2%
Nominator: Lower quartile (Actual versus Forecast)	2,107	70.7%	24.2%	5.1%
Nominator: Upper quartile (Actual versus Forecast)	1,011	55.2%	37.0%	7.8%
Denominator: Actual versus Median	1,650	67.0%	27.9%	5.1%
Denominator: Actual versus Preceding period	634	45.3%	0.0%	54.7%

The unanticipated increase in the market copper price by the CM panel participants was reflected in the high RMSE values for the aggregated consensus metal price forecasting estimates, including the consensus median. The lower RMSE for the prior period was because the preceding period average more effectively captured the upward trending copper price with a shorter lead-time between the forecasting period. The incorrectly predicted market copper price direction by the CM panel participants was reflected in the substantial Theil Um (bias element) contribution to the forecasting error. The forecasting error could have been reduced by having more participants, assuming they did not share similar copper price expectations. The lesser but still significant contribution of the Theil Us (variance element) to the forecasting error of the CM panel participants reflected the slow change in their copper price expectations over the successive cycles of forecasting. Between October 2020 and December 2021, the increasing copper price resulted in a lagged change in the preceding period value compared to the actual period value, reflected in the covariance error. The longer duration of the CM panel forecasts had a lesser Theil Uc (covariance element) forecasting error, as the participants could make cyclical forecasts.

Table 4-5 CM panel Forecasts - Gold Consensus Estimates Summary summarizes the gold forecasts made from October 2020 to December 2021 by the CM panel participants. The AR project's quantitative element focused on analyzing the collective forecasts to evaluate the reliability of the aggregated consensus estimates compared to the participating forecaster's predictions for each period. A significant relevant indicator in analyzing the CM panel forecasts was correctly predicting the future market gold price direction. The actual market gold price percentage changes highlight the actual market gold price direction compared to the anticipated market direction of the forecasters (*Table 4-5*; Column 2; Row a). The gold price aggregated consensus estimates forecasts errors are reported as the APE compared to the actual market gold price for the respective periods (*Table 4-5*; Columns 3-6; Row a). The consensus lower quartile was the smallest APE consistently for the seven periods (*Table 4-5*; Column 3; Row a). For the seven periods combined, the lowest APE was for the consensus lower quartile (*Table 4-5*; Column 6; Row b). The Theil U(II) Index compared to the consensus median was reported, with a value of less than 100% reflecting a more reliable consensus estimate than the consensus median, with only the consensus lower quartile yielding a value of less than 100% of 62.6% (*Table 4-5*; Column 3; Row c). The ReLMAE for each aggregated consensus estimate was calculated compared to the prior year's actual market gold price. Based on the ReLMAE, the consensus lower quartile (103.5%) was the most reliable estimate for the seven forecasting periods, compared to the consensus median (118.7%) (*Table 4-5*; Columns 3 & 5; Row e).

Table 4-5 CM panel Forecasts - Gold Consensus Estimates Summary

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		**LBMA average price for the prior period**	**LBMA average price for the period**	**Lower Quartile**	**Average**	**Median**	**Upper Quartile**
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	October 2020	$1,923/oz	$1,901/oz	$1,910/oz	$1,930/oz	$1,940/oz	$1,948/oz
a	Actual % Δ / APE %	**Neutral**	**-1.1%**	**0.5%**	1.5%	2.0%	2.4%
	November 2020	$1,901/oz	$1,866/oz	$1,870/oz	$1,910/oz	$1,930/oz	$1,955/oz
a	Actual % Δ / APE %	**Neutral**	**-1.9%**	**0.2%**	2.3%	3.4%	4.7%
	December 2020	$,,866/oz	$1,856/oz	$1,860/oz	$1,911/oz	$1,930/oz	$1,950/oz
a	Actual % Δ / APE %	**Neutral**	**-0.6%**	**0.2%**	2.9%	4.0%	5.1%
	March 2011	$1,856/oz	$1,796/oz	$1,850/oz	$1,906/oz	$1,860/oz	$1,960/oz
a	Actual % Δ / APE %	**Down**	**-3.4%**	**3.0%**	6.1%	3.6%	9.1%
	June 2021	$1,796/oz	$1,815/oz	$1,850/oz	$1,963/oz	$1,940/oz	$2,030/oz
a	Actual % Δ / APE %	**Neutral**	**1.1%**	**1.9%**	8.1%	6.9%	11.8%
	September 2021	$1,815/oz	$1,790/oz	$1,940/oz	$2,029/oz	$2,010/oz	$2,100/oz
a	Actual % Δ / APE %	**Neutral**	**-1.4%**	**8.4%**	13.4%	12.3%	17.3%
	December 2021	$1,790/oz	$1,795/oz	$1,950/oz	$2,058/oz	$2,020/oz	$2,160/oz
a	Actual % Δ / APE %	**Neutral**	**0.3%**	**8.6%**	14.7%	12.5%	20.3%
	Oct '20 – Dec '21	$1,850/oz	$1,832/oz	$1,890/oz	$1,958/oz	$1,947/oz	$2,015/oz
b	Actual % Δ / APE %	**Neutral**	**-1.0%**	**3.2%**	6.9%	6.3%	10.0%

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LBMA average price for the prior period	LBMA average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
c	Theil U(II) – Median			**62.6%**	113.2%	100.0%	157.4%
d	Theil U(II) – Prior			**280.2%**	507.0%	448.0%	705.1%
e	ReLMAE - Prior			**103.5%**	116.8%	118.7%	111.0%

From a cooperative inquiry perspective, to evaluate the possibility of using aggregated consensus metal price estimates in the workplace, the quantitative analysis did highlight a potential aggregated consensus estimate that could merit further consideration as a reliable forecasting alternative. For four of the seven forecasting periods, the lower quartile proved to be a more reliable forecast of the future gold price than the consensus average, median, or upper quartile. The CM panel gold average APE for the seven forecasting periods was average, median, upper quartile, and lower quartile 6.9%, 6.3%, 10.0%, and 3.2%, compared to the actual market gold price change of 1.0% (*Table 4-5*; Row b).

Between September 2020 and December 2021, the average market gold price decreased by around 7% (LBMA, 2020). The CM panel participants had not expected the marginally declining market gold price. Uncertainty surrounding the United States elections at the end of 2020 and its possible impact on the relative value of the United States Dollar was cited by several participants as potentially positive for a gold price rally. The economic fallout from COVID-19 was also mentioned by participants, with the expectation gold would act as a safe-haven investment. The expected lackluster world economy and increasing uncertainty because of the dominant events mentioned were expected to be positive for gold. Of the 84 forecasts made by the CM panel gold forecasters for the seven forecasting periods, 49 (58.3%) participants predicted a gold price more than 1% higher than the preceding period, and 35 (41.7%) that a gold price more than 1% lower than the preceding period. The market gold price remained subdued during the seven forecasting periods, so the further in the future the gold price forecasts were made, the less accurate they were. The CM panel gold forecasts in earlier periods were more accurate, reliably reflecting the shorter-term market gold price trends. The disparity between the CM panel gold participants' expectations of gold and market gold prices negatively affected the aggregated consensus estimates. The aggregated consensus estimates were prone to the forecasted gold price direction errors of the CM panel participants.

Table 4-6 CM panel – Gold Consensus Estimates Reliability

Cm panel gold participant	No. periods forecasts made	No. period most accurate	No. forecasts submitted	APE measure	APE (%)	Average Rank	Rank z-Score
Part. 1	7	0	19	Participant	13.9%	6	(1.00)
				Median	6.3%	5	(0.50)
				Q1	**3.2%**	3	0.50
Part. 2	7	3	15	**Participant**	**3.2%**	2	1.00
				Median	6.3%	5	(0.50)
				Q1	3.2%	3	0.50
Part. 3	7	2	15	Participant	3.7%	2	1.00
				Median	6.3%	5	(0.50)
				Q1	**3.2%**	3	0.50
Part. 4	7	3	8	Participant	4.4%	4	0.00
				Median	6.3%	5	(0.50)
				Q1	**3.2%**	3	0.50
Part. 5	3	0	7	Participant	2.0%	3	0.66
				Median	3.7%	6	(0.66)
				Q1	**1.1%**	2	1.09
Part. 6	6	1	6	Participant	9.0%	6	(0.66)
				Median	7.1%	6	(0.66)
				Q1	**3.7%**	3	0.66
Part. 7	5	1	5	Participant	6.0%	5	(0.22)
				Median	4.0%	5	(0.22)
				Q1	**1.1%**	2	1.09
Part. 8	4	0	4	Participant	8.8%	6	(0.66)
				Median	8.8%	6	(0.66)
				Q1	**5.5%**	3	0.66
Part. 9	3	1	3	Participant	4.5%	3	0.66
				Median	7.5%	6	(0.66)
				Q1	**4.3%**	3	0.66
Part. 10	1	0	1	Participant	2.0%	2	1.00
				Median	2.0%	5	(0.50)
				Q1	**0.5%**	4	0.00
Part. 11	1	0	1	Participant	11.5%	5	(0.22)
				Median	4.0%	6	(0.66)
				Q1	**0.2%**	3	0.66
11	**51**	**11**	**84**	**Total**			
Average APE / Average Rank / z-Score <1.4> (CL 90%)				**Participant**	**6.3%**	**4**	**0**
				Median	**5.7%**	**5**	**0**
				Q1	**2.7%**	**3**	**0**

Table 4-5 CM panel Forecasts - Gold Consensus Estimates Summary examined the CM panel gold forecasts from the perspective of the reliability of the aggregated consensus estimates. *Table 4-6 CM panel – Gold Consensus Estimates Reliability* considered the reliability of the individual participants. Around half, 5

(45.5%) of the 11 CM panel gold participants achieved a lower APE than the equivalent aggregated median APE. By comparison, only 1 (9.1%) of the 11 CM panel gold participants had a lower APE than the equivalent APE for the Lower Quartile aggregated estimate. For the 3 participants with the lowest APE (*Table 4-6*, Participants 5,10 & 2), the lowest two only submitted copper price forecasts in 3 and 1 of the seven forecasting periods, respectively. Using the CM panel gold participants' APE in each forecasting period, they submitted forecasts; a ranking was assigned from most to least accurate. The participant's rankings achieved for each period they participated were averaged to obtain an overall ranking. Considering the comparable average number of participants, a z-Score of the ranking was calculated for each participant. At a 90% confidence limit, no participants (*Table 4-6*) had a positive z-Score exceeding the 1.4 confidence limit threshold. Significantly, the average APE (3.16%) for the single participant (*Table 4-6*, Participant 2) was only marginally lower than the equivalent APE for the Lower Quartile (3.19%).

The CM panel gold aggregated consensus metal price estimates from October 2020 to December 2021 were evaluated for accuracy using the Theil U(II) Index based on the consensus median (*Table 4-5*; Row d). The RMSE forecast error for the actual versus forecast (Nominator) and actual versus median (Denominator) was decomposed into their respective bias (Um), variance (Us), and covariance (Uc) error elements to highlight the relevant impact of each in the overall forecast error. The formulation of Theil's U(II) Index is comparable to the Relative Mean Absolute Error (RelMAE) when the prior period is used as the reference comparison. The Theil U(II) Index based on the preceding period and the RelMAE were computed to evaluate the forecasting error (*Table 4-5*; Row d & e).

Table 4-7 CM Gold Theil U(II) Index RMSE Decompositions

Theil U(II) Error Elements	RMSE	Um	Us	Uc
Nominator: Average (Actual versus Forecast)	154	67.6%	1.3%	31.1%
Nominator: Lower quartile (Actual versus Forecast)	85	47.2%	0.0%	52.7%
Nominator: Upper quartile (Actual versus Forecast)	214	73.2%	3.3%	23.5%
Denominator: Actual versus Median	136	72.2%	0.5%	27.2%
Denominator: Actual versus Preceding period	30	36.1%	6.3%	57.6%

The unanticipated decrease in the market gold price by the CM panel participants was reflected in the higher RMSE values for the aggregated consensus metal price forecasting estimates, including the consensus median. The lower RMSE for the prior period was because the downward trending gold price was more effectively captured

by the preceding period average, which had a shorter lead time before the forecasting period. The incorrectly predicted market gold price direction by the CM panel participants was reflected in the substantial Theil Um (bias element) contribution to the forecasting error. The forecasting error could have been reduced by having more participants, assuming they did not share similar gold price expectations. The minimal contribution of Theil Us (variance element) to the forecasting error of the CM panel participants reflected the slight change in their gold price expectations over the successive cycles of forecasting. Between October 2020 and December 2021, the decreasing gold price resulted in a lagged change in the preceding period value compared to the actual period value, reflected in the significant covariance error.

4.4. Consensus Metals Forecasts Qualitative Results

In structuring the cooperative inquiry, an incorporated element was the influence interaction could have on the CM panel participants in making their metal price forecasts. The opportunity for the CM panel participants to act, see the interactions of other participants, reflect on the shared insights, and then act again by making further cyclical forecasts was part of the cooperative inquiry's intention of evaluating the importance of shared interactions. For the period October 2020 to December 2021, the CM panel participants made copper and gold price forecast, together with a brief justification for the basis of their forecasts. The following summarizes the forecast justifications posted by the CM panel participants:

- 86 (~80%) of the 107 copper forecasts included a justification.
- 75 (~77%) of the 84 gold forecasts included a justification.
- For the 86 copper justifications made, the average word count was 27.
- For the 65 gold justifications made, the average word count was 25.
- None of the participants took the opportunity to comment specifically on the metal price forecasts of the other CM panel participants.
- Two copper CM panel participants referred to the emerging consensus average, and only one gold CM panel participant made a similar reference.
- The correlation between the CM panel participants' ranks, based on their forecast APE, and justification word count was ~31% for copper and ~39% for gold, reflecting a weak positive relationship between APE (rank) and word count.

Table 4-8 CM panel Keyword Summary tabulates the occurrence of the most frequent words used by the CM panel participants in their copper and gold justifications, with the percentages compared to the respective total word counts for the two metals.

Table 4-8 CM panel Keyword Summary

Keyword	Copper			Gold		
	Rank	Frequency	%	Rank	Frequency	%
Economic	1	77	3.1%	3	38	2.29%
Demand	2	68	2.8%	4	33	1.99%
Covid	3	56	2.3%	5	32	1.93%
World	4	56	2.3%	1	70	4.23%
Supply	5	33	1.3%	7	10	0.60%
USA	6	26	1.1%	10	5	0.30%
US election	7	22	0.9%	2	43	2.60%
Uncertainty	8	15	0.6%	6	14	0.85%
Trend	9	12	0.5%	7	10	0.60%
Metals	10	9	0.4%	11	1	0.06%
US Dollar	11	5	0.2%	9	7	0.42%

The unfolding world events were the most significant factor cited for the CM panel copper justifications, the state of the global economy, and the CM panel gold justifications. During the cooperative inquiry, the copper and gold CM panel participants often noted the topical issues of COVID-19 and the United States election. The economic impact of COVID-19 was seen as detrimental to the copper price, and the uncertainty surrounding the United States elections was positive for the gold price. Demand and supply fundamentals were mentioned in the justifications for both metals; however, demand was emphasized more frequently.

The CM panel participants were encouraged to make cyclical metal price forecasts based on their evolving insights from the cooperative inquiry's observed interactions. *Table 4-9 CM panel Cyclical Forecasts - Copper and Gold* summarize the CM panel participants' cyclical forecasts for each forecasting period.

Table 4-9 CM panel Cyclical Forecasts - Copper and Gold

Copper	Cyclical Forecast Frequency					Total Number of Forecasts
	One	Two	Three	Four	Five	
October 2020	7	1	0	0	0	**9**
November 2020	6	3	0	0	0	**12**
December 2020	4	4	1	1	0	**19**
March 2021	4	2	3	0	1	**22**
June 2021	3	5	1	0	0	**16**
September 2021	4	4	1	0	0	**15**

Copper	Cyclical Forecast Frequency					Total Number of Forecasts
	One	Two	Three	Four	Five	
December 2021	3	4	1	0	0	**14**
Total	**31**	**46**	**21**	**4**	**5**	**107**

Gold	Cyclical Forecast Frequency					Total Number of Forecasts
	One	Two	Three	Four	Five	
October 2020	6	0	0	0	0	**6**
November 2020	3	4	0	0	0	**11**
December 2020	3	3	2	0	0	**15**
March 2021	5	0	4	0	0	**17**
June 2021	5	1	2	0	0	**13**
September 2021	4	2	1	0	0	**11**
December 2021	3	2	0	1	0	**11**
Total	**29**	**24**	**27**	**4**	**0**	**84**

Of the copper forecasts made in 31 (~29%) instances, the CM panel participants only made one forecast in a particular forecasting period. The most common occurrence was 46 (~43%) instances of making two forecasts in a forecasting period. In 30 (~28%) instances, the copper CM panel participants made more than two forecasts per forecasting period. 30 of the 32 cases when the copper CM panel participants made more than one forecast in a forecasting period, the last forecast was more accurate than the first. The average improvement for the 30 instances of the better forecast was 10.7%, while the average deterioration for the two worse instances of forecasts was 2.5%.

For the gold forecasts made by the CM panel participants, the instance in which the participants only made one forecast in a particular forecasting period was 29 or about ~35% of the total forecasts made. The most common practice was making two (24) or three (27) forecasts in each forecasting period, accounting for 51 (~61%) of the total forecasts made. Only one gold CM panel participant made four (~5%) forecasts in December 2020. 21 Out of the 22 cases, when the gold CM panel participants made more than one forecast per forecasting period was more accurate than the first. The average improvement for the 21 instances of the better forecast was 5%, while the average deterioration for the worse forecast was 2.8%.

4.5. **LBMA Precious Metals Survey Results**

Between 2000 and 2020, 127 different analysts participated in the annual LBMA-PMS, with 545 forecasts made over the twenty-one years. The matching comments justifying their predictions totaled around 92’700 words. Of the analysts taking part, circa 38% only took part for one year, and approximately 37% participated for four or more years at the other end of the participation occasions. *Table 4-10 LBMA-PMS*

Analysts' Participation Summary tabulates the average APE and word count for the twenty-one years of the LMBA-PMS participants aggregated by the analysts' annual participation frequency rate.

- The average APE for the one to two years of participation was approximately 9%, and the average word count of their forecast justifications was 146.
- The average APE for participation between three and five years declined to around 6.9%, and the average word count of their forecast justifications was 163.
- The average APE for participation between six and fifteen years worsened to approximately 8.4%, and the average word count of their forecast justifications was 173.
- The average APE for participation between sixteen and twenty-one years improved to around 7.3%, and the average word count of their forecast justifications was 189.

The average word count for the participants with minimal versus maximum years of participation had an average difference of around 30%.

Table 4-10 LBMA-PMS Analysts' Participation Summary

Number of years participated	Number of participants	% Of Participants	Average APE	Average word count of justifications
1 Year	48	37.8%	9.2%	145
2 Years	23	18.1%	8.6%	149
3 Years	9	7.1%	7.3%	144
4 Years	8	6.3%	6.5%	177
5 Years	5	3.9%	6.9%	175
6 Years	5	3.9%	8.0%	169
7 Years	6	4.7%	7.8%	167
8 Years	2	1.6%	8.5%	208
9 Years	6	4.7%	9.0%	175
10 Years	1	0.8%	9.8%	161
11 Years	1	0.8%	9.0%	120
12 Years	3	2.4%	7.9%	178
13 Years	1	0.8%	8.3%	183
14 Years	2	1.6%	8.9%	211
15 Years	1	0.8%	8.6%	100
16 Years	2	1.6%	6.9%	222
19 Years	2	1.6%	7.5%	145
21 Years	2	1.6%	7.4%	200
Total	**127**	**100.0%**	**8.4%**	**171**

The correlation between years taken part and APE was approximately 21%, reflecting the inconsistent variation in APE compared to years of participation (*Figure*

4-1 LBMA-PMS Gold Forecasters' APE). The correlation between years participated, and word count was around 23%, reflecting a marginal increase with years of participation.

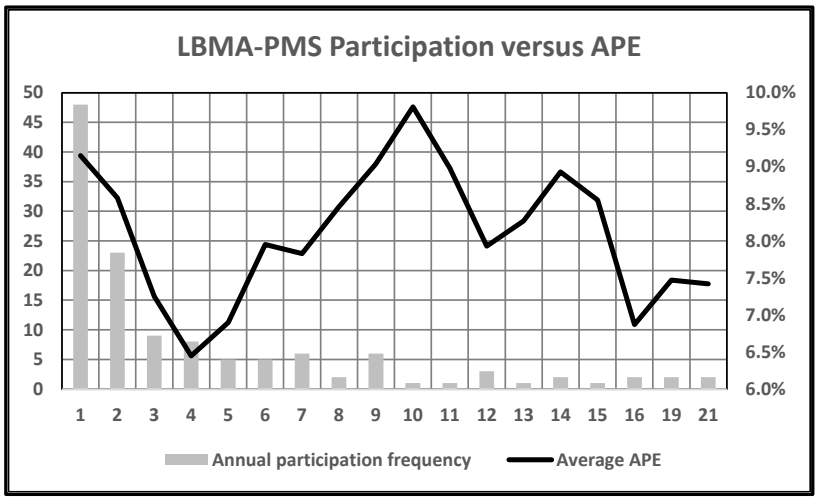


Figure 4-1 LBMA-PMS Gold Forecasters' APE

4.6. **LBMA Precious Metals Survey Quantitative Results**

The prediction-realization diagram (Koutsoyiannis, 1977) *Figure 4-2 LBMA-PMS Gold Prediction Realization Diagram* reflects the percentage error of the average consensus forecasts made by participants in the annual LBMA-PMS between 2000 and 2020 compared to the actual year-on-year change in the gold price. In the prediction-realization diagram in *Figure 4-2*, the 45-degree line through the origin is "*the line of perfect forecasts*." It distinguishes between forecasts above or below the actual subsequent outcome, with the perfect forecasts falling on the line.

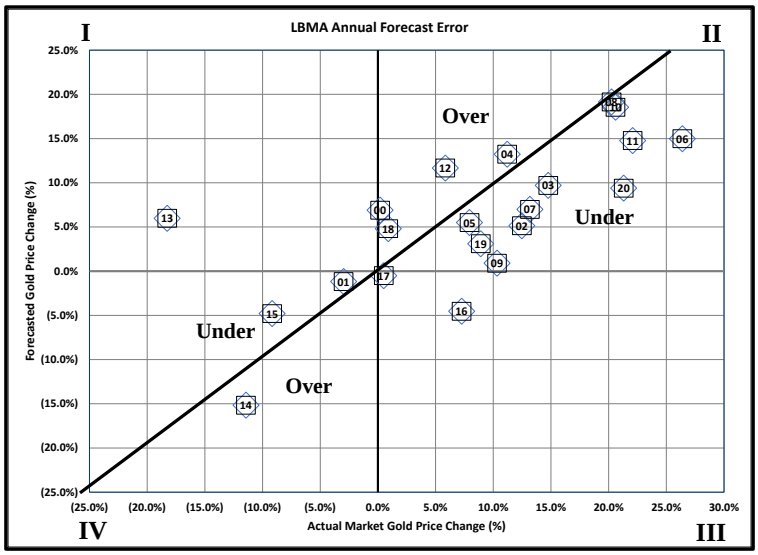


Figure 4-2 LBMA-PMS Gold Prediction Realization Diagram

Points in either quadrants II or IV indicate when the LBMA-PMS participants correctly predicted the annual direction of the market gold price. For quadrants I and III, the points are when the LBMA-PMS forecasters incorrectly predicted the market's gold price direction. The details of the two outliers in quadrants I (2013) and III (2016) are tabulated below, highlighting the forecasters' directional error, albeit small for 2016.

Table 4-11 LBMA-PMS Actual and Forecasted Gold Price Direction

Year	Prior year market gold price	Actual year's market gold price	Actual market direction change	Consensus median forecasted gold price	Forecasted market direction change	% Of participants anticipated direction incorrectly
2013	$1'699/oz	$1'411/oz	**Down (-17%)**	$1'753/oz	**Up (3.2%)**	**78.3%**
2016	$1'160/oz	$1'251/oz	**Up (7.8%)**	$1'120/oz	**Down (-3.4%)**	**87.1%**

From 2002 to 2012 and from 2016 to 2020, the market gold price rose. For the sixteen years, the gold price increased by more than 1%, and on only three occasions was the consensus median forecasted higher than the actual gold price, 2004, 2012, and 2018. For the other thirteen (81%) years, the LBMA-PMS underestimated the future gold price compared to the prior year. In 2000, 2001, and from 2013 to 2015, the market gold price increased by less than 1%, and only in 2014 (20%) did the LBMA-PMS forecasters correctly predict a gold price decrease.

The quantitative analysis evaluated the reliability and consistency of the respective aggregated consensus estimates compared to the participating LBMA-PMS gold forecasters for the twenty-one years analyzed. A significant factor in analyzing the LBMA-PMS data was the importance of predicting the market direction of the forecasted gold price, with seventeen (81%) of the twenty-one years analyzed correctly predicted. The LBMA-PMS forecasts were evaluated by calculating the APE compared to the actual market gold price and checking whether the participants correctly predicted the future market gold price direction. The annual LBMA-PMS forecasts were made after the first week of each year, making it possible to differentiate the January market gold price from the prior year's average as either 1% higher or lower. The January gold price allowed the LMBA-PMS participants' gold price predictions to be categorized based on the actual relative price change. If the LBMA-PMS forecasters predicted higher or lower expected gold prices, it was to calculate the posterior probabilities of the different gold price directional changes over the twenty-one years. Calculating the posterior probabilities for the gold price forecasting

directional changes opens the possibility of using the estimated probabilities as prior probabilities in a Bayesian approach to assess the reliability of future LBMA-PMS forecasts (Annexure F: LBMA-PMS gold forecasts probabilities) (Greenberg, 2012).

Over the period 2000 to 2020, the annual average market price of gold increased by more than 1% fourteen times, decreased by less than 1% four times, and on three occasions, the annual change was less than 1%. *Table 4-12* tabulates the actual and forecasted gold price change for the twenty-one years by the LBMA-PMS participants, categorized into either a gold price change of more than 1% or less than 1%.

Table 4-12 LBMA-PMS Gold Forecasted and Actual Direction

Description	Units	Total	More than 1% Δ	Less than 1% Δ
LBMA annual gold price change				
Market gold price change	Years	21	14	7
Market gold price change	%	100.0%	66.7%	33.3%
Average forecasted gold price change	Years	21	16	5
Average forecasted gold price change	%	100.0%	76.1%	23.8%
Market gold price change correctly forecasted	Years	17	13	4
Market gold price change correctly forecasted	**%**	**80.9%**	**81.3%**	**80.0%**
LBMA total forecasts made				
Total forecasts by gold price direction change	No.	545	397	148
Total forecasts by gold price direction change	%	100.0%	72.8%	27.2%
Market gold price change correctly forecasted	No.	413	316	97
Market gold price change correctly forecasted	**%**	**75.8%**	**79.6%**	**65.5%**
The first week of January change positive compared to the prior year's average				
Market gold price change	Years	13	12	1
Market gold price change	%	100.0%	92.3%	7.7%
Average forecasted gold price change	Years	13	13	0
Average forecasted gold price change	%	100.0%	100.0%	0.0%
Market gold price change correctly forecasted	Years	12	12	0
Market gold price change correctly forecasted	**%**	**92.3%**	**92.3%**	**0.0%**
Total forecasts by gold price direction change	No.	336	318	18
Total forecasts by gold price direction change	%	100.0%	94.6%	5.4%
Market gold price change correctly forecasted	No.	300	297	3
Market gold price change correctly forecasted	**%**	**89.3%**	**93.4%**	**16.7%**
The first week of January change negative compared to the prior year's average				
Market gold price change	Years	8	2	6
Market gold price change	%	100.0%	25.0%	75.0%
Average forecasted gold price change	Years	8	3	5
Average forecasted gold price change	%	100.0%	37.5%	62.5%
Market gold price change correctly forecasted	Years	5	1	4
Market gold price change correctly forecasted	**%**	**62.5%**	**33.3%**	**80.0%**
Total forecasts by gold price direction change	No.	209	79	130
Total forecasts by gold price direction change	%	100.0%	37.8%	62.2%
Market gold price change correctly forecasted	No.	113	19	94

Description	Units	Total	More than 1% Δ	Less than 1% Δ
Market gold price change correctly forecasted	**%**	**54.1%**	**24.1%**	**72.3%**

Of the sixteen years, the LBMA-PMS participants predicted a market gold price change of more than 1%; they were correct in thirteen (81.3%) years. Out of the five years, the LBMA-PMS participants expected a market gold price change of less than 1%; they were correct in four (80%) years. The comparable ratio of correctly predicting the market gold price change in terms of the annual number of forecasts made was similar for gold price changes of more than 1% but less so for the converse of gold price of less than 1%. Of the LBMA-PMS forecasts predicting a market gold price change of more than 1%, 316 (79.6%) out of 397 were correct, but only 97 (65.5%) out of 148 were right for a market gold price change of less than 1%.

The gold price direction change and the gold price directional change in the first week of January improved the accuracy ratio of the LBMA-PMS participants if they chose the same direction for their forecasted gold prices. When the January relative gold price change was more than 1%, the LBMA-PMS participants correctly predicted a gold price change of more than 1% in twelve (92.3%) out of thirteen years and 297 (93.4%) out of 318 forecasts. When the January relative gold price change was less than 1%, the LBMA-PMS participants correctly predicted the gold price change of less than 1% in four (80.0%) out of five years and 94 (72.3%) out of 130 forecasts. When the January relative gold price change was more than 1%, and the subsequent year's market gold price change was less than 1% (2018), only 3 (12.5%) out of 24 LBMA-PMS participants correctly forecasted the gold price change for the year. When the January relative gold price change was less than 1%, and the subsequent year's market gold price change was more than 1% (2009 and 2016), only 19 (35.5%) out of 55 LBMA-PMS participants correctly predicted the gold price change direction in for the year.

In 2000 and 2018, the average year-on-year market gold price change was marginal, 0.2% and 0.9%, respectively, while the LBMA-PMS participants had predicted more significant increases. When the LBMA-PMS participants incorrectly anticipated the market gold price in 2013 and 2016, the gold market experienced an inflection direction change for two years. Following eleven years in which the gold price had increased on average by around 12%, in 2013, it fell by 18.3%. In 2016 the market gold price rose by 7.3%, following the previous three years in which the gold price had decreased on average by circa 13%.

Table 4-13 LBMA-PMS Gold Forecasts - Consensus Estimates Summary summarizes the gold forecasts made between 2000 and 2020 by the LBMA-PMS participants. The LBMA-PMS gold forecasts were analyzed comparably to allow evaluation of the reliability of the aggregated consensus estimates for each of the twenty-one years, to triangulate the AR project quantitative results from the CM panel. As with the CM panel forecasts, correctly predicting the market metal price direction was a significant factor in finding a reliable aggregated consensus estimate that could be considered for use in the workplace. The actual annual average market gold price percentage changes compared to the prior year's average market gold prices were calculated to analyze the accuracy of the LBMA-PMS forecasters in predicting the market gold price direction change (*Table 4-13*; Column 2; Row a). The gold price aggregated consensus estimates forecasts errors are reported as the APE compared to the actual market gold price for the respective years (*Table 4-13*; Columns 3-6; Row a). The APE for the combined twenty-one years of the LBMA-PMS is reported for the aggregated consensus estimates compared to the actual year-on-year change in the market gold price (*Table 4-13*; Columns 3-6; Row b). The RMSE used to calculate the Theil U(II) Index for the aggregated consensus estimates is reported (*Table 4-13*; Columns 3-6; Row c). Compared to the consensus median, the Theil U(II) Index, with a less than 100% value, reflects a more reliable consensus estimate than the median (*Table 4-13*; Columns 3-6; Row d). The Theil U(II) Index, compared to the prior year's actual average gold price, is reported for comparative purposes to the ReLMAE (*Table 4-13*; Columns 3-6; Row e). The ReLMAE for each aggregated consensus estimate compared to the prior year's actual market copper price was calculated (*Table 4-13*; Columns 3-6; Row f).

Table 4-13 LBMA-PMS Gold Forecasts - Consensus Estimates Summary

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		**LBMA average price for the prior period**	**LBMA average price for the period**	**Lower Quartile**	**Average**	**Median**	**Upper Quartile**
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	2000	$279/oz	$279/oz	$285/oz	$298/oz	$295/oz	$305/oz
a	Actual % Δ / APE %	**Neutral**	**0.2%**	**2.1%**	6.7%	5.7%	9.3%
	2001	$279/oz	$271/oz	$268/oz	$276/oz	$275/oz	$283/oz
a	Actual % Δ / APE %	**Down**	**(3.0%)**	**1.1%**	1.8%	1.5%	4.3%
	2002	$271/oz	$310/oz	$280/oz	$287/oz	$290/oz	$293/oz
a	Actual % Δ / APE %	**Up**	**12.5%**	9.6%	7.3%	6.4%	**5.6%**
	2003	$310/oz	$363/oz	$335/oz	$345/oz	$345/oz	$351/oz
a	Actual % Δ / APE %	**Up**	**14.8%**	7.8%	5.1%	5.0%	**3.5%**
	2004	$363/oz	$409/oz	$403/oz	$417/oz	$419/oz	$429/oz
a	Actual % Δ / APE %	**Up**	**11.2%**	**1.5%**	2.0%	2.4%	4.7%

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LBMA average price for the prior period	LBMA average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	2005	$409/oz	$444/oz	$422/oz	$434/oz	$438/oz	$450/oz
a	Actual % Δ / APE %	**Up**	**7.9%**	5.1%	2.4%	1.5%	**1.3%**
	2006	$444/oz	$604/oz	$520/oz	$535/oz	$525/oz	$554/oz
a	Actual % Δ / APE %	**Up**	**26.4%**	13.9%	11.4%	13.1%	**8.2%**
	2007	$604/oz	$695/oz	$620/oz	$652/oz	$650/oz	$675/oz
a	Actual % Δ / APE %	**Up**	**13.2%**	10.8%	6.2%	6.5%	**3.0%**
	2008	$695/oz	$872/oz	$825/oz	$862/oz	$850/oz	$893/oz
a	Actual % Δ / APE %	**Up**	**20.3%**	5.4%	**1.1%**	2.5%	2.4%
	2009	$872/oz	$972/oz	$836/oz	$881/oz	$901/oz	$946/oz
a	Actual % Δ / APE %	**Up**	**10.3%**	14.1%	9.4%	7.4%	**2.7%**
	2010	$972/oz	$1,225/oz	$1,165/oz	$1,199/oz	$1,199/oz	$1,233/oz
a	Actual % Δ / APE %	**Up**	**20.6%**	4.8%	2.1%	2.1%	**0.7%**
	2011	$1,225/oz	$1,572/oz	$1,449/oz	$1,457/oz	$1,464/oz	$1,491/oz
a	Actual % Δ / APE %	**Up**	**22.1%**	7.8%	7.3%	6.8%	**5.1%**
	2012	$1,572/oz	$1,669/oz	$1,728/oz	$1,766/oz	$1,770/oz	$1,833/oz
a	Actual % Δ / APE %	**Up**	**5.8%**	**3.5%**	5.8%	6.1%	9.8%
	2013	$1,669/oz	$1,411/oz	$1,714/oz	$1,753/oz	$1,753/oz	$1,782/oz
a	Actual % Δ / APE %	**Down**	**(18.3%)**	**21.4%**	24.3%	24.2%	26.2%
	2014	$1,411/oz	$1,266/oz	$1,176/oz	$1,219/oz	$1,230/oz	$1,265/oz
a	Actual % Δ / APE %	**Down**	**(11.4%)**	7.2%	3.7%	2.9%	**0.1%**
	2015	$1,266/oz	$1,160/oz	$1,188/oz	$1,211/oz	$1,230/oz	$1,255/oz
a	Actual % Δ / APE %	**Down**	**(9.2%)**	**2.4%**	4.4%	6.0%	8.2%
	2016	$1,160/oz	$1,251/oz	$1,058/oz	$1,103/oz	$1,120/oz	$1,145/oz
a	Actual % Δ / APE %	**Up**	**7.3%**	15.5%	11.8%	10.5%	**8.5%**
	2017	$1,251/oz	$1,257/oz	$1,209/oz	$1,244/oz	$1,260/oz	$1,285/oz
a	Actual % Δ / APE %	**Neutral**	**0.5%**	3.8%	1.0%	**0.2%**	2.2%
	2018	$1,257/oz	$1,268/oz	$1,287/oz	$1,318/oz	$1,321/oz	$1,359/oz
a	Actual % Δ / APE %	**Neutral**	**0.9%**	**1.5%**	3.9%	4.1%	7.1%
	2019	$1,268/oz	$1,393/oz	$1,300/oz	$1,312/oz	$1,315/oz	$1,326/oz
a	Actual % Δ / APE %	**Up**	**8.9%**	6.7%	5.8%	5.6%	**4.8%**
	2020	$1,393/oz	$1,770/oz	$1,521/oz	$1,559/oz	$1,559/oz	$1,593/oz
a	Actual % Δ / APE %	**Up**	**21.3%**	14.0%	11.9%	11.9%	**10.0%**
	2000-2020	$903/oz	$974/oz	$933/oz	$959/oz	$962/oz	$988/oz
b	Actual % Δ / APE %	**Up**	**12.3%**	8.3%	7.2%	7.1%	**6.8%**
c	Root Mean Square error	160		113	106	**105**	107
d	Theil U(II) – Median	153.1%		107.9%	**101.4%**	100.0%	102.2%
e	Theil U(II) – Prior			70.5%	66.2%	**65.3%**	66.7%
f	ReLMAE - Prior			104.4%	82.6%	82.3%	**65.9%**

A general pattern can be discerned, albeit not a perfect fit, with sixteen (76.2%) of the twenty-one years conforming with the following two prediction rules.

- The most accurate aggregated consensus estimate is the upper quartile if the market change year-on-year is greater than 1%.
- The most accurate aggregated consensus estimate is the lower quartile if the market direction year-on-year is less than 1%.

For the twenty-one years of the annual LBMA-PMS, the aggregated consensus estimate with the lowest absolute percentage error is tabulated in *Table 4-14*.

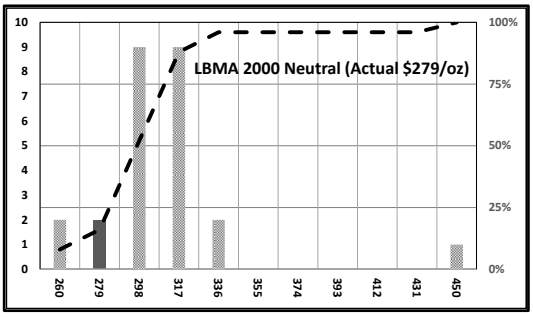
Table 4-14 LBMA-PMS Aggregated Consensus Estimates Reliability

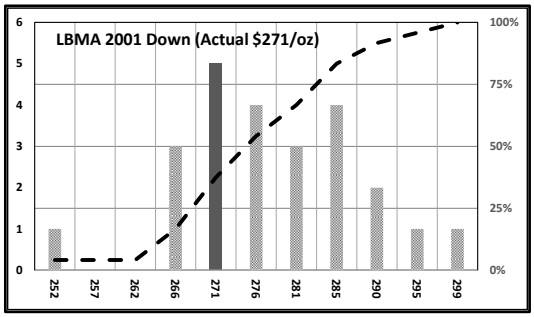
Direction	Average	Median	Lower Quartile	Upper Quartile	Total
Up by more than 1%	1 (2008)		2 (2004, 2012)	11	**14**
Down by more than 1%		1 (2017)	5	1 (2014)	**7**
Total	**1**	**1**	**7**	**12**	**21**

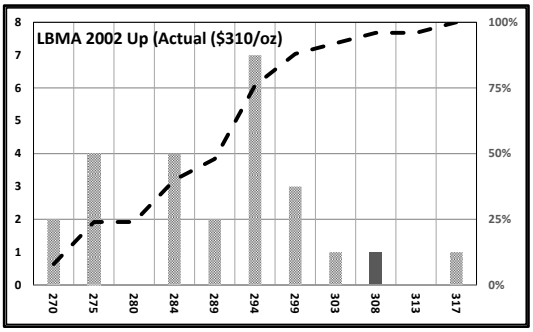
The lower quartile accounted for 7 (33.3%) of the twenty-one years, with 5 (71.4%) of the seven years occurring in years when the market gold price declined by more than 1%. The upper quartile accounted for 12 (57.1%) of the twenty-one years, with 11 (78.6%) of the fourteen years occurring in years when the market gold price increased by more than 1%. The five exceptions to the lower / upper quartile rules were:

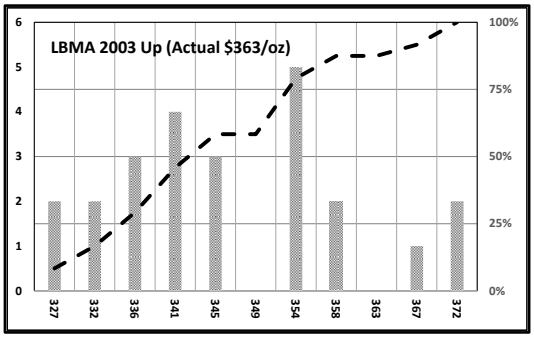
- 2004 (Annual average gold price increase of 11.2%): The lower quartile had a forecast error of 1.5%. The median forecast error of 2.4%, and the upper quartile's forecast error of 4.7%.
- 2008 (Annual average gold price increase of 20.3%): The average had a forecast error of 1.1%. The median forecast error was 2.5%, and the upper quartile forecast error was 2.4%.
- 2012 (Annual average gold price increase of 5.8%): The lower quartile had a forecast error of 3.5%. The median forecast error was 6.1%, and the upper quartile forecast error was 9.8%.
- 2014 (Annual average gold price decrease of 11.4%): The upper quartile had a forecast error of 0.1%. The median forecast error was 2.9%, and the lower quartile forecast error was 7.2%.
- 2017 (Annual average gold price increase of 0.5%): The median had a forecast error of 0.2%. The median forecast error was 0.2%, and the lower quartile forecast error was 3.8%.

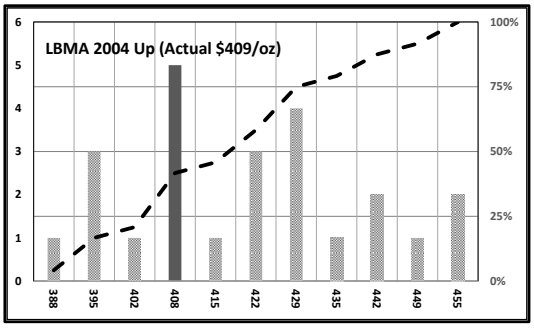
The LBMA-PMS annual forecasts were grouped into eleven equal annual gold price bands and graphed with the cumulative frequency percent. The point was to see if the spread of forecasts above and below the central value of the average LBMA gold price for the twenty-one years conformed with the expectation of balanced error dispersion (*Figure 4-3 LBMA-PMS Gold Forecasts Distribution in Incremental Price Buckets*).

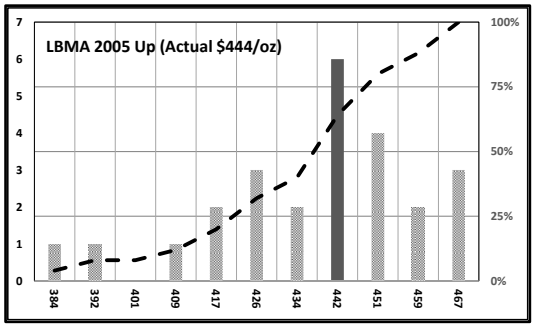

LBMA 2000 Neutral (Actual $279/oz)
0 1 2 3 4 5 6 7 8 9 10
0% 25% 50% 75% 100%
260 279 298 317 336 355 374 393 412 431 450

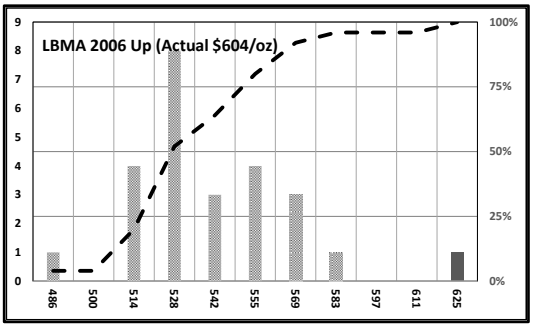

LBMA 2001 Down (Actual $271/oz)
0 1 2 3 4 5 6
0% 25% 50% 75% 100%
252 257 262 266 271 276 281 285 290 295 299

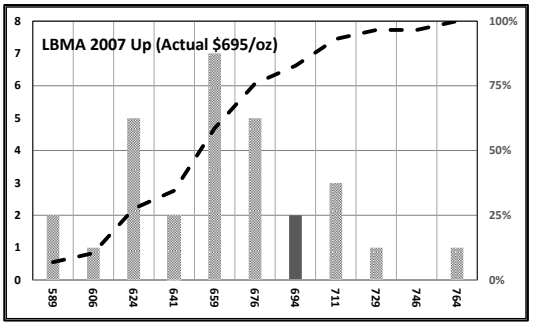

LBMA 2002 Up (Actual ($310/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
270 275 280 284 289 294 299 303 308 313 317

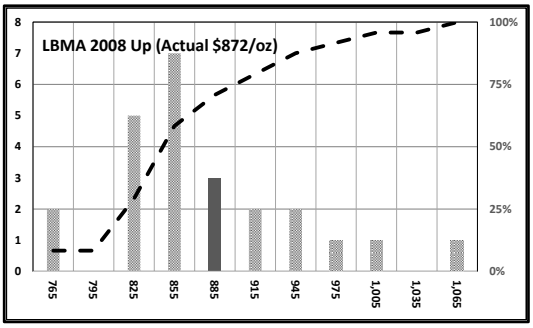

LBMA 2003 Up (Actual $363/oz)
0 1 2 3 4 5 6
0% 25% 50% 75% 100%
327 332 336 341 345 349 354 358 363 367 372

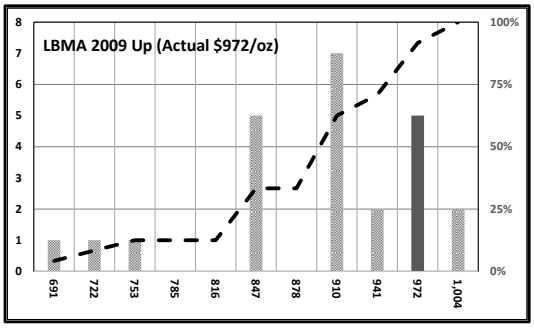

LBMA 2004 Up (Actual $409/oz)
0 1 2 3 4 5 6
0% 25% 50% 75% 100%
388 395 402 408 415 422 429 435 442 449 455

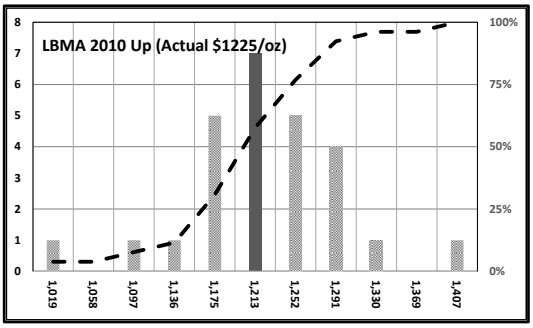

LBMA 2005 Up (Actual $444/oz)
0 1 2 3 4 5 6 7
0% 25% 50% 75% 100%
384 392 401 409 417 426 434 442 451 459 467

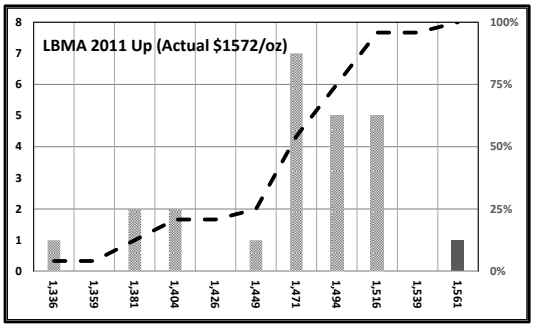

LBMA 2006 Up (Actual $604/oz)
0 1 2 3 4 5 6 7 8 9
0% 25% 50% 75% 100%
486 500 514 528 542 555 569 583 597 611 625

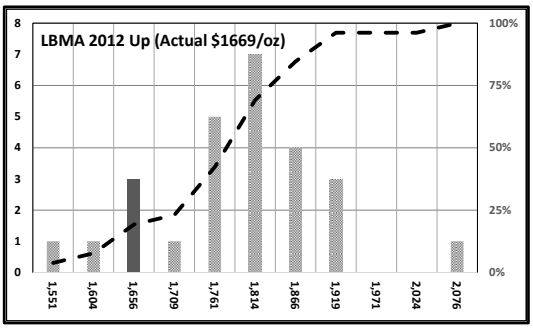

LBMA 2007 Up (Actual $695/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
589 606 624 641 659 676 694 711 729 746 764

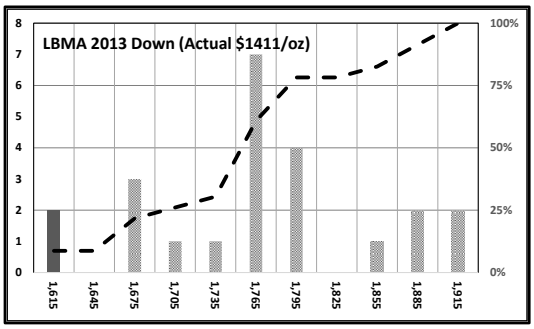

LBMA 2008 Up (Actual $872/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
765 795 825 855 885 915 945 975 1,005 1,035 1,065

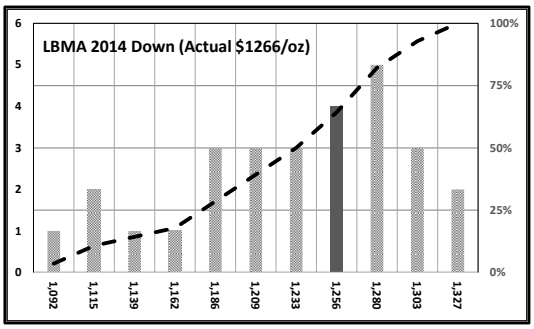

LBMA 2009 Up (Actual $972/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
691 722 753 785 816 847 878 910 941 972 1,004

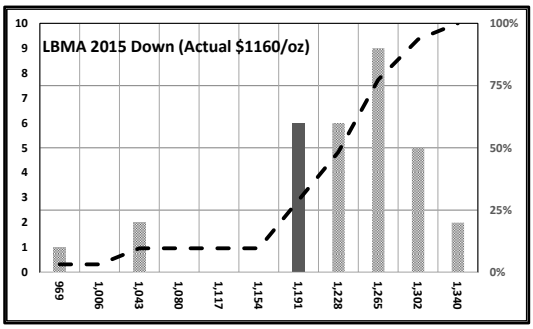

LBMA 2010 Up (Actual $1225/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
1,019 1,058 1,097 1,136 1,175 1,213 1,252 1,291 1,330 1,369 1,407

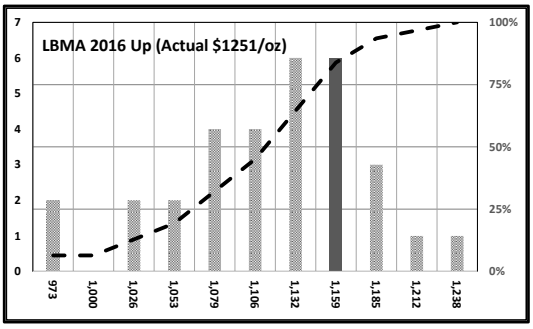

LBMA 2011 Up (Actual $1572/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
1,336 1,359 1,381 1,404 1,426 1,449 1,471 1,494 1,516 1,539 1,561

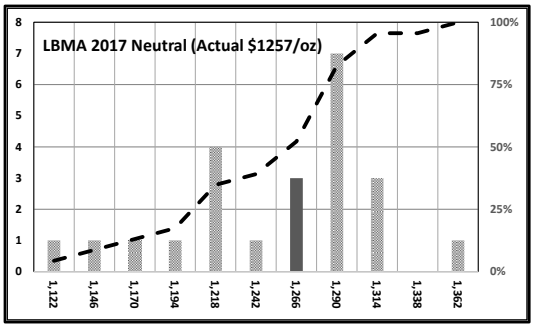

LBMA 2012 Up (Actual $1669/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
1,551 1,604 1,656 1,709 1,761 1,814 1,866 1,919 1,971 2,024 2,076

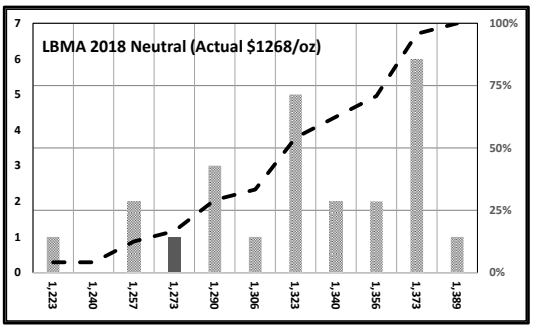

LBMA 2013 Down (Actual $1411/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
1,615 1,645 1,675 1,705 1,735 1,765 1,795 1,825 1,855 1,885 1,915

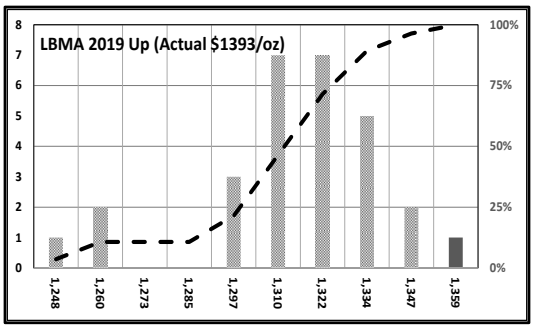

LBMA 2014 Down (Actual $1266/oz)
0 1 2 3 4 5 6
0% 25% 50% 75% 100%
1,092 1,115 1,139 1,162 1,186 1,209 1,233 1,256 1,280 1,303 1,327


LBMA 2015 Down (Actual $1160/oz)
0 1 2 3 4 5 6 7 8 9 10
0% 25% 50% 75% 100%
969 1,006 1,043 1,080 1,117 1,154 1,191 1,228 1,265 1,302 1,340


LBMA 2016 Up (Actual $1251/oz)
0 1 2 3 4 5 6 7
0% 25% 50% 75% 100%
973 1,000 1,026 1,053 1,079 1,106 1,132 1,159 1,185 1,212 1,238


LBMA 2017 Neutral (Actual $1257/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
1,122 1,146 1,170 1,194 1,218 1,242 1,266 1,290 1,314 1,338 1,362


LBMA 2018 Neutral (Actual $1268/oz)
0 1 2 3 4 5 6 7
0% 25% 50% 75% 100%
1,223 1,240 1,257 1,273 1,290 1,306 1,323 1,340 1,356 1,373 1,389


LBMA 2019 Up (Actual $1393/oz)
0 1 2 3 4 5 6 7 8
0% 25% 50% 75% 100%
1,248 1,260 1,273 1,285 1,297 1,310 1,322 1,334 1,347 1,359

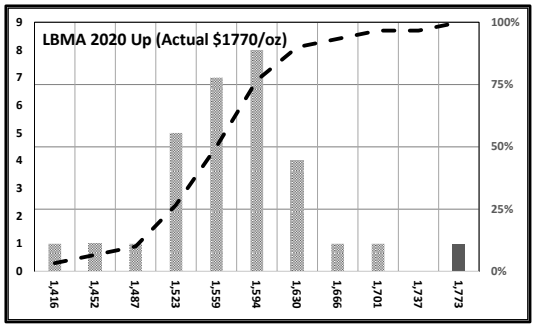


Figure 4-3 LBMA-PMS Gold Forecasts Distribution in Incremental Price Buckets

As observed from the twenty-one graphs, the distribution of the forecasts is not equally distributed above and below the LMBA annual average gold price. The extent of the skew in the predictions can be seen in *Table 4-15*, which tabulates the LBMA-PMS forecasts above and below the LMBA annual average gold price. The comparable distributions of the LBMA-PMS forecasts relative to the LBMA-PMS forecast average are included for comparative purposes.

Table 4-15 LBMA-PMS Annual Forecasts Distributions

Year	Market gold direction	LBMA Annual Average	Below LBMA Annual Average	Above LBMA Annual Average	Lowest LBMA-PMS Forecast	Highest LBMA-PMS Forecast	Lowest APE consensus estimate
2000	Neutral	$279/oz	16.0%	84.0%	$250/oz	$440/oz	Quartile 1
2001	Down	$271/oz	37.5%	62.5%	$250/oz	$297/oz	Quartile 1
2002	Up	$310/oz	96.0%	4.0%	$268/oz	$315/oz	Quartile 3
2003	Up	$363/oz	87.5%	12.5%	$325/oz	$370/oz	Quartile 3
2004	**Up**	**$409/oz**	**41.7%**	**58.3%**	**$385/oz**	**$452/oz**	**Quartile 1**
2005	Up	$444/oz	64.0%	36.0%	$380/oz	$463/oz	Quartile 3
2006	Up	$604/oz	96.0%	4.0%	$479/oz	$618/oz	Quartile 3
2007	Up	$695/oz	82.8%	17.2%	$580/oz	$755/oz	Quartile 3
2008	**Up**	**$872/oz**	**62.5%**	**37.5%**	**$750/oz**	**$1'050/oz**	**Average**
2009	Up	$972/oz	91.7%	8.3%	$675/oz	$988/oz	Quartile 3
2010	Up	$1'225/oz	69.2%	30.8%	$1'000/oz	$1'388/oz	Quartile 3
2011	Up	$1'572/oz	100.0%	0.0%	$1'325/oz	$1'550/oz	Quartile 3
2012	**Up**	**$1'669/oz**	**19.2%**	**80.8%**	**$1'525/oz**	**$2'050/oz**	**Quartile 1**
2013	Down	$1'411/oz	0.0%	100.0%	$1'600/oz	$1'900/oz	Quartile 1
2014	**Down**	**$1'266/oz**	**75.0%**	**25.0%**	**$1'080/oz**	**$1'315/oz**	**Quartile 3**
2015	Down	$1'160/oz	12.9%	87.1%	$950/oz	$1'321/oz	Quartile 1
2016	Up	$1'251/oz	100.0%	0.0%	$960/oz	$1'225/oz	Quartile 3
2017	**Neutral**	**$1'257/oz**	**43.5%**	**56.5%**	**$1'110/oz**	**$1'350/oz**	**Quartile 3**
2018	Neutral	$1'268/oz	12.5%	87.5%	$1'215/oz	$1'381/oz	Quartile 1
2019	Up	$1'393/oz	100.0%	0.0%	$1'242/oz	$1'365/oz	Quartile 4
2020	Down	$1'770/oz	100.0%	0.0%	$1'398/oz	$1'755/oz	Quartile 4
Average			**63.3%**	**36.7%**			

The distribution of the LBMA-PMS forecasts was skewed below the LBMA annual average gold price, with a notable tendency for the vast majority of the LBMA-PMS forecasters to share similar expectations about the anticipated LBMA average yearly gold price in the forthcoming year. The conformity in gold price forecasts was

also reflected in the high standard deviation of 34%, showing a high degree of variability. The forecasting spread below and above the LBMA annual average gold price in 2004, 2008, and 2017 was less than the overall twenty-one 34% standard deviation. In these three years, the median or average was among the more accurate aggregated consensus estimates, reflecting the expectation of balanced forecasting distribution errors.

Table 4-16 LBMA-PMS Gold Aggregated Consensus Estimates Selection Errors

Year	Actual annual gold price change	LBMA-PMS Forecasted gold price change	Expected most accurate aggregated consensus estimate	Actual most correct aggregated consensus estimate	LBMA-PMS participant relative ascending rank and percent
2004	Up (11.2%)	Up (100%)	Q3 (4.7%)	Q1 (1.5%)	9 (38%)
2008	Up (20.3%)	Up (100%)	Q3 (2.4%)	Average (1.1%)	16 (67%)
2012	Up (5.8%)	Up (96.2%)	Q3 (9.8%)	Q1 (3.5%)	5 (19%)
2014	Down (11.4%)	Down (100%)	Q1 (7.2%)	Q3 (0.1%)	21 (75%)
2017	Neutral (0.5%)	Neutral (48%/52%)	Q1 (3.8%)	Median (0.2%)	10 (43%)

For all five years, when the expected aggregated consensus estimate did not conform to the indicated selection criteria, the LBMA-PMS participants correctly predicted the market gold price direction. In 2004 relative to the average market gold price, 14 of the 24 LBMA-PMS participants made forecasts skewed to the upside. The successive year-on-year gold price increases since 2001 perhaps led to the overly optimistic expectations in 2004, resulting in the lower quartile being the most correct aggregated consensus estimate. The winning LBMA-PMS participant's ranking was in the second quartile, with 15 instead of the expected 6 participants making higher gold price forecasts.

Relative to the average market gold price in 2008, of the LBMA-PMS forecasts, 15 were lower, and 9 were higher. The margin of error between the consensus average and the upper quartile as the most accurate aggregated consensus estimate was $11/oz or 1.3%. The impact of the Global Financial Crisis could have contributed to the lower-than-expected gold price in 2008, which dropped in the second half relative to the first half. The winning LBMA-PMS participant's ranking was in the third quartile, with 15 rather than the expected 18 participants making lower gold price forecasts. In 2012, relative to the average market gold price, the LBMA-PMS forecasts were skewed to the upside, with 21 higher and 5 lower. Following the significant increase over the preceding nine years, the expectation was maybe for a similar increase in 2012, with

the gold price rising, falling, rising, and falling progressively in a volatile year. The winning LBMA-PMS participant's ranking was in the first quartile, with 21 instead of the expected 6 participants making higher gold price forecasts.

The LBMA-PMS forecasts in 2014 relative to the average market gold price were skewed to the downside, with 21 lower and 7 higher. Following the significant gold price decline in 2013 (21.4%), the expectation based on the LBMA-PMS forecasts appeared they expected a similar fall in 2014. The winning LBMA-PMS participant's ranking was in the third quartile, with 20 instead of the expected 7 participants making lower gold price forecasts. The LBMA-PMS forecasts in 2017 relative to the actual market gold price were skewed to the upside, with 13 higher and 10 lower. The winning LBMA-PMS participant's ranking was in the second quartile, with 9 instead of the expected 6 participants making lower gold price forecasts.

The LBMA-PMS aggregated consensus estimates were analyzed for twenty-one years to test the reliability of applying the two indicative rules for average annual market gold price increasing by more than 1% or declining by more than 1%. The lower/upper quartile selection criteria yielded a smaller APE than the aggregated consensus average or median in sixteen years of the twenty-one years. The lower quartile yielded the lowest APE for two (2004; 2012) of the other five years rather than the expected upper quartile. In one (2014) instance of the other five years, the upper quartile had the lowest APE instead of the expected lower quartile.

The RMSE was computed for each aggregated consensus estimate, showing the average forecasting error for each alternative reported (*Table 4-13*; Row c). The computed RMSE was used for the numerator to calculate the Theil U(II) Index. The RMSE numerator was divided by the RMSE based on the actual market gold price less the aggregated consensus median (*Table 4-13*; Row d). The RMSE numerator was divided by the RMSE based on the actual market gold price less than the prior year's value (*Table 4-13*; Row e). The RelMAE based on the actual consensus average, median, lower quartile, upper quartile, and the rule-based lower and upper quartile is reported (*Table 4-13*; Row f). *Table 4-17 LBMA-PMS Gold Aggregated Consensus Estimates Forecasting Errors* tabulates the aggregated APE, RMSE, Theil U(II) Index, and RelMAE for the alternative aggregated consensus estimates for the twenty-one years as reported (*Table 4-13*; Columns 3-6, Row f).

Table 4-17 LBMA-PMS Gold Aggregated Consensus Estimates Forecasting Errors

Aggregated consensus estimate	Absolute Percentage Error (APE)	Root Mean Square Error (RMSE)	Theil U(II) Median	Theil U(II) Prior year	ReLMAE prior year
Average	7.2%	106	101.4%	66.2%	82.6%
Median	7.1%	105	100.0%	65.3%	82.3%
Lower quartile (25%)	8.3%	113	107.9%	70.5%	104.4%
Upper quartile (75%)	6.8%	107	102.2%	66.7%	65.9%
Lower & Upper quartiles – LBMA forecasted direction[1]	7.4%	113	107.9%	70.4%	63.8%
Lower & Upper quartiles – Actual market direction[2]	6.2%	95	90.3%	59.0%	59.6%

1. *Based on LBMA-PMS forecasted direction, select the upper quartile for rising gold markets and the lower quartile for neutral or falling gold markets.*
2. *Based on the actual direction, the upper quartile is for rising gold markets, and the lower quartile is for neutral or falling gold markets.*

The APE difference between the average (7.2%) and the median (7.1%) is insignificant. The forecasting accuracy of the upper quartile (6.8%) is better than the average or median, reflecting the underestimation of the market gold price increase over the twenty-one years that included fourteen years of rising gold prices. Based on the actual gold market price direction, the most accurate forecast estimate (6.2%) would be the rule-based strategy of selecting either the lower or upper quartile based on an average annual market gold price of less than 1% or more than 1%, respectively. Looking at the reliability of the alternative aggregated consensus estimates and applying the rule-based selection based on the forecasted market direction yields a level of accuracy (7.4%), marginally worse than the aggregated consensus average (7.2%) or median (7.1%). The difference between the rule-based strategy depending on the actual market direction versus the LBMA-PMS forecasted market direction is a measure of the LBMA-PMS forecast direction error (1.2%). The RMSE used to calculate the Theil U(II) Index mirrored the APE results and indicated the expected reliability of aggregated consensus estimates.

The two aggregated consensus estimates based on the degree of centrality of the forecasts, the median, and the average yielded similar results across all the forecasting error calculations. The advantage of the median and the average measures is that the offsetting of forecasting errors above and below the central point helped compensate for market gold price directional forecasting errors. Using the consensus median as the

comparative benchmark in the Theil U(II) Index calculation highlighted how the alternative of the lower or upper quartile choice was more reliable. In only seventeen of the twenty-one years, the LBMA-PMS participants correctly forecasted the future market gold price direction, influencing the evaluation of the lower and upper quartiles selection rule dependent on the actual market gold price direction. Using the LBMA-PMS forecasted market gold price directions for the twenty-one years resulted in a less reliable outcome than the consensus median or average, as measured by the RMSE and reflected in the Theil U(II) Indexes. When using the actual market gold price change for selecting the lower or upper quartile, their reliability outperforms the consensus median or average based on the RMSE and Theil U(II) Index values. Using the actual market gold price change eliminates the LBMA-PMS participants' directional forecasting errors. The extent of the market gold price directional error can be measured as the difference in the RMSE error for the two lower and upper quartile selection options. The results from the RelMAE calculations confirm the Theil U(II) Indexes results.

A factor in evaluating the reliability of aggregated consensus estimates was whether any individual LBMA-PMS forecaster could consistently be more reliable with a lower overall APE across multiple years. *Table 4-18 LBMA-PMS Gold Analysts' Forecasting* tabulates the annual LBMA-PMS winning forecasters' performance for all the years they took part, compared to the comparative aggregated consensus estimates.

Table 4-18 LBMA-PMS Gold Analysts' Forecasting Reliability

LBMA-PMS Participant	No. of years ranked most accurate[1]	No. of years part.	Average APE for all years part.	Median APE	Theil U(II) – Median	ReLMAE – prior year	Mann-Kendall statistic ($Z_{crit95\%}$ 1.65)[2]
Aubertin, Philip	1 (100%)	1	0.0%	2.1%	1.9%	0.2%	N/A
Doshi, Aakash	1 (100%)	1	0.1%	4.1%	2.9%	11.8%	N/A
Holmes, David	1 (33%)	3	6.1%	6.5%	103.0%	58.1%	N/A
Levine, Howard	1 (25%)	4	3.7%	4.7%	84.0%	47.8%	N/A
Dahdah, Bernard	1 (17%)	6	10.9%	6.7%	179.9%	172.2%	-0.38
Tully, Edel	1 (17%)	6	8.7%	5.5%	226.1%	46.9%	0.00
Teves, Joni	1 (14%)	7	9.6%	9.3%	122.2%	124.1%	0.87
Melek, Bart	1 (13%)	8	10.5%	8.6%	127.7%	141.3%	0.75
Reade, John	1 (11%)	9	9.9%	5.6%	240.5%	73.7%	-0.73
Hochreiter, Rene	3 (20%)	15	8.6%	7.1%	107.4%	73.5%	0.54
Turner, Matthew	1 (6%)	16	6.2%	5.2%	109.3%	72.6%	0.30

LBMA-PMS Participant	No. of years ranked most accurate[1]	No. of years part.	Average APE for all years part.	Median APE	Theil U(II) – Median	ReLMAE – prior year	Mann-Kendall statistic ($Z_{crit95\%}$ 1.65)[2]
Norman, Ross	**3 (16%)**	**19**	**6.1%**	**7.2%**	**91.2%**	**48.0%**	**1.00**
Klapwijk, Philip	**3 (14%)**	**21**	**7.8%**	**7.1%**	**120.5%**	**69.8%**	**-1.99**
Panizzutti, Frederic	**2 (10%)**	**21**	**7.0%**	**7.1%**	**97.8%**	**68.6%**	**-0.60**
Average	0 (0%)	21	7.2%	7.1%	101.4%	82.6%	-0.63
Median	0 (0%)	21	7.1%	7.1%	100.0%	82.3%	-0.82
Lower Quartile	0 (0%)	21	8.3%	7.1%	107.9%	104.4%	-0.69
Upper Quartile	1 (5%)	21	6.8%	7.1%	102.2%	65.9%	-0.45
Q1/Q3 – LBMA[3]	0 (0%)	21	7.4%	7.1%	107.9%	90.3%	-1.24
Q1/Q3 – Actual[4]	**0 (0%)**	**21**	**6.2%**	**7.1%**	**63.8%**	**59.6%**	**-1.48**

1. *The number of years ranked first in LBMA-PMS and the percent of first compared to all years participated.*
2. *Based on the APE, the z-Critical value for a 95% confidence interval for a one-sided test is 1.65. The z-Critical value for a 90% confidence interval for a one-sided test is 1.28.*
3. *Lower and upper quartile selection based on LBMA-PMS participants' forecasted direction.*
4. *Lower and upper quartile selection is based on the actual market gold price direction.*

For the top LBMA-PMS participants tabulated in *Table 4-18*, 11 of the 14 had a ReLMAE lower than 100%, indicating they were more reliable than the prior year's average for predicting the following year's gold price. The same result applies to aggregated consensus estimates other than the lower quartile (104.4%) and selection rule based on the actual market price direction of 59.6%. For the two participants who have taken part in all twenty-one years of the LBMA-PMS analyzed, and who previously ranked as the most accurate forecaster in at least one year, Klapwijk (3) and Panizzutti (2), their respective APE of 7.8% and 7.0% were comparable to the average (7.2%) and median (7.1%) APE. The most accurate participant in the annual LBMA-PMS over the twenty-one years analyzed was Norman (3), with an APE of 6.1%, RelMAE of 48.0%, and Theil U(II) Index value of 91.2% for the nineteen years he took part. The Theil U(II) Index scores below 100% for Panizzutti (97.8%) and Norman (91.2%) reflect the same result as for the APE, as the only two top participants who had an overall performance better than the aggregated consensus median. The lower and upper quartile selection rule based on the actual market price direction yielded the lowest Theil U(II) Index value (63.8%). It provided an ideal benchmark for the most reliable forecasting aggregated consensus estimate.

The Mann-Kendall test was used to test if there was any trend in the reliability of the LBMA-PMS forecasters' APE values and the aggregated consensus estimates

(Mann, 1945; Gilbert, 1987). Because the APE ignores the directional error of the forecasting error, a one-sided confidence value of 1.65 was used based on a 95% confidence interval. The comparable confidence value for a 90% confidence interval is 1.28. Only 1 (Klapwijk) of the top 14 LBMA-PMS forecasters for the twenty-one years analyzed had a Mann-Kendall statistic greater than 1.65 z-Critical value. Klapwijk's Mann-Kendall statistic of -1.99 indicated a decreasing accuracy trend over the twenty-one years he participated. At the 90% confidence interval, the lower and upper quartile selection rule based on the actual market price direction had a Mann-Kendall statistic value of -1.48, greater than the 1.28 z-Critical value, indicating a significant decreasing trend over the twenty-one years.

Table 4-19 LBMA-PMS Consensus Estimates and Top Analysts Reliability, summarizes the reliability of the individual LBMA-PMS participants relative to the aggregated consensus estimates. Of the top 3 participants, only Norman (6.1%) achieved a lower annual average APE than the average yearly aggregated consensus median (6.2%). Over the 21 years, Klapwijk and Panizzutti participated in the LBMA-PMS. Each achieved a lower APE than the aggregated consensus median in 12 (57.1%) of the 21 years. Norman achieved a lower APE than the aggregated consensus median in 15 (78.9%) of the 19 years he participated in the LBMA-PMS. The selection rule for the Lower/Upper quartile (*Table 4-19*, Rule Q1/Q3) based on the actual LBMA (2020) annual average gold direction, achieved an APE of 6.3% and a lower APE than the aggregated consensus median in 17 (81%) of the 21 years. An indicative measure of the performance of the top 3 LBMA-PMS gold analysts and the RuleQ1/Q3 is the average rank over the years participated. The average number of participants over the twenty-one years analyzed was 26. As expected, the aggregated consensus median achieved an average rank of 11, comparable to the midpoint value of 13. Klapwijk and Panizzutti similarly achieved average ranks close to the midpoint value. Norman achieved an average rank of 9, representing a binomial probability of 11.5% based on the 19 years he participated. The Rule Q1/Q3 average rank was 8 and represented a binomial probability of 5.4% occurring over the 21 years.

Table 4-19 LBMA-PMS Consensus Estimates and Top Analysts Reliability

Year / No. participants	Lowest APE (%) / Market change (%) / Highest APE (%)	APE measure	APE (%)	APE (%) Rank	Rank z-Score
2000	Lowest	Median	5.7%	11	0.28
	10.4%	Rule Q1/Q3	2.1%	4	1.25
Market	**0.2%**	Norman			
Participants	Highest	**Klapwijk**	**0.3%**	**1**	1.66
25	57.7%	Panizzutti	8.2%	16	-0.42
2001	Lowest	Median	1.5%	6	0.94
	7.8%	Rule Q1/Q3	1.1%	4	1.23
Market	**-3.0%**	Norman			
Participants	Highest	**Klapwijk**	**0.0%**	**1**	1.66
24	9.6%	Panizzutti	4.0%	16	-0.51
2002	Lowest	Median	6.4%	10	0.42
	13.5%	Rule Q1/Q3	5.6%	7	0.83
Market	**12.5%**	**Norman**	**1.5%**	**1**	1.66
Participants	Highest	Klapwijk	5.6%	7	0.83
25	1.7%	Panizzutti	4.7%	5	1.11
2003	Lowest	Median	5.0%	11	0.22
	10.6%	Rule Q1/Q3	3.5%	6	0.94
Market	**14.8%**	**Norman**	**1.3%**	**2**	1.52
Participants	Highest	Klapwijk	1.7%	3	1.37
24	1.7%	Panizzutti	1.7%	4	1.23
2004	Lowest	**Median**	**2.4%**	**8**	0.65
	5.9%	Rule Q1/Q3	4.7%	17	-0.65
Market	**11.2%**	Norman	10.5%	24	-1.66
Participants	Highest	Klapwijk	9.4%	22	-1.37
24	10.5%	Panizzutti	7.5%	20	-1.08
2005	Lowest	Median	1.5%	8	0.69
	14.5%	Rule Q1/Q3	1.3%	6	0.97
Market	**7.9%**	**Norman**	**0.6%**	**3**	1.39
Participants	Highest	Klapwijk	3.7%	16	-0.42
25	4.2%	Panizzutti	3.3%	13	0.00
2006	Lowest	Median	13.1%	13	0.00
	20.7%	Rule Q1/Q3	8.2%	7	0.83
Market	**26.4%**	**Norman**	**2.4%**	**1**	1.66
Participants	Highest	Klapwijk	11.9%	12	0.14
25	2.4%	Panizzutti	9.9%	9	0.55
2007	Lowest	Median	6.5%	13	0.24
	16.6%	Rule Q1/Q3	3.0%	7	0.96
Market	**13.2%**	**Norman**	**3.0%**	**6**	1.08
Participants	Highest	Klapwijk	3.1%	8	0.84
29	8.6%	Panizzutti	6.2%	12	0.36
2008	Lowest	Median	2.5%	5	1.08
	14.0%	Rule Q1/Q3	2.4%	4	1.23

Year / No. participants	Lowest APE (%) / Market change (%) / Highest APE (%)	APE measure	APE (%)	APE (%) Rank	Rank z-Score
Market	**20.3%**	Norman	11.9%	21	-1.23
Participants	Highest	Klapwijk	0.7%	2	1.52
24	20.4%	**Panizzutti**	**0.0%**	**1**	1.66
2009	Lowest	Median	7.4%	12	0.07
	30.6%	Rule Q1/Q3	2.7%	6	0.94
Market	**10.3%**	Norman	1.6%	4	1.23
Participants	Highest	**Klapwijk**	**0.2%**	**1**	1.66
24	1.6%	Panizzutti	7.3%	12	0.07
2010	Lowest	Median	2.1%	7	0.87
	18.3%	**Rule Q1/Q3**	**0.7%**	**2**	1.53
Market	**20.6%**	Norman	0.9%	5	1.13
Participants	Highest	Klapwijk	4.3%	14	-0.07
26	13.4%	Panizzutti	3.6%	10	0.47
2011	Lowest	Median	6.8%	12	0.07
	15.7%	Rule Q1/Q3	5.1%	6	0.94
Market	**22.1%**	**Norman**	**3.7%**	**2**	1.52
Participants	Highest	Klapwijk	6.0%	10	0.36
24	1.4%	Panizzutti	4.4%	3	1.37
2012	Lowest	Median	6.1%	12	0.20
	8.6%	Rule Q1/Q3	9.8%	19	-0.73
Market	**5.8%**	Norman	5.8%	11	0.33
Participants	Highest	**Klapwijk**	**5.5%**	**10**	0.47
26	22.8%	Panizzutti	8.3%	17	-0.47
2013	Lowest	Median	24.2%	12	0.00
	13.4%	**Rule Q1/Q3**	**21.4%**	**6**	0.90
Market	**-18.3%**	Norman	23.0%	8	0.60
Participants	Highest	Klapwijk	30.9%	19	-1.06
23	34.6%	Panizzutti	24.2%	12	0.00
2014	Lowest	Median	2.9%	12	0.31
	14.7%	Rule Q1/Q3	7.2%	21	-0.80
Market	**-11.4%**	Norman	0.6%	4	1.30
Participants	Highest	Klapwijk	7.1%	21	-0.80
28	3.8%	**Panizzutti**	**0.4%**	**1**	1.67
2015	Lowest	Median	6.0%	13	0.34
	18.1%	Rule Q1/Q3	2.4%	5	1.23
Market	**-9.2%**	Norman	13.9%	30	-1.57
Participants	Highest	**Klapwijk**	**1.1%**	**4**	1.34
31	13.9%	Panizzutti	11.4%	27	-1.23
2016	Lowest	Median	10.5%	16	0.00
	23.3%	**Rule Q1/Q3**	**8.5%**	**8**	0.89
Market	**7.3%**	Norman	11.2%	17	-0.11
Participants	Highest	Klapwijk	16.6%	27	-1.23

Year / No. participants	Lowest APE (%) / Market change (%) / Highest APE (%)	APE measure	APE (%)	APE (%) Rank	Rank z-Score
31	2.1%	Panizzutti	10.5%	16	0.00
2017	Lowest	**Median**	**0.2%**	**3**	1.36
	11.7%	Rule Q1/Q3	3.8%	15	-0.45
Market	**0.5%**	Norman	4.2%	17	-0.75
Participants	Highest	Klapwijk	2.5%	10	0.30
23	7.4%	Panizzutti	1.2%	5	1.06
2018	Lowest	Median	4.1%	11	0.22
	4.2%	**Rule Q1/Q3**	**1.5%**	**5**	1.08
Market	**0.9%**	Norman	7.1%	18	-0.79
Participants	Highest	Klapwijk	8.9%	24	-1.66
24	8.9%	Panizzutti	7.6%	21	-1.23
2019	Lowest	Median	5.6%	15	0.06
	10.8%	Rule Q1/Q3	4.8%	8	0.87
Market	**8.9%**	**Norman**	**4.0%**	**4**	1.33
Participants	Highest	Klapwijk	5.3%	13	0.29
30	2.0%	Panizzutti	4.1%	5	1.21
2020	Lowest	Median	11.9%	15	0.06
	21.0%	Rule Q1/Q3	10.0%	8	0.87
Market	**21.3%**	**Norman**	**0.8%**	**1**	1.68
Participants	Highest	Klapwijk	12.2%	18	-0.29
30	0.8%	Panizzutti	7.6%	3	1.44
26	**7.7%**	**Average**			
Average APE / Average Rank / z-Score <1.65> (CL 90%)		**Median**	**7.1%**	**11**	**0.28**
		Rule Q1/Q3	**6.2%**	**8**	**0.69**
		Norman	**6.1%**	**9**	**0.54**
		Klapwijk	**7.8%**	**12**	**0.14**
		Panizzutti	**7.0%**	**11**	**0.28**

An inference that could be drawn from Norman's relative APE performance was that he could be considered a more reliable forecaster than the aggregated consensus median. An assumption could similarly be drawn about the Rule Q1/Q3 aggregated consensus estimate. The average z-Scores for Norman's ranks over the 19 years he participated was 0.54, with 3 (15.8%) years ranking above the critical limit of 1.65 and 1 (5.3%) year below the critical limit of -1.65. The cumulative probability of 3 years' ranks exceeding the critical limit is 10.1%, remembering the offsetting 1 year below the critical limit. The same conclusion can be drawn for the other two top LBMA-PMS analysts, with Klapwijk having 3 z-Scores above 1.65, a cumulative probability of 9.3%, and 1 below -1.65. Panizzutti had 2 z-Scores above 1.65 with a cumulative probability of 6.9%.

Figure 4-4 LBMA-PMS Gold Top 3 Forecasters and Consensus Median shows the forecasted annual gold prices of the top 3 LBMA-PMS participants and the aggregated consensus median compared to the actual market gold price and annual ranks for all the LBMA-PMS participants each year. Given the low APE results achieved by the LBMA-PMS top 3 participants (Average 6.9%), their gold price forecasts tracked the actual market gold price closely, with a few exceptions, such as in 2012, 2013, and 2016. The annual variation in the rank of the top 3 LBMA-PMS forecasters and the aggregated consensus median illustrates the variation in their relative performance. When the LBMA-PMS forecasters and the aggregated consensus median are ranked according to their annual APE, their yearly average ranking tends to revert towards the midpoint rank of the average of all the LBMA-PMS participants.

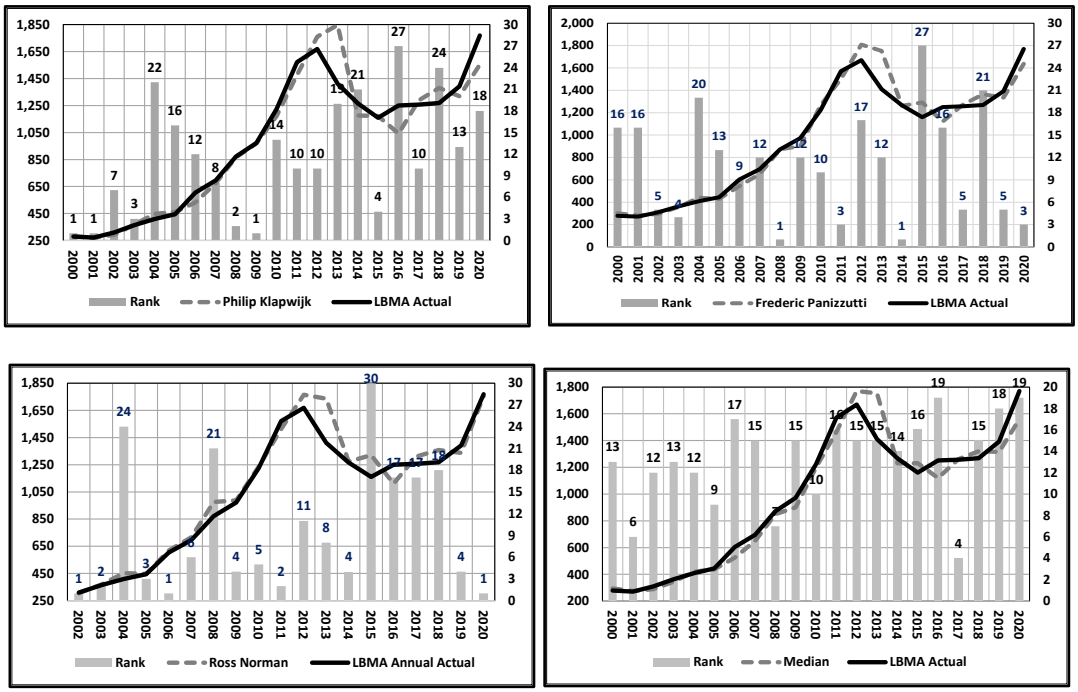


Figure 4-4 LBMA-PMS Gold Top 3 Forecasters and Consensus Median Reliability

For the Test-Retest correlation analysis of the APE of the winning LBMA-PMS forecasters and the aggregated consensus estimates, the correlation average was around 15% (Guttman, 1945; Weir, 2005). The Mann-Kendall statistic showed similar results, indicating no significant sign of a trend over successive years of forecasting, either improving or worsening (Mann, 1945; Gilbert, 1987). *Figure 4-5 LBMA-PMS Top 3 Analysts' Annual APE and the Consensus Median* graphs the annual variation in

the APE for the top 3 LBMA-PMS forecasters and the aggregated consensus median and illustrates the lack of clear trends.

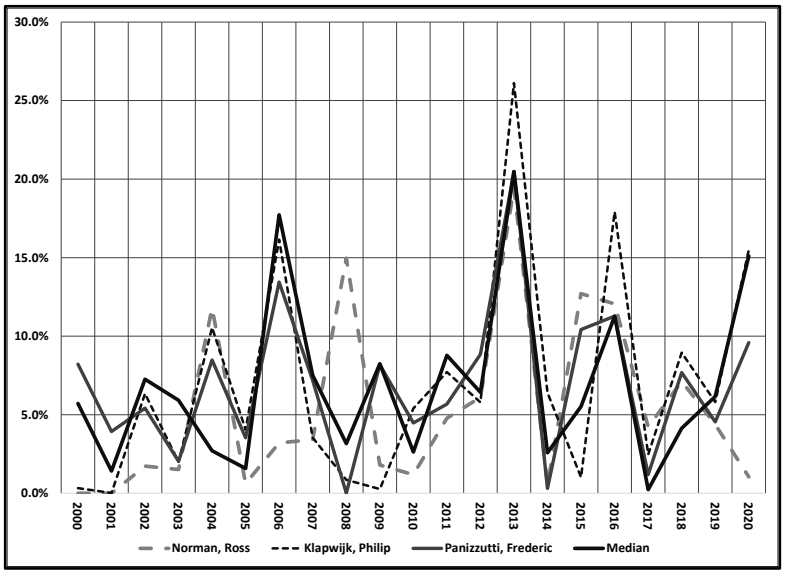


Figure 4-5 LBMA-PMS Top 3 Analysts' Annual APE and the Consensus Median

The performance of the top 3 LBMA-PMS forecasters in predicting the future market gold price direction was comparable to the aggregated consensus median of all the LBMA-PMS participants. The possibility of selecting the best forecast depends on correctly knowing the expected market gold price direction, regardless of the past performance of the LBMA-PMS forecaster, the aggregated consensus median, or applying a selection rule such as the Lower/Upper Quartile depending on the expected gold price direction.

4.7. **LBMA Precious Metals Survey Qualitative Results**

The annual LBMA-PMS requires the participants to explain the basis for their gold forecasts. Using the key factors considered significant for forecasting the future gold price by the World Gold Council (Gold, 2021), the explanations of the LBMA-PMS participants over the twenty-one years were standardized for terminology and analyzed for frequency of keywords and collocates using WordSmith Tools. A consideration in the qualitative analysis was if the qualitative data mirrored the reliability of the aggregated consensus estimates based on the quantitative data. The *Table 4-20* below tabulates the frequency of the keywords or concepts for the twenty-one years of LBMA-PMS linked to the World Gold Council (Gold, 2021) list. *Table 4-20 LBMA-PMS World Gold Council Keyword Frequency* lists the identified keywords' relative occurrence, which accounted for around 10% of the total word count.

Table 4-20 LBMA-PMS World Gold Council Keyword Frequency

World Gold Council keywords occurrences	Economy	Demand	Risk / Uncertainty	Opportunity cost	Trends / Supply
Economic	0.80%				
USA	0.59%				
World	0.35%				
Oil	0.10%				
Demand		0.62%			
Market		0.61%			
Metal		0.43%			
Central Banks		0.32%			
China		0.21%			
India		0.14%			
Asia		0.11%			
Production hedge		0.16%			
Asset		0.15%			
Jewellery		0.14%			
ETF		0.10%			
Reserve		0.09%			
US Dollar			0.78%		
Policies			0.31%		
Risk			0.24%		
Inflation			0.22%		
Currency			0.19%		
Geopolitical			0.14%		
EUR			0.11%		
Safe-haven			0.10%		
Environment			0.10%		
Political			0.09%		
Interest Rates				0.43%	
Investor				0.37%	
Investment				0.27%	
Federal Reserve Bank				0.26%	
Equity Exchange				0.17%	
Fund				0.09%	
Supply					0.19%
Producer					0.14%
Mine					0.13%
Production					0.08%
TOTAL	**1.84%**	**3.07%**	**2.29%**	**1.59%**	**0.54%**

The explanations provided by the LBMA-PMS participants focused on the buy-side of the gold market, economy, demand, risk and uncertainty, and opportunity cost, accounting for 94% of the keyword occurrences. The sell-side or supply of the gold market accounted for a far lesser 6%. For the buy-side of the gold market, the demand (35%) and risk and uncertainty (26%) were seen as the most significant factors, accounting for 61% of the total buy-side keywords.

Table 4-21 LBMA-PMS World Gold Council Keyword Annual Occurrences tabulates the occurrence of the World Gold Council (Gold, 2021) keywords annually for all the LBMA-PMS participants and the corresponding aggregated consensus median APE for the year. The correlation between the APE and the keyword occurrences, average word count, and the number of LBMA-PMS participants was calculated to determine if the explanations supplied informed the reliability of the aggregated consensus median. The years in which the LBMA-PMS participants incorrectly predicted the market gold price direction are highlighted in bold. In 2017, the split of LBMA-PMS participants expecting a gold price increase (11) versus a decrease (12) in the gold price was almost balanced.

Table 4-21 LBMA-PMS World Gold Council Keyword Annual Occurrences

World Gold Council keyword annual occurrences	Total keyword occurrences (%)	LBMA-PMS average word count	LBMA-PMS No. of participants	LBMA-PMS Median APE
2000	**6.7%**	**163**	**25**	**5.7%**
2001	9.2%	185	24	1.5%
2002	8.8%	165	25	6.4%
2003	8.5%	175	24	5.0%
2004	8.9%	186	24	2.4%
2005	9.6%	137	25	1.5%
2006	8.5%	174	25	13.1%
2007	9.2%	178	29	6.5%
2008	10.1%	195	24	2.5%
2009	9.3%	179	24	7.4%
2010	9.1%	185	26	2.1%
2011	9.7%	184	24	6.8%
2012	8.1%	172	26	6.1%
2013	**8.8%**	**185**	**23**	**24.2%**
2014	9.7%	157	28	2.9%
2015	10.0%	173	31	6.0%
2016	**11.0%**	**168**	**31**	**10.5%**
2017	**9.9%**	**180**	**23**	**0.2%**

World Gold Council keyword annual occurrences	Total keyword occurrences (%)	LBMA-PMS average word count	LBMA-PMS No. of participants	LBMA-PMS Median APE
2018	**10.7%**	**140**	**24**	**4.1%**
2019	10.5%	142	30	5.6%
2020	9.6%	162	30	11.9%
Average	**9.3%**	**171**	**26**	**7.1%**
Correlation	**-12.4%**	**11.2%**	**8.7%**	**100.0%**

The correlation between the aggregated consensus median APE and the keyword occurrences, the average explanations word count, and the number of annual LBMA-PMS participants was low, indicating no significant relationship between these factors. In 2017 when the number of LBMA-PMS participants was equally divided in the expected gold price direction, the aggregated consensus median was the most reliable. The keyword occurrences percentage (9.9%) was comparable to the twenty-one-year average (9.3%). The 2017 forecasting spread was the closest to the general expectation of almost equivalent forecasting errors above and below the central measure of the group. Still, from a qualitative perspective, the same trend was not discernible.

An explanation for the lack of a correlation between the keywords and the aggregated consensus median was that the possible diversity of occurrences across the keywords offset each other, leading to a lower level of correlation with the aggregated consensus median APE. *Table 4-22 LBMA-PMS World Gold Council Keyword Correlations* tabulates the correlation between the World Gold Council (Gold, 2021) keyword groupings, word count, number of participants, and aggregated consensus median APE.

Table 4-22 LBMA-PMS World Gold Council Keyword Correlations

World Gold Council keyword correlation matrix	Economy	Demand	Risk / Uncertainty	Opportunity cost	Trends / Supply	Total keyword occurrences	Average word count	No. Participants	Median APE
Economy		(70.5%)	53.5%	63.6%	(50.5%)	69.9%	(18.8%)	42.0%	15.0%
Demand	(70.5%)		(51.5%)	(35.4%)	31.1%	(30.9%)	24.4%	(12.6%)	1.5%
Risk / Uncertainty	53.5%	(51.5%)		47.7%	(48.9%)	74.8%	(18.1%)	7.1%	(23.7%)
Opportunity cost	63.6%	(35.4%)	47.7%		(56.6%)	81.7%	(39.4%)	44.8%	4.2%
Trends / Supply	(50.5%)	31.1%	(48.9%)	(56.6%)		(38.3%)	23.7%	(24.1%)	(24.8%)
Total keyword occurrences	69.9%	(30.9%)	74.8%	81.7%	(38.3%)		(23.2%)	36.8%	(12.4%)

World Gold Council keyword correlation matrix	Economy	Demand	Risk / Uncertainty	Opportunity cost	Trends / Supply	Total keyword occurrences	Average word count	No. Participants	Median APE
Average word count	(18.8%)	24.4%	(18.1%)	(39.4%)	23.7%	(23.2%)		(32.8%)	11.2%
No. Participants	42.0%	(12.6%)	7.1%	44.8%	(24.1%)	36.8%	(32.8%)		8.7%
Median APE	15.0%	1.5%	(23.7%)	4.2%	(24.8%)	(12.4%)	11.2%	8.7%	

The World Gold Council (Gold, 2021) keyword groups and the aggregated consensus median APE were negatively and positively correlated, albeit not strongly, perhaps accounting for the low overall correlation observed. The buy-side factors of economy, risk, uncertainty, and opportunity cost showed strong positive cross-correlations mirrored in the correlations with the total occurrences, indicating a common association between the LBMA-PMS participants in their expectations about these factors. The demand and supply factors displayed a low correlation, a negative correlation with the economy, risk and uncertainty, and opportunity cost. The correlation between the number of LBMA-PMS participants and economy and opportunity cost was positive and more significant than the negative correlation between demand and supply. The implication is that increasing the number of participants resulted in more of the same explanations rather than a greater diversity of opinions. For the correlations of the keyword groups and the average word count, the higher average word counts were associated with lesser occurrences of economy, risk and uncertainty, opportunity cost, and demand and supply. The combination of a negative correlation between the number of participants and demand and supply and a positive correlation of the two factors with average word count would appear to show as participation drops off, the overall diversity of opinions becomes more balanced.

The low correlation between the World Gold Council (Gold, 2021) keyword groups and the aggregated consensus median APE was unexpected. The combination of negative and positive correlations for some of the World Gold Council (Gold, 2021) keyword groupings may account for the low overall correlations. Another explanation was that the justifications supplied were intended to reflect the uncertainty around the judgmental gold price forecasts made, aiming to explain their basis and perhaps the unknown factors that could cause the estimates to be wrong. An implication from the qualitative analysis questions the assertion that requiring an explanation for a judgmental forecast improves its reliability. It could also be that the LBMA-PMS

forecasters have an intuitive sense of the value of the judgmental estimate they are making, and the justifications were more of a practiced response addressing the typical issues.

The LBMA-PMS explanations were analyzed for neighboring collocates, twelve words on either side of the keyword, using WordSmith Tools. *Table 4-23 LBMA-PMS World Gold Council Main Keyword Collocates* summarizes the ten main keywords that accounted for approximately 30% of the collocate occurrences between the keywords listed and around 60% of the total occurrence for the identified keywords.

Table 4-23 LBMA-PMS World Gold Council Main Keyword Collocates

Keyword Collocates	Economic	USA	Demand	Market	US Dollar	Inflation	Interest Rates	Investment	Federal Reserve	% of Total
Economic		**1.44%**	0.27%	0.32%	0.54%	0.36%	0.39%	0.10%	0.22%	**3.52%**
USA	**1.43%**		0.24%	0.29%	**0.96%**	0.45%	**1.40%**	0.08%	**0.71%**	**4.36%**
Demand	0.30%	0.23%		0.61%	0.25%	0.16%	0.15%	**1.11%**	0.08%	**3.16%**
Market	0.32%	0.30%	0.60%		0.44%	0.14%	0.37%	0.35%	0.21%	**2.74%**
US Dollar	0.60%	**0.93%**	0.28%	0.43%		0.29%	**0.83%**	0.12%	0.30%	**3.21%**
Inflation	0.35%	0.44%	0.18%	0.16%	0.24%		0.44%	0.09%	0.15%	**1.83%**
Interest Rates	0.38%	**1.36%**	0.12%	0.35%	**0.82%**	0.46%		0.07%	**0.98%**	**4.66%**
Investment	0.13%	0.09%	**1.09%**	0.38%	0.13%	0.10%	0.07%		0.04%	**2.13%**
Federal Reserve	0.21%	0.68%	0.09%	0.20%	0.34%	0.16%	**1.13%**	0.05%		**1.73%**
% of Total	**3.70%**	**5.48%**	**2.87%**	**2.74%**	**3.71%**	**2.12%**	**4.78%**	**1.97%**	**2.70%**	**27.35%**

The most significant associations centered around the economy, the United States, and demand. The occurrences of the collocates focused on the United States economy, dollar, interest rates, and Federal Reserve Bank. For demand, the main cross collocation occurred with investment. The correlation between the aggregated consensus median APE and the nine keywords in *Table 4-23* was 18.3%, compared to 12.4% for all thirty-six keywords from the World Gold Council (Gold, 2021) list. It may improve the reliability of the aggregated gold price forecasting if participants are asked to focus on the fewer but more significant factors or even make it easier and ask the forecasters to assign a weighting.

4.8. Research Findings

The cooperative inquiry involved the collection of 190 metal price forecasts from 23 participants over seven periods on the CM panel website. In addition, 545 gold price forecasts made by 127 participants in the LBMA-PMS over twenty-one years were sourced to triangulate with the CM panel results. The data from both forecasting forums were analyzed to evaluate the possibility of finding a reliable aggregated consensus estimate that could be used in the workplace in the future. The quantitative

analysis yielded an alternative perspective on aggregated consensus estimates that warrants further consideration in the workplace of natural resource organizations. The qualitative analysis produced less definitive results, suggesting that the metal price forecasting participants' explanations of their judgmental forecasts were not fully informed by their numerical estimates. From a cooperative inquiry perspective of the researcher, orchestrating the collection and making metal price forecasts allowed insights to be gained that could be applied in considering a metal price forecasting approach in the workplace, structured by considering some of the insights gained from the AR project.

The AR project questioned three aspects of the possibility of using an aggregated consensus approach for making metal price forecasts for use in the workplace by natural resource organizations.

Hypothesis 1

The aggregated consensus median will be the most reliable metal price forecast over multiple periods, capturing the diversity of estimates around the statistical midpoint.

Hypothesis 1 Conclusion

For the CM panel copper forecasts, the lowest APE (8.9%) over the seven forecasting periods and 107 forecasts was the upper quartile (*Table 4-2 CM panel Forecasts - Copper Consensus Estimates Summary*). The comparable aggregated consensus average and median APE were 14.5% and 16.1%. The CM panel copper Theil U(II) Index for the upper quartile, compared to the median (100%), was 61.3%. The z-Score based on the average ranking for the aggregated consensus median and upper quartile were below the critical threshold of 1.4 at a 90% confidence limit. The inference that can be drawn from the CM panel copper data, the aggregated consensus median is not the most reliable forecast. Despite having the lowest APE, the z-Score for the aggregated consensus upper quartile would infer it is not a reliable forecast. If the Theil U(II) Index is used to measure forecasting reliability, the aggregated consensus upper quartile performed better than the aggregated consensus median in forecasting the copper price in a rising metal price environment.

For the CM panel gold forecasts, the lowest APE (3.2%) over the seven forecasting periods and 84 forecasts was the lower quartile (*Table 4-5 CM panel Forecasts - Gold*

Consensus Estimates Summary). The comparable aggregated consensus average and median APE were 6.9% and 6.3%. The CM panel gold Theil U(II) Index for the lower quartile, compared to the median (100%), was 62.6%. The z-Score based on the average ranking for the aggregated consensus median and lower quartile was below the critical threshold of 1.4 at a 90% confidence limit. The inference that can be drawn from the CM panel gold data, the aggregated consensus median is not the most reliable forecast. Despite having the lowest APE, the z-Score for the aggregated consensus lower quartile would infer it is not a reliable forecast. If the Theil U(II) Index is used to measure forecasting reliability, the aggregated consensus lower quartile performed better than the aggregated consensus median in forecasting the gold price in a declining metal price environment.

For the LBMA-PMS, the consensus average (7.2%) and median (7.1%) of the twenty-one years and 545 forecasts only accounted for one year with the lowest APE. The lower quartile (8.3%) had the lowest APE for seven of the twenty-one years, and the upper quartile (6.8%) had the lowest APE for the remaining twelve years (*Table 4-13 LBMA-PMS Gold Forecasts - Consensus Estimates Summary*). The LBMA-PMS Theil U(II) Index for the election rule Q1/Q3 aggregated consensus lower/upper quartile, compared to the median (100%), was 90.3%. The z-Score based on the average ranking for the aggregated consensus median was similarly below the critical threshold of 1.65 at the 90% critical limit. The inference that can be drawn from the LBMA-PMS data, the aggregated consensus median is not the most reliable forecast. Despite having a low APE, the z-Score for the selection rule Q1/Q3 aggregated consensus lower/upper quartile would infer it is not a reliable forecast. Suppose the Theil U(II) Index is used as a measure of forecasting reliability. In that case, the selection rule Q1/Q3 aggregated consensus lower/upper quartile potentially provides a more reliable forecast than the aggregated consensus median, subject to correctly anticipating the future market metal direction.

The CM panel data reject the hypothesis that the aggregated consensus median is reliable for predicting metal prices. The LBMA-PMS data reinforces the CM panel findings that the aggregated consensus median is not a reliable forecast of metal prices. The analysis of the CM panel and the LBMA-PMS data, based on the achieved APE and Theil U(II) Index values, point to the possibility of using one of the other aggregated consensus estimates as a more reliable estimate of future metal prices. The

lower and upper quartile aggregated consensus estimates offer a workable alternative to the aggregated consensus average or median. The reliability of the lower and upper quartile is subject to the proviso of correctly predicting the future market metal price direction. The CM panel and LBMA-PMS data analysis did not support the possibility of an individual expert consistently proving reliable in forecasting metal prices. Hypothesis 1, advocating the consensus median as a reliable forecaster of metal prices, is rejected.

Null Hypothesis 1 - Alternative 1

An individual expert can reliably forecast metal prices over multiple periods more than the aggregated consensus median.

Null Hypothesis 1 – Alternative 1 Conclusion

Among the CM panel copper participants who submitted forecasts in all seven forecasting periods, 2 achieved a lower APE than the aggregated consensus average or median, and 2 performed worse (*Table 4-3 CM panel - Copper Consensus Estimates Reliability*). None of the 4 CM panel copper participants achieved a lower APE than the aggregated consensus upper quartile. The CM panel copper Theil U(II) Index for the upper quartile (61.3%), compared to the median (100%), which was lower than any of the 4 CM panel copper participants who submitted forecasts in all seven forecasting periods. The z-Score based on the average ranking for the 4 top participating CM panel copper participants was below the critical threshold of 1.4 at the 90% critical limit. However, around half of the CM panel copper participants achieved a lower APE than the aggregated consensus median. However, none could be considered a reliable forecaster based on their ranking performance. If the Theil U(II) Index is used to measure forecasting reliability, 2 of the CM panel copper participants performed better than the aggregated consensus median but not the aggregated consensus upper quartile.

Among the CM panel gold participants who submitted forecasts in all seven forecasting periods, 3 achieved a lower APE than the aggregated consensus average or median, and 1 performed worse (*Table 4-6 CM panel – Gold Consensus Estimates Reliability*). Of the 4 CM panel gold participants, only 1 achieved an APE marginally lower than the aggregated consensus lower quartile. The CM panel gold Theil U(II) Index for the lower quartile (62.6%), compared to the median (100%), which was

lower than 3 of the 4 CM panel gold participants who submitted forecasts in all seven forecasting periods. The z-Score based on the average ranking for the 4 top participating CM panel gold participants was below the critical threshold of 1.4 at the 90% critical limit. However, around half of the CM panel gold participants achieved a lower APE than the aggregated consensus median. However, none could be considered a reliable forecaster based on their ranking performance. Despite having the lowest APE, the z-Score for the aggregated consensus lower quartile would infer it is not a reliable forecast. If the Theil U(II) Index is used to measure forecasting reliability, 3 of the CM panel gold participants performed better than the aggregated consensus median and one better than the aggregated consensus lower quartile.

For the top 3 LBMA-PMS analysts who had submitted forecasts in nineteen or more years, 1 achieved a lower APE than the aggregated consensus median, and 2 performed worse (*Table 4-18 LBMA-PMS Gold Analysts' Forecasting*). Of the 3 LBMA-PMS analysts, 1 achieved an APE (6.1%) marginally lower than the selection rule Q1/Q3 aggregated consensus lower/upper quartile (6.2%). The LBMA-PMS Theil U(II) Index for the best-performing analyst was 91.2%, compared to the median (100%), and was comparable to selection rule Q1/Q3 aggregated consensus lower/upper quartile 90.3%. The z-Score based on the average ranking for the 3 top performing LBMA-PMS analysts was below the critical threshold of 1.65 at the 90% critical limit. Based on their ranking performance, the inference that can be drawn from the LBMA-PMS data is that none of the top 3 LBMA-PMS analysts could consistently be considered a reliable forecaster. If the Theil U(II) Index is used to measure forecasting reliability, the top LBMA-PMS analyst (Norman) performed better than the aggregated consensus median and similar to the selection rule Q1/Q3 aggregated consensus lower/upper quartile.

Over the nineteen years, Norman's best APE was 0.6% in 2014, and his worst APE was 23.0% in 2013, with a standard deviation of 5.8% for all nineteen years. The comparable relative APE performance of lower or upper quartile rule-based strategy over the same 19 years of the LBMA-PMS was a minimum APE of 0.7% in 2010 and a maximum APE of 21.4% in 2013, with a standard deviation over the nineteen years of 4.6%. The Test-Retest correlations for Norman and the lower or upper quartile rule-based strategy were around negative 17% and positive 24%, respectively, showing low consistency levels. The positive lower or upper-quartile rule-based strategy correlation

implied greater consistency in successive years' performances. The Mann-Kendall test was used to test if there was any trend in the reliability of Norman's APE values over the nineteen years he participated in the LBMA-PMS (Mann, 1945; Gilbert, 1987). Because the APE ignores the directional error of the forecasting error, a one-sided confidence value of 1.65 was used based on a 95% confidence interval. The comparable confidence value for a 90% confidence interval is 1.28. Norman's Mann-Kendall statistic value of 1.00 indicated an increasing accuracy trend over the nineteen years he participated, but not sufficient to support the hypothesis that he reliably predicted the future gold price.

For the CM panel data, with only one participant performing marginally better than the aggregated consensus lower quartile in forecasting the gold price, it supports rejecting the hypothesis that an individual expert can reliably predict metal prices. The LBMA-PMS data reinforces the CM panel findings that a particular expert can reliably forecast metal prices. The insight gained from the LBMA-PMS longer forecasting data horizon analyzed was an individual expert might perform well compared to the aggregated consensus estimates. However, the comparable performance may not be significantly better than the alternative aggregated consensus estimates. The expert's possible inconsistent performance over successive forecasting periods may undermine the reliability looked for by the natural resource industry organizations, supporting the consideration of an aggregated consensus estimate that reflects a broader group of insights. The CM panel and LBMA-PMS data analysis did not support the possibility of an individual expert consistently proving reliable in forecasting metal prices. Null Hypothesis 1 – Alternative 1 advocating an individual expert as a reliable forecaster of metal prices is rejected.

Null Hypothesis 1 - Alternative 2

As alternative aggregated consensus estimates, the lower and upper quartiles can be used as more reliable metal price forecasts than the aggregated consensus median or specific experts' forecasts.

Null Hypothesis 1 - Alternative 2 Conclusion

The data analysis from the CM panel and LBMA-PMS showed that the lower quartile or upper quartile represented workable aggregated consensus estimates, subject to the proviso that it depended on correctly predicting the market direction of

the metal price being forecasted. Compared to the aggregated consensus average or median, the more reliable performance of the two quartile measures indicated forecasters are prone to bias, such as an anchoring bias linked to the prior period or groupthink reflecting the prevailing metal price expectations based on past trends (Drummond, 2001; Jannis, 1973). The herding behavior observed among analysts is ubiquitous, cited in numerous stock market studies (De Bondt and Forbes, 1999; Rülke, 2016; Trueman, 1994; Welch, 2000). The crucial element in choosing the most reliable aggregated consensus estimates is correctly anticipating the forecasted metal price direction (Armstrong, 2008, Greenberg, 2012).

On average, 3 (20%) out of 15 forecasts made by the CM panel copper participants in the seven forecasting periods correctly predicted the rising copper price. Applying the Condorcet Jury Theorem, the calculated average probability for a declining copper price was 98% relative to the increasing market copper price (Mueller, 2003). On average, 5 (~42%) out of 12 forecasts made by the CM panel gold participants in the seven forecasting periods correctly predicted the declining gold price. Applying the Condorcet Jury Theorem, the calculated average probability for an increasing gold price was ~81% compared to the declining market gold price (Mueller, 2003). The performance of the lower or upper quartiles for the CM panel forecasts yielded aggregated consensus estimates that were ~80% more accurate than the performance of the individual forecasters. As aggregated consensus estimates, the lower and upper quartiles were also more reliable than the aggregated consensus average or median, assuming the direction of the market metal price was correctly anticipated.

The possibility of incorrectly anticipating the gold price direction was significant for the LBMA-PMS participants with a 1-year time horizon for the gold price forecasts made over twenty-one years. For 16 (~76%) of the twenty-one years, the LBMA-PMS participants correctly forecasted the gold price direction. Their forecasting performance was better for years when the gold price rose by more than 1% (13/14 ~92%) and worse for years when the gold price fell by more than 1% (3/7 ~43%). Of the actual LBMA-PMS forecasts, ~80% (316/397) correctly predicted the gold price increasing by more than 1%, with a lesser ~49% (73/148) correctly predicted the gold price decreasing by more than 1%. For the fourteen years when the market gold price increased by more than 1%, ~23 (~88%) out of the 26 LBMA-PMS participants predicted the market gold price direction correctly. Using the Condorcet Jury

Theorem, the probability of the LBMA-PMS participants reliably anticipating the market gold price direction was calculated (Mueller, 2003). For the seven years when the market gold price decreased by more than 1%, ~12 (~48%) out of the 25 LBMA-PMS participants correctly predicted the market gold price direction. Using the Condorcet Jury Theorem, the probability of the LBMA-PMS participants reliably anticipating the market gold price direction was calculated at 50% (Mueller, 2003). In 2013 when 18 (~78%) of the 23 LBMA-PMS participants had incorrectly predicted the gold price would increase by more than 1%, the calculated probability using the Condorcet Jury Theorem was 99.9% in favor of an increase, highlighting the possibility of incorrectly predicting the market gold price direction (Mueller, 2003). In 2016 when 27 (~87%) of the 31 LBMA-PMS participants had incorrectly predicted the gold price decreased by more than 1%, the calculated probability using the Condorcet Jury Theorem was 100% in favor of a decrease, similarly reflecting the possibility of incorrectly predicting the market gold price direction (Mueller, 2003). (See Annexure G: LBMA-PMS Condorcet Jury Theorem probabilities).

Based on the CM panel and the LBMA-PMS forecasts analyzed, the possibility of using the aggregated consensus lower or upper quartile depending on the market metal price direction showed better reliability than the other aggregated consensus estimates. The reliability of the lower and upper quartiles as aggregated consensus estimates indicates that they could prove more dependable than relying on a single forecaster. Despite the diversity in consensus forums, the possibility of groupthink may lead to inaccurate forecasts of future market metal price directions, reflected in significant probabilities calculated using the Condorcet Jury Theorem (Mueller, 2003). The possibility of using a Bayesian approach to estimate the prior probabilities of an expected outcome based on the past performance of the consensus panel participants may be a more reliable alternative than purely relying on the participants' forecasted metal price direction (Greenberg, 2012).

Table 4-24 LBMA-PMS Conditional Probabilities summarizes the conditional probabilities for the LBMA-PMS forecasting data over the twenty-one years analyzed. The forecasts can be grouped into two categories. Firstly, the estimates made by the LBMA-PMS participants when the forecasted metal price direction and the actual market annual gold price direction coincided. Secondly, the forecasts made by the LBMA-PMS participants when the forecasted metal price direction and the market

annual average gold price direction were the opposite. The LBMA-PMS forecasts were made at the beginning of January each year, which allowed the estimates to be differentiated into years when the January gold price was more than 1% higher than the prior year's average and the converse of an increase of less than 1%. The LBMA-PMS forecasts are tabulated in *Table 4-24* into those when the January gold price was higher or lower than the prior year and the total forecasts made. In *Table 4-24*, quadrants 1, 5, and 9 reflected the probabilities when the LBMA-PMS participants forecasted an increase of more than 1% in the expected market gold. Their forecasts were later proven correct by the annual average gold price change. In *Table 4-24*, quadrants 4, 8, and 12 reflected the probabilities when the LBMA-PMS participants forecasted an increase of less than 1% in the expected market gold. Their forecasts were later proven correct by the annual average gold price change.

In *Table 4-24*, quadrants 3, 7, and 11 reflected the probabilities when the LBMA-PMS participants forecasted an increase of more than 1% in the expected market gold. Their forecasts were later proven incorrect by the annual average gold price change. In *Table 4-24*, quadrants 2, 6, and 10 reflected the probabilities when the LBMA-PMS participants forecasted an increase of less than 1% in the expected market gold. Their forecasts were later proven incorrect by the annual average gold price change. The shaded quadrants (*Table 4-24*; Quadrants 1,4,5,8,9,12) were the probabilities when the LBMA-PMS participants were correct in their forecasted predictions. From a forecasting reliability perspective, the rest of the unshaded quadrants (*Table 4-24*; Quadrants 2,3,6,7,10,11) represent the challenge in selecting a reliable aggregated consensus estimate.

Table 4-24 LBMA-PMS Conditional Probabilities

LBMA Gold Forecasts Probabilities	**1st Week January increase >1%**		**1st Week January increase <1%**		**All Forecasts Combined**	
Market gold price direction	**Year Average increase >1%**	**Year Average increase <1%**	**Year Average increase >1%**	**Year Average increase <1%**	**Year Average increase >1%**	**Year Average increase <1%**
Forecast increase >1%	1 12 years 93.4%	3 2018 6.6%	5 2009 24.1%	7 2000/13 75.9%	9 13 years 79.6%	11 2000/13/18 20.4%
Forecast increase <1%	2 No specific 83.3%	4 No specific 16.7%	6 2016 27.7%	8 2001/14/15/17 72.3%	10 2016 34.5%	12 2001/14/15/17 65.5%

By differentiating the forecasts according to the January gold price direction, the conditional probabilities for the correct estimates can be improved, as seen in *Table 4-24*, quadrant 1 versus 9 and quadrant 8 versus 12. The probability of the average annual gold price direction mirroring the first week of January gold price direction was ~86% for the twenty-one years of LBMA-PMS data analyzed. Following the January price direction would have resulted in only 2016 being correctly forecasted as declining when the later market price increased by more than 1%, a more reliable outcome than the 4 incorrect forecasted gold price directions of the LBMA-PMS participants.

The average copper and gold forecasts per forecasting period for the CM panel were 15 and 12, respectively. The LBMA-PMS averaged 26 participants annually over the twenty-one years analyzed. Achieving a reliable aggregated consensus estimate did not require a substantial number of participants compared to the many industry forecasters. A core number of participants (e.g., 12) and a similar number of less frequent contributors ensure a diversity of opinions to achieve a reliable aggregated consensus estimate (Armstrong, 2008; Batchelor and Dua, 1995). The number of participants is influenced by the reciprocity for making forecasts (Fisher et al., 2012; Rijshouwer, 2019; Ury, 2013). If the forecasting process includes cyclical revisions and transparency of participants' estimates, it also affects the number of participants (Armstrong, 2001a; Brannick and Coghlan, 2014; Rowe and Wright, 1999; Rowe and Wright, 2011). A crucial element in achieving reliability from an aggregated consensus approach is correctly predicting the future market metal direction (Greenberg, 2012).

The aggregated lower or upper-quartile consensus estimates consistently achieved low APE and Theil U(II) Index values in the CM panel and LBMA-PMS data. The results support the possibility of using either the lower or upper quartile as a reliable metal price forecast (Armstrong, 2008; Batchelor and Dua, 1995). The longer time horizon of the LBMA-PMS data, which covered periods of falling and rising gold prices, shows the need for a selection rule depending on the anticipated metal price direction. The aggregated consensus lower quartile is likely the most reliable metal price estimate for a neutral or declining metal price environment. The aggregated consensus upper quartile will likely be the most reliable estimate for a rising metal price environment. Null Hypothesis 1 – Alternative 2 advocating the selection rule

Q1/Q3 based on the expected market metal price direction is accepted as a reliable alternative metal price forecast.

Hypothesis 2

Repetitive forecasting cycles will improve the reliability of metal price forecasts made by an individual expert.

Hypothesis 2 Conclusion

For the CM panel forecasts made, ~71% of the copper and ~66% of the gold of the participants made more than one forecast in the seven forecasting periods (*Table 4-9 CM panel Cyclical Forecasts - Copper and Gold*). The accuracy gains for the successive rounds of estimates for the CM panel copper participants outstripped the worsening forecasts by ~4 times. The comparable accuracy improvement for the CM panel gold participants was ~2 times. For the LBMA-PMS, ~62% of the participants participated for over a year (*Table 4-10 LBMA-PMS Analysts' Participation Summary*). The average APE for the participants who only took part for one year was ~9.2%. At the opposite end of the participating rate, the participants who took part for eighteen or more years had an average APE of ~7.5%. The overall correlation between participation and APE was ~70%, reflecting an improvement for recursive participation, mirroring the lower APE for participants with multiple years of taking part compared to those who took part for only one year.

The Mann-Kendall statistic tabulated in *Table 4-18 LBMA-PMS Gold Analysts' Forecasting Reliability* for the LBMA-PMS past top analysts shows some consistency with increasing participation, with a negative correlation of ~31%. As participation increases, the extent of deviation from the median diminishes. The decreasing trend of APE and Theil U(II) Index with participation is also evident in *Table 4-18*. The trend is for the analyst's APE to decline with participation and rise again before declining again (*Figure 4-6*). Why the APE declines, rises, and declines again is possibly linked to the reciprocity observed by Wikipedia (Rijshouwer, 2019). The pressure to perform well or face growing criticism consistently is a form of negative reciprocity (Fisher et al., 2012; Ury, 2013). From a reliability perspective, the implication would be to combine numerous metal price forecasts. The need for diversity that comes from having participants with varying levels of historical participation is recommended (Armstong, 2008; Batchelor and Dua; Hong and Page, 2012).

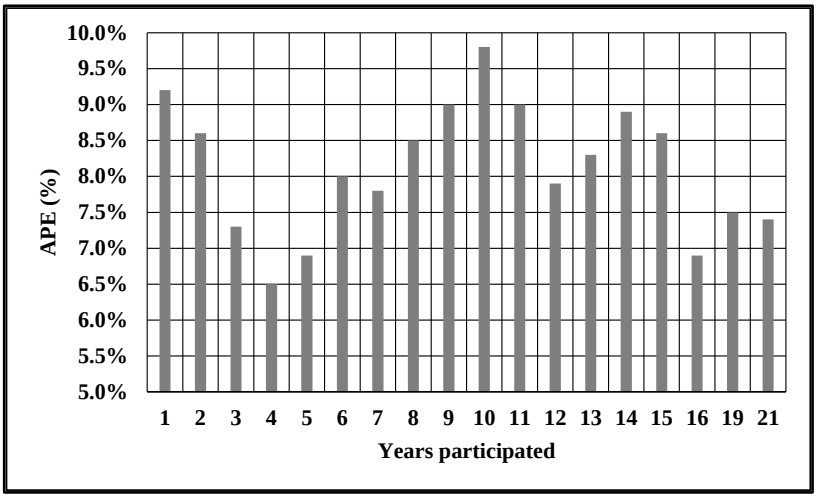


Figure 4-6 LBMA-PMS Analyst Participation versus average APE achieved

The inference drawn from the CM panel and LBMA-PMS forecasting forums, with repetitive participating learning occurs, both within the process through seeing how other participants formulate their forecasts or from the process through developing and understanding the basis for making future predictions (Argyris, 1977, Argyris et al., 1985, Brannick and Coghlan, 2014). The action of forecasting and the reflection on the action of predicting can lead to a better understanding of the dynamics of making predictions and a grasp of the influencing market factors that affect future metal prices. Based on the quantitative analysis, diversity of participation added robustness to the aggregated consensus estimates, which surfaced reliable aggregated consensus estimates, subject to the limitation of correctly predicting the future market metal price direction (Armstrong, 2008; Batchelor and Dua, 1995; Hong and Page, 2012). Hypothesis 2, advocating reliance on an individual expert to make reliable metal price forecasts consistently, is rejected. Despite the evident improvement for participants with the most years of participation, the varying trend of forecasting accuracy cannot be reliably known beforehand.

Null Hypothesis 2 - Alternative 1

Using a collective decision-making approach to forecast metal prices will yield a reliable result.

Null Hypothesis 2 - Alternative 1 Conclusion

With its collective participation in a cooperative inquiry, the CM panel AR project had similarities to a Delphi Method (*Table 3-5 Approach Comparison of CM panel, Delphi Method, and LBMA-PMS*). The CM panel cooperative inquiry differed most

significantly from the Delphi Method in allowing and encouraging direct and transparent interactions between the participants (Brannick and Coghlan, 2014; Heron and Reason, 2008; Pedler, 2012; Wright and Rowe, 1999). The moderator acted as a referee, ensuring participants showed consideration and respect towards each other in their interactions. The moderator's role of aggregating opinions and providing participant feedback was eliminated. The CM panel structure continually allowed observable interactions and the automated aggregation of the emerging consensus average by all participants. On balance, the CM panel showed the possibility of using a cooperative inquiry approach for collecting and aggregating consensus estimates (Brannick and Coghlan, 2014; Heron and Reason, 2008; Hong and Page, 2012).

The CM panel participants did not use the opportunity despite the ability to reflect and comment on their peers' metal price forecasts. In an ad hoc verbal discussion with some of the CM panel participants, they commented on the estimates of others but saw no need to do so formally within the CM panel forum. If the moderator role on the CM panel had been more interactive, as in a Delphi Method, the participants would have felt compelled to respond to the feedback provided by the moderator (Rowe and Wright, 1999). The forecasters implied they had an intuitive feel for future metal prices and saw no need to be influenced by the opinion of other participants, focusing instead on exogenous metal market events and emerging metal price trends. The implication is the CM panel participants have a mental model they rely on to make their forecasts, which is not swayed by sources other than those they consider authoritative (Drummond, 2001; Hong and Page, 2012; Rowe and Wright, 2011). The motivation to make revised forecasts was also driven by the competitive nature of the forecasters to be ranked top or higher than most other participants (Fisher et al., 2012; Rijshouwer, 2019; Ury, 2013). The benefit of the cooperative inquiry structure was in the ongoing participation, which aided in the emergence of more reliable aggregated consensus estimates. The cooperative inquiry proved the possibility of using the approach in the workplace to surface reliable aggregated consensus estimates of future metal prices collectively from a group (Armstrong, 2001a; Batchelor and Dua, 1995; Heron and Reason, 2008).

For the LBMA-PMS, the top participants' performance is published at the beginning of the following year, just before the next LBMA-PMS. The published review focuses on the relative accuracy of the top participants rather than the

aggregated consensus estimates of all the metal price forecasts submitted. The peer and industry recognition given to the LBMA-PMS participants potentially encourages and discourages participation by publicly displaying the relative performance of the participants (Fisher et al., 2012; Rijshouwer, 2019; Ury, 2013). The ~38% (~9% of forecasts made) of the participants only took part for one year. Or ~63% (~22% of forecasts submitted) who took part for three years or less may have been discouraged by the public scrutiny of their specific forecasting performance (*Table 4-10 LBMA-PMS Analysts' Participation Summary*). At the other end of the participation spectrum, 15 (~12%) participants had taken part for ten or more years and accounted for 225 (~41%) of the 545 forecasts, encouraged by LBMA-PMS's public recognition. As the quantitative analysis showed, aggregated consensus estimates offer workable alternative metal price forecasts, made possible by both the short-duration and longstanding participation of the LBMA-PMS forecasters making metal price forecasts (*Table 4-14 LBMA-PMS Aggregated Consensus Estimates Reliability*). The LBMA-PMS aggregated consensus estimates reinforced that both the collective outcome and the individual competitive involvement contributed to the possibility of using a collective approach for forecasting metal prices reliably (Armstrong, 2001a, Batchelor and Dau, 1995; Coghlan and Brannick, 2014; Heron and Reason, 2008; Hong and Page, 2012).

For the CM panel, the impetus to revise forecasts was driven more by exogenous market dynamics. Still, the outcome was participants made recursive forecasts that led to more reliability of the aggregated consensus estimates than the estimates of the individual participants (*Table 4-3 CM panel - Copper Consensus Estimates Reliability; Table 4-6 CM panel – Gold Consensus Estimates Reliability*). Participation in the cooperative inquiry yielded an outcome without the need to compel participants to interact collectively in making a decision (Brannick and Coghlan, 2014; Heron and Reason, 2008). The crucial elements were access to the participation forum and independently identifying a reliable aggregation method (*Figure 2-1 Collective Decision-Making Classification*). The viability of the collective decision-making approach of the CM panel approach for aggregating metal price consensus estimates was confirmed by pooling the LBMA-PMS data and comparing the reliability against the individual participants. Null Hypothesis 2 – Alternative 1 advocating reliance on a

collective decision-making approach to collect and aggregate metal price forecasts and identify a reliable aggregated consensus estimate is accepted.

Null Hypothesis 2 - Alternative 2

Over repetitive cycles of forecasting metal prices, alternative aggregated consensus estimates will prove more reliable than the median or a specific expert.

Null Hypothesis 2 - Alternative 2 Conclusion

The CM panel based on the copper and gold forecasts submitted, the aggregated consensus upper quartile and lower quartile respectively yielded more reliable predictions than the other aggregated consensus estimates (*Table 4-3 CM panel - Copper Consensus Estimates Reliability; Table 4-6 CM panel – Gold Consensus Estimates Reliability*). For the twenty-one years (2000-2020) of LBMA-PMS data analyzed, the quantitative results identified the aggregated consensus lower quartile as the most reliable forecast for neutral or declining metal price environments. Conversely, for a rising metal price environment, the aggregated consensus upper quartile was identified as a reliable forecast (*Table 4-14 LBMA-PMS Aggregated Consensus Estimates Reliability*). The lower or upper quartile selection was proposed as a forecasting rule, subject to the proviso of correctly predicting the future market metal price (*Table 4-16 LBMA-PMS Gold Aggregated Consensus Estimates Selection Errors*).

The reliability of combing collective forecasts to yield a reliable result is identified by Armstrong (2001a), Batchelor and Dua (1995), Galton (1907a), Surowiecki (2004), and Tetlock and Gardner (2016). An extension of the value of combining collective estimates is provided by Hong and Page (2012) in their explanation of how the diversity of mental models for interpreting data leads to more reliable outcomes. An assumption underlying the collective decision-making recommendations of the preceding authors is some implicit understanding of the participants' forecasting method and capability. The expectation of necessary expertise is reflected in Aristotle's view that deliberating decisions is best left to those more capable of understanding the decision's implications, '*the wise man rule*' (Flyvbjerg et al., 2014).

The focus in collective decision-making is on access and why and which participants should be allowed to participate (Sunstein, 2006; Rowe and Wright, 1999). The democracy of participation supersedes the process of the aggregation of the

outcome. Galton (1907a) advocated the median, as it equally split the views above and below the statistical midpoint. The support for the median is made by Levy and Peart (2002), while the preference of Hong and Page (2012) see the average as more reliable as it more equally balances the forecasting errors on either side of the central point. A limitation of either the median or average as the most reliable estimate of future metal prices is market metal prices are directional. Asking a group to guess the weight of an ox does not allow for the possibility of a zero or negative outcome. Market prices, by contrast, can remain unchanged, decrease, or increase. Hayek (1945) recognized the potential of market prices moving in multiple directions, as the participants with situation-specific insights acted according to their views. Surowiecki (2004) differentiates collective decisions to be made as either "*cognition*," a definitive outcome; or "*coordination*," to achieve a common objective; or "*cooperation*," to balance conflicting interests. Forecasting metal prices do not fall into one of the three categories identified by Surowiecki. The forecaster is making a judgmental prediction, which cannot be expected to determine the future market metal price and does not require either the coordination or cooperation of other forecasters to be made.

The quantitative results from the CM panel and LBMA-PMS forecasting forums point to the possibility of using a collaborative approach for surfacing aggregated consensus estimates, which are reliable when the market direction is correctly anticipated (Ahlburg, 1984; Armstrong, 2001a; Bliemel, 1973; Hyndman and Koehler, 2006; Koutsoyiannis, 1977). The aggregated consensus estimates achieved reliability without specific recursive action, reflection, and action cycles by both the participants of the CM panel and LBMA-PMS. Making judgmental metal price forecasts appears to be an internal rather than external reflective practice, implying the aggregation of the participants' forecasts is paramount for a more reliable outcome than the ability to observe the opinions and forecasting processes of other forecasters (Hayek, 1945). The implication is that the participation structure that allows access is crucial to the viability of a collective approach (Badaracco,1992; Goodpaster, 1991). The need to be democratic and use an aggregation method that balances the opinions of those involved in making the predictions sees the aggregated consensus outcome as a "*common good*" when the aggregated consensus outcome is a derivation of the collective actions of those involved, without detracting from their personal views (Ostrom, 1973).

Null Hypothesis 2 – Alternative 2 advocating reliance on a collective decision-making approach to collect and aggregate make reliable judgmental metal price forecasts is accepted. The consistent reliability of aggregated consensus estimates, capturing the diversity of views of the participants without having to draw an inference about their capabilities, makes a cooperative inquiry approach viable for forecasting metal prices.

Hypothesis 3:

In making metal price forecasts, forecasters have a consistent expectation and explanation of the future metal price outcome.

Hypothesis 3 Conclusion

The qualitative analysis of the CM panel copper forecast explanations showed a negative correlation of circa 31% for both the word count and the number of justifications used to support the forecasts made and the rank based on an ascending APE (*Table 4-8 CM panel Keyword Summary*). The qualitative analysis showed a negative correlation of around 39% for the CM panel gold forecast explanations (*Table 4-8 CM panel Keyword Summary*). The implication from the qualitative analysis shows a weak relationship between forecast reliability and forecast accuracy, supporting the expectations of Armstrong (2001a) and Hyndman and Koehler (2006) for forecasting reliability. The use of keywords by the copper and gold participants in their respective CM panel justifications was high, with approximately 80% of the main occurrences accounted for by 5 keywords (Bazerman and Moore, 2008; Drummond, 2001). The commonality in the keyword occurrences of the CM panel participants shows a similar belief amongst participants of the influencing factors for the metal prices rather than a wide diversity of opinions (Janis, 1973).

For the LBMA-PMS qualitative analysis, the correlation between the aggregated consensus median APE was low for keyword occurrences (~12%), word count (~11%), and the number of participants (~8%), showing a marginal relationship between the justifications and forecasting accuracy. When the LBMA-PMS keywords were analyzed according to the World Gold Council (Gold, 2021) factors, the correlations from some groups improved, but not all, similarly pointing to the possibility of a low association between explanations and forecasting accuracy. The narrower choice of keywords based on the collocates helped to improve the correlation

with the aggregated consensus median marginally but not significantly to support the possibility of a relationship between requiring justifications and forecasting accuracy (Armstrong, 2001a; Hyndman and Koehler, 2006). It is also possible the high correlation year-on-year between the LBMA-PMS explanations might reflect a rote explanation approach, accounting for the weak correlation (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973). The pooling of the LBMA-PMS justifications could have resulted in the explanations of the more correct forecasters being obscured by the other forecasters' diverse opinions. It would require further research to verify the possibility.

The qualitative correlation results for the CM panel and LBMA-PMS are not comparable. The LBMA-PMS keyword analysis was done collectively for all participants, compared to the CM panel participants individually. However, given the relative consistency of the LBMA-PMS aggregated consensus estimates forecasting performance accuracy of around ~7%, the diverse insights were expected to reflect a correlation comparable to the CM panel correlation results (*Table 4-10 LBMA-PMS Analysts' Participation Summary*). A fundamental difference was that the CM panel approach had transparency in the explanations made and allowed participants to make recursive forecasts. Further research would be needed to verify if the different methods accounted for the difference.

The consistency in the keyword usage among participants, the significant correlation between succeeding years, and the low correlation with forecasting accuracy make the possibility that participants make quantitative forecasts more reliably than the justifications they supply, which are more intuitive than supportive of their judgmental forecasts. Further research would be needed to test the difference between requiring and not requiring explanations and forecasting accuracy to assess the importance of making metal price forecasts. Hypothesis 3, advocating consistency between forecasting reliability and requiring justifications, is rejected. The inconsistency of the forecasting accuracy as measured by the APE and keyword frequency indicated the possibility of an intuitive judgmental forecast justified by explanations expected by users of the estimates (Bazerman and Moore, 2008; Drummond, 2001).

Null Hypothesis 3 - Alternative 1

Metal price forecasters focus more on the quantitative rather than the qualitative element of their predictions.

Null Hypothesis 3 - Alternative 1 Conclusion

The CM panel copper and gold participants did not correctly predict the direction of the metal price, with more distant forecasts more prone to error (*Table 4-3 CM panel - Copper Consensus Estimates Reliability; Table 4-6 CM panel – Gold Consensus Estimates Reliability*). The correlation between the forecast accuracy and justifications averaged in the ~30% range for copper and gold (*Table 4-8 CM panel Keyword Summary*). The divergence between the forecasted metal price direction and the justification correlations shows that the benefit of considering the shared insights is insignificant in selecting a particular forecaster as a reliable metal price predictor (Bazerman and Moore, 2008; Drummond, 2001).

The low correlation for the overall LBMA-PMS justifications classified according to the World Gold Council (Gold, 2021) categories indicated limited value in considering the qualitative analysis of explanations provided as an indicator for selecting an aggregated consensus estimate (*Table 4-22 LBMA-PMS World Gold Council Keyword Correlations*). As it cannot be known beforehand which individual forecasting participant would be the most accurate, analyzing the participant's explanations may be of limited benefit in selecting a specific forecaster's prediction (Armstrong, 2001a; Batchelor and Dua, 1995).

The better reliability of the aggregated consensus lower or upper quartiles, when considering the market metal price direction, would imply that the outlying forecasters from the center towards the lower and upper limits, the outliers, are more likely to be correct. The herding observed in forecasting and the prevalence of groupthink supports the importance of focusing on the outliers (De Bondt and Forbes, 1999; Rülke, 2016; Trueman, 1994; Welch, 2000). It may be that the opinions expressed by the less reliable forecasters, the outliers, could offer insights into more reliable forecasts. However, the significant conformity in the justifications shows that the outliers do not express their views well or are not considered well enough by the other forecasters for their comments to become mainstream. Consequently, it would be more informative to use a quantitative approach to find a possible, reliable aggregated consensus estimate and then consider their justifications' qualitative merits.

The merits of either the aggregated consensus lower or upper quartiles as reliable estimates of the predicted market gold price suggest the LBMA-PMS justifications should be similarly stratified to test the level of coherence with the relevant aggregated consensus estimates. Without analyzing the LBMA-PMS justifications according to the quartile in which the matching forecast fell, it is impossible to confirm or discard the possibility that the participants' explanations offered any insight into the choice of the best-aggregated consensus estimate. The aggregated median and average offered limited reliability. The lack of correlation with the justifications showed the aggregated average keyword occurrences were less valuable for informing the aggregated consensus median's reliability.

Null Hypothesis 3 – Alternative 1, advocating more reliance on the quantitative instead of the qualitative elements of metal price forecasters' predictions, is accepted. The low correlation between keyword occurrences and forecasting reliability measured by the APE over multiple periods indicates an inconsistency between the quantitative and qualitative elements of metal price forecasts. A possible explanation is using typical explanations expected by users of metal price forecasts instead of more reflective opinions (Argyris, 1977; Argyris et al., 1985).

Null Hypothesis 3 - Alternative 2

Metal price forecasters focus on similar exogenous factors when making forecasts and supporting justifications.

Null Hypothesis 3 - Alternative 2 Conclusion

The number of CM panel forecasts per period (average copper 15 and gold 12) combined with no justifications provided in some instances (copper ~20% and gold ~23%) influenced the participants' ability to assess the coherence explanations fully. A sign of the CM panel coherence for the copper and gold participants could be deduced, with 5 of the 11 most frequent keywords accounting for ~80% of the primary keyword occurrences recorded (*Table 4-8 CM panel Keyword Summary*). The correlation between the CM panel copper and the gold keyword rankings was ~64%, showing that participants used similar keywords in their explanations in successive forecasting periods. The possibility of bias in the coherence of participants' expectations was noticeable in the conformity of their forecasted metal price directions, reflected in the calculated Condorcet Jury Theorem probabilities (Mueller,

2003). Combining the shared forecasted metal price directions and common usage of keywords could show a groupthink bias amongst the forecasters linked to their shared interest in the metal markets (Janis, 1973). Further research comparing the typical views with those of the participants falling in the lower or upper quartiles would need to be undertaken to confirm a general bias by comparing their explanations for different keyword frequencies.

The low correlation between the keywords and aggregated consensus median APE (~12%) for the LBMA-PMS would imply a low coherence with the forecasting reliability (*Table 4-22 LBMA-PMS World Gold Council Keyword Correlations*). The year-on-year average correlation of ~83% for the LMBA-PMS participants' reuse of keywords in consecutive years reflects a strong consistency in the participants' perceptions of the gold market (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973). However, given the low correlation between the forecasting reliability, keyword occurrences, and repetitive use of keywords over successive years, it does not necessarily support a reliable association. The herding seen in the forecasting behavior of the LBMA-PMS participants, as reflected in the significant Condorcet Jury Theorem probabilities, indicated the quantitative expectations of the forecasters were frequently similar, which is also reflected in the continued use of similar keywords (De Bondt and Forbes, 1999; Rülke, 2016; Trueman, 1994; Welch, 2000).

The significant cross-correlations between the key factors identified by the World Gold Council (Gold, 21) based on the frequencies of occurrences indicates the consistency of use in the explanations provided by the LBMA-PMS analysts (*Table 4-22 LBMA-PMS World Gold Council Keyword Correlations).* The ten keywords listed in *Table 4-23 LBMA-PMS World Gold Council Main Keyword Collocates* that accounted for around 60% of all the keywords identified indicate a significant overlap in the explanations provided by the LBMA-PMS analysts. Null Hypothesis 3 – Alternative 2 advocating forecasters focus on similar explanations for their differing metal price forecasts is accepted. The significant correlation between keyword occurrences over successive years and a select few keywords representing most of the occurrences point to a consistency in the explanations provided by forecasters (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973). The herding in explanations goes contrary to the expectations of Hong and Page (2012) that forecasters use different mental models.

4.9. **Conclusion: Results and Findings**

The AR project, structured as a cooperative inquiry, evaluated the possibility of using a collective decision-making approach to improve metal price forecasts' reliability. The hypotheses analyzed focused on the reliability of an individual expert compared to the different aggregated consensus estimates. The importance of requiring forecasters to justify their quantitative forecasts was analyzed to assess if the requirement improved the metal price forecasts' reliability.

Table 4-25 Research Hypotheses Findings

Hypothesis	Result
Hypothesis 1 The aggregated consensus median will be the most reliable metal price forecast over multiple periods, capturing the diversity of estimates around the statistical midpoint.	Rejected
Null Hypothesis 1 - Alternative 1 An individual expert can reliably forecast metal prices over multiple periods more than the aggregated consensus median.	Rejected
Null Hypothesis 1 - Alternative 2 As alternative aggregated consensus estimates, the lower and upper quartiles can be used as more reliable metal price forecasts than the aggregated consensus median or specific experts' forecasts.	Accepted
Hypothesis 2 Repetitive forecasting cycles will improve the reliability of metal price forecasts made by an individual expert.	Rejected
Null Hypothesis 2 - Alternative 1 Using a collective decision-making approach to forecast metal prices will yield a reliable result.	Accepted
Null Hypothesis 2 - Alternative 2 Over repetitive cycles of forecasting metal prices, alternative aggregated consensus estimates will prove more reliable than the median or a specific expert.	Accepted
Hypothesis 3: In making metal price forecasts, forecasters have a consistent expectation and explanation of the future metal price outcome.	Rejected
Null Hypothesis 3 - Alternative 1 Metal price forecasters focus more on the quantitative rather than the qualitative element of their predictions.	Accepted
Null Hypothesis 3 - Alternative 2 Metal price forecasters focus on similar exogenous factors when making forecasts and supporting justifications.	Accepted

The learning outcome from the AR project as it pertains to the possibility of using a cooperative inquiry approach for forecasting metal prices in the workplace were:

- The most significant bias to consider is not the possible anchoring by participants on current metal price levels or trends but the phenomenon of industry participants holding a common expectation about the future metal price direction (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973). The possibility of participants gravitating toward an incorrect outcome is discussed by Ackerman and Fishkin (2005) and Sunstein (2006), with the probability of the wrong outcome occurring reflected in the Condorcet Jury Theorem (Baker, 1976; Mueller, 2003). The herding phenomenon is noted as a common occurrence among analysts in financial markets (De Bondt and Forbes, 1999; Rülke, 2016; Trueman, 1994; Welch, 2000).
- Metal prices are dynamic, and allowing participants to make cyclical forecasts can improve the forecasted consensus estimate's reliability over time (Argyris, 1977; Argyris et al., 1985). Making a forecast, reflecting on the newly emerging information, and taking the opportunity to make a revised forecast mirrors the recursive interactions action learning (Brannick and Coghlan, 2014; Coghlan, 2007).
- Asking participants to explain their metal price forecast does not necessarily improve the reliability of the estimates provided (Argyris, 1977; Bazerman and Moore, 2008; Drummond, 2001). The expectation that justifying an opinion will improve reliability was not evident (Armstrong, 2001a; Hyndman and Koehler, 2006). If what is being judged is the quantitative result, that is more likely to be well considered.
- The consensus average and median are less reliable as aggregated consensus estimates than the lower and upper quartile, depending on the future market direction of the forecasted metal prices. The distribution of forecasting errors around either the average or median is equally distributed when the nature of the forecast is judgmental, as expected (Galton, 1907a; Hong and Page, 2012; Levy and Peart, 2002).
- Participation in the cooperative inquiry by metal price forecasting participants allows a reliable consensus estimate to surface (Heron and Reason, 2008; Rowe and Wright, 2011). A balance of shorter-duration and longer-duration forecast contributions is a workable approach, allowing for the reciprocity expected by participants to continue participating (Fisher et al., 2012; Rijshouwer, 2019; Ury, 2013).

Further aspects that could be included in the list of lessons learned from the cooperative inquiry and should be considered in the implementation in practice are:

- The CM panel approach used a probability distribution to collect the metal price forecasts instead of a single-point estimate. The metal price probability distribution made it possible to evaluate the trend in the metal price forecasts and the relative contribution to the aggregated consensus average (Armstrong, 2001a; Greenberg, 2012; Hyndman and Koehler, 2006). The LBMA-PMS required that participants supply an expected Low and High for the forecasted year. Taking the average of the Low and High values provided a second forecast estimate for the LBMA-PMS. The comparative accuracy measures for all the aggregated consensus estimates were lower than for the actual forecasts (See Annexure H: LBMA-PMS gold high-low aggregated consensus estimates).
- The explanations for both the CM panel and LBMA-PMS forecasting forums showed a high degree of consistency in the occurrences of the keywords (Bazerman and Moore, 2008; Drummond, 2001; Janis, 1973). Combined with the low correlation between the keywords and the forecast accuracy, it indicates a possible standard response to the justifications provided (Argyris, 1977; Argyris et al., 1985). From a consistency perspective, if the keywords are grouped, and the participants are asked to give a weighted influence percentage, be it negative or positive, such that the absolute total sums to 100%, the participants' collective expectations could be collated and analyzed more accurately (Armstrong 2001a; Hyndman and Koehler, 2006).
- The most crucial aspect of finding the most reliable aggregated consensus estimate is the accuracy of the forecasted metal price direction (Greenberg, 2012; Mueller, 2003). If the participants provide an explanation, it should focus primarily on this crucial issue. The reliability of the consensus approach rests on correctly anticipating the market metal price direction. Asking participants to justify their specific expectations may lessen the similar predicted market metal price direction seen in the analyzed CM panel and LBMA-PMS data (Argyris, 1977; Gioia and Chittipeddi, 1991; Weick, 1988).
- For a sustainable consensus panel to gather the metal price forecasts and calculate the most reliable aggregated consensus estimate, participants with a suitable skill set and willingness to participate are required (Batchelor and Dua, 1995; Hong and

Page, 2012; Tetlock and Gardner, 2016). The attrition rate amongst the CM panel and the LBMA-PMS participants requires a strategy to balance ongoing contributions from short-term and long-term participants. The act of participation was identified as being linked to reciprocity, peer recognition within the forecasting forum, and industry recognition for being part of the forecasting forum. Gaining and keeping participation requires reciprocity, particularly recognition (Fisher et al., 2012; Rijshouwer, 2019; Ury, 2013). The aspects underlying recognition would appear to be appreciated for participating, the opportunity to compete with peers, and acknowledgment for accurate forecasting. The importance of collective participation to surface reliable aggregated consensus estimates requires giving specific recognition for contributing instead of focusing on the most accurate forecaster. It requires a balance between the need for a group to aggregate their forecasts and the recognition of being the best to encourage accurate forecasts that foster a more reliable aggregated consensus estimate (Badaracco, 1992; Ostrom, 1988).

The CM panel used online communication technology, which allowed asynchronous interaction and complete visibility of the emerging aggregated consensus average. It enabled the participation of 23 participants from 6 different world regions, 7 different workplaces, 5 different professional backgrounds, 6 different forecasting approaches, and ranging in age from mid-thirties to mid-seventies. The CM panel cooperative inquiry achieved the aims of a pilot AR project of showing the possibility of the approach as a workable alternative metal price forecasting practice in the workplace. The method could be sustainable with more time and greater recognition for participation. The reliability of the cooperative inquiry approach used for the CM panel could potentially be improved through some suggested changes (Heron and Reason, 2008; Rowe and Wright, 2011):

- Include an input section for the 5 or 6 main keyword groups found and ask the participants to assign a negative or positive percentage influence it will have on the metal price in the forecasting period.
- Include a table that summarizes the keyword group expectations of the participants and the indicated metal price direction it implies to act as a feedback loop for the quantitative metal price forecasts.

- Have only one section to justify the metal price forecasts, comment on the emerging aggregated consensus estimates, and comment on the estimates of the other participants.
- Include the aggregated consensus lower and upper quartiles in the reported emerging metal price tables and graphs.
- Investigate the possibility of making participation more public, such that the participants can get recognition for their contribution (Fisher et al., 2012; Rijshouwer, 2019; Ury, 2013).

The suggested improvements are doable; however, how the last suggestion of recognition is tackled is crucial. The approach should ideally have the support of a leading industry organization to be sustainable.

The cooperative inquiry was structured to collect and aggregate metal price forecasts, an inherently complex activity prone to herding by forecasters (De Bondt and Forbes, 1999; Rülke, 2016; Trueman, 1994; Welch, 2000). A potential panacea for this organizational problem lies in collaborating and aggregating the collective contributions to surface reliable aggregated consensus estimates. The CM panel cooperative inquiry found elements that could lead to more reliable organizational metal price forecasts and how such an approach could be structured to improve the outcome (Heron and Reason, 2008; Rowe and Wright, 2011).

CHAPTER FIVE – SUMMARY, CONCLUSION, AND FUTURE RESEARCH

5. Summary And Conclusion

The cooperative inquiry discussed in this doctoral thesis was influenced by the consequences of unreliably forecasted metal prices, a problem experienced in the workplace of mining organizations. As a critical planning metric in natural resource organizations and a significant driver of operating performance, the expectation would be that responsible organizational executives and stakeholders would use those resources that could consistently ensure reliable metal price forecasts over multiple years. A resource used by some industry practitioners is to obtain several forecasts from research organizations that supply consensus metal price forecasts or from the research departments of financial institutions that collate consensus metal price forecasts for their clients. Despite the availability of consensus metal price forecasts, industry experts have not fully embraced their use as a solution for forecasting metal prices reliably. The standard practice is to use either the consensus average or median as the most reliable aggregated consensus estimate to avoid the possibility of biased outliers. The AR project found that for judgmental forecasts such as metal prices, either the aggregated consensus lower or upper quartiles were the most reliable forecasted metal price estimates depending on the direction of the market metal prices.

5.1. Summary

Collective decision-making has origins dating back to the 4th Century BC and growing significance in the 21st Century with the advances in internet technology. Aristotle (Aristotle and Everson, 1988) advocated the idea of deliberation to reach a consensus outcome that included the opinions of those best placed to understand and solve the problem. The move towards more collaborative business practices that embraced deliberation was bolstered by Ackerman and Fishkin (2002) and Sunstein (2006). The growing usage of the internet as a resource for collaboration facilitated the growth of organizations such as Wikipedia, which permitted numerous contributors to work collectively on a project (Brabham, 2013; Wales, 2004). Sunstein (2006) focused on the importance of deliberation in shared activities and how collective efforts could achieve more than an individual alone. Tetlock and Garder (2016) explored how the shared insights of a "*crowd*" could surface a reliable outcome based on their collective interactions. The collective decision-making approach is based on the expectation that

diversity will improve the reliability of the result by incorporating different insights into one aggregated consensus estimate (Hong and Page, 2012).

Sir Galton (1907a) approached achieving the consensus outcomes by taking the median of all participating in the process. Condorcet (Baker, 1976; Mueller, 2003) and the law of averages relied on the principle that the weighted opinions of participants should decide the outcome, as, on average, they will reflect the majority consensus opinions. Hong and Page (2012) considered combining diversity to capture individual views. Hayek (1945) saw the value of surfacing diverse insights to facilitate an agreed exchange value by the market participants. Surowiecki (2004) leaned towards the approach of Sir Galton and Condorcet, stressing the importance of the general opinion of the "*crowd*." Rowe and Wright (2011) the moderated average or median. Collective decision-making requires balancing the need for access to the group and how the group outcome is decided (*Figure 2-1 Collective Decision-Making Classification*).

Elements from the alternative collective decision-making approaches were used in the CM panel cooperative inquiry to make metal price forecasts and find a reliable aggregated consensus estimate (Brannick and Coghlan, 2014; Heron and Reason, 2008). Engaging multiple participants is fundamental to surfacing an aggregated consensus estimate. The Delphi Method, which epitomizes a balance between deliberation and collaboration, and with the emerging consensus average or median shared through a moderator with all participants, was replicated in the AR project to facilitate the collection of metal prices forecasts from dispersed participants (Rowe and Wright, 1999; 2011). The cooperative inquiry allowed the collection and aggregation of metal price forecasts from several participants interacting openly and with the ability to act, reflect, and act again in recursive cycles (Coghlan, 2007; Coghlan and Brannick, 2014). The cooperative inquiry reflected that consensus forecasts could only be done "*with people*" with a similar expectation that their collective contribution could be an effective alternative workplace practice (Heron and Reason, 2008).

Two sets of metal price forecasts were sourced during the AR project. The forecasts collected from the CM panel website were the primary data sourced by the researcher by inviting participants to collaborate in a cooperative inquiry focused on making cyclical metal price forecasts. Additional public metal price forecasting data was sourced from the LBMA-PMS for the twenty-one years from 2000 to 2020. Both

data sets of metal price forecasts included a numerical estimate of the metal price for a specific future period and a justification supporting the forecast. A significant difference between the CM panel and LBMA-PMS metal price forecasts was the degree of interaction between the participants. The CM panel participants could see all other participants' forecasts, justification, the progressive aggregated consensus average metal price, and the opportunity to make further cyclical forecasts before the forecasting period ended. The LBMA-PMS participants made their annual forecast and justifications in early January each year without interacting and without the participants having the opportunity to make further revised forecasts. Based on the quantitative analysis of the CM panel and LBMA-PMS forums forecasts, correctly predicting the future direction of the metal prices was critical for selecting the most reliable aggregated consensus estimate. The lower quartile was the most reliable aggregated consensus estimate for a neutral or declining metal price (less than a 1% increase). The upper quartile was the most reliable aggregated consensus estimate for a rising metal price (greater than 1%). The qualitative analysis of justifications for the CM panel participants showed some correlation (mid 30%) with forecasting accuracy (APE), while the similar LBMA-PMS analysis was less conclusive. The finding was that the numerical estimate made by the forecasting participants was more informative than the accompanying justifications provided.

5.2. Conclusion

The CM panel was set up for the online cooperative inquiry, resulting in collecting and analyzing 191 metal price forecasts and 4’000 words of justifications from 23 participants over seven forecasting periods. The CM panel data was triangulated with 545 gold price forecasts and approximately 92’700 words of explanations made by 127 participants over twenty-one years in the LBMA-PMS. The CM panel findings, confirmed by the LBMA-PMS findings, were informative, albeit not exactly as expected when the AR project began. Based on the literature reviewed discussing aggregated consensus estimates, the predominant expectation focused on the average or median as the most reliable consensus measure of collective group forecasts. Based on the literature reviewed, an expectation was held that including a justification would improve the reliability of the metal price forecast. The outcome of the quantitative analysis pointed to either the lower or upper quartiles as the most reliable aggregated

consensus estimates. The qualitative research did not support the importance of requiring a justification to improve the reliability of an aggregated consensus estimate.

The quantitative analysis revealed the most crucial aspect of using aggregated consensus estimates was correctly anticipation of the future direction of the forecasted metal price. If the forecasted metal price direction was correctly predicted, two related selection rules could be used to improve the choice of the most reliable aggregated consensus estimate:

- The lower quartile will be the more reliable aggregated consensus estimate when the predicted metal prices are expected to be neutral or down (less than a 1% increase).
- The upper quartile will be the more reliable aggregated consensus estimate when the predicted metal prices are expected to increase (more than a 1% increase).

The low or high outliers tend to offset each other when using the consensus average or median as the aggregated consensus estimate. At the same time, many forecasters similarly focus their estimates on a commonly perceived value based on recent events, influencing the central aggregated estimates. The implication for selecting the most reliable aggregated consensus estimates shows the estimate will lie either halfway above or below the median and the corresponding outliers, depending on the expected metal price direction. The indicated quartile value will reduce the biases occurring because of the centralized groupthink and the extremes of over-inflated expectations, reflected in a lower MSE and APE for the aggregated consensus relevant quartile estimates.

The AR project qualitative analysis proved unclear, showing a low association between forecast justifications and the submitted forecasts' APE reliability. The qualitative analysis of the forecast justifications showed significant coherence between the beliefs of the forecasters about factors most pertinent to future metal prices. The keyword analysis also showed that the forecasters displayed a high degree of consistency in the factors that affected the metal prices in the successive forecasting period. The low correlation between the keyword groups and forecasting reliability, coupled with the consistent usage of keywords by the forecasters, could show the possibility of rote explanations. The quantitative forecast estimate is more meticulously formulated than the explanation when the forecasters predict future metal

prices. A possible alternative approach that may yield more consistent results is to have the forecasters quantify their expectations of the influence of the critical factors as negative and positive percentages. The modification could allow a more concise analysis of the critical elements, improve the ability to aggregate the estimated influences, and highlight expectations as a bias indicator. The explanations should focus on the anticipated market metal price direction, given the significance of making reliable forecasts.

From a cooperative inquiry perspective, getting active participation, which could influence the level of collaboration, was identified as an element in achieving the AR project's goal of identifying a reliable aggregated consensus estimate. Participation was divided between the participants who remained committed to the metal price forecasting process and those who only took part for a short duration for both the CM panel and the LBMA-PMS. The issue of participation motivation surfaced in interactions with some CM panel participants. It led to the understanding that reciprocity was an expected value exchange that influenced the decision and degree of participation. Reciprocity was associated with the recognition given by competing peers sharing a similar experience of contributing to something important to those involved. Although the collective actions of all participants contributed to the AR project’s results and findings, their willingness to participate made the endeavor successful. The importance of anonymous participation in the cooperative inquiry could have limited participation, and the primary focus of the LBMA-PMS on the top performers could have contributed to the dropout of the less well-recognized participants' contributions. The importance of reciprocity opens the debate around anonymous participation in surveys and panels. Limiting the recognition awarded to participants for their contributions in an observable manner can curtail their ongoing participation. Focusing primarily on the top performers without recognizing the value created from being able to aggregate all the collective contributions could also be seen as limiting the potential of finding reliable aggregated consensus estimates from group participation settings.

5.3. Future Research

An aspect linked directly to the AR project that calls for further analysis is testing the keyword frequency in the LBMA-PMS explanations against the individual participants' annual forecasting performance (APE). The uncertain outcome of the AR

project qualitative analysis cannot be accepted without exploring the possibility that the joint analysis of all participants' explanations caused canceling “noise” that obscured the actual relationship between the keywords and forecasting reliability.

A finding of the AR project was the selection rule for the lower and upper quartiles. The reliability of the proposed lower and upper quartile selection rule needs to be tested with other metal price forecasting data to either confirm the conclusion reached or refine the selection rule. A preliminary quantitative analysis of the LBMA-PMS data for silver (Annexure I), platinum (Annexure J), and palladium (Annexure K) were done to evaluate the most reliable aggregated consensus estimate. The lower and upper quartile rule proved most reliable for the silver and platinum LBMA-PMS data. The median was the most reliable aggregated consensus estimate for the LBMA-PMS palladium data. Although Consensus Economics (2020) was approached to access their data comparable to the LBMA-PMS data, they were unwilling to share. Thomson Reuters GFMS (2020) has similar data on metal price forecasts that could be analyzed. Numerous consensus judgmental financial estimates are reported by institutions, which could also possibly be accessed to test the selection hypothesis rule.

A common expectation about forecasting is that its accuracy will be improved by asking for a justification for the estimate. The aspect needs further research based on the inconclusive findings about explanations in the AR project. The suggestion that the justification requirement instead is set as a table of critical factors with a negative and positive percentage assigned needs to be explored to verify if it offers a more reliable gauge of forecasting reliability in alternative settings and the significance of transparency across all participants.

Correctly predicting the future direction of forecasted metal prices was flagged as a crucial element in improving the reliability of aggregated consensus estimates. How workable it would be to achieve better reliability in predicting future metal market directions is debatable, given the myriad of influencing factors. Classifying the LBMA-PMS forecasts improved the calculated prior probabilities compared to the overall data set. It may be possible to find other indicators that could similarly improve the likelihood of correctly choosing the future market price direction in advance, aiding in selecting the most reliable aggregated consensus estimate.

Researching alternative routes to reach your destination will be possible when embarking on a new journey. An alternative approach may involve asking some experienced travelers which route they would recommend. Setting off in the right direction would be helpful when leaving on the journey. Upon arriving at your destination, the expectations created about the trip would unlikely match your expectations before setting out. As with all road trips, the driver's experience would have differed from the other passengers' experiences and involved making many decisions alone. The AR project had similar overtones, a workplace issue, and a prior expectation of the best alternative approach. Researching the relevant literature created certain expectations, while asking a panel of experts did not precisely confirm the initial expectations held. During the AR project, the researcher needed to make some decisions that influenced the structure of the cooperative inquiry without changing the outcome. After reaching the destination, new insights have been gained on the most reliable route to take, but the findings need to be tested by other travelers to confirm the results.

"*None of us acting alone can achieve success*" – Nelson Mandela (2018-2013)

6. REFERENCES

Abramowicz, M. (2004) 'Prediction Markets, Administrative Decision making, and Predictive Cost-Benefit Analysis' *University of Chicago Law Review*, 71(3), pp. 933-1020.

Ackerman, B. and Fishkin, J. S. (2005) *Deliberation Day*. New Haven & London: Yale University Press.

Adams, Z. and Glück, T. (2015) 'Financialization in commodity markets: A passing trend or the new normal', *Journal of Banking & Finance*, 60, pp. 93-111.

Aggarwal, R., Lucey, B. M. and O'Connor, F. A. (2014) 'Rationality in Precious Metals Forward Markets Evidence of Behavioural Deviations in the Gold Markets', *Social Science Research Network*, pp. 1-29. Available from http://ssrn.com/abstract=2437444 (Accessed April 6, 2018).

Ahlburg, D. A. (1984) 'Forecast Evaluation and improvement Using Theil's Decomposition', *Journal of Forecasting*, 3, pp. 345-351.

Alvin, W.G. (1960) 'The norm of reciprocity: a preliminary statement', *American Sociological Review*, 25(2), pp. 161-178.

Andler, D. (2012) 'What has collective wisdom to do with wisdom', In Landemore, H. and Elster, J. eds. *Collective Wisdom: Principles and Mechanisms*. Cambridge University Press. Kindle Edition.

Arbatli, E. C. and Vasishtha, G. (2012) 'Growth in emerging market economies and the commodity boom of 2003 - 2008 Evidence from growth forecast revisions', *Bank of Canada Working Paper* (2012-8), pp. 1-36. Available from https://www.researchgate.net/publication/241752338_Growth_in_Emerging_Market_Economies_and_the_Commodity_Boom_of_2003-2008_Evidence_from_Growth_Forecast_Revisions (Accessed September 24, 2018).

Argyris, C. (1977) 'Double loop learning in organizations', *Harvard Business Review*, 55(5), pp. 115-125.

Argyris, C., Putnam, R., and Smith, D. (1985) *Action science*. San Francisco: Jossey-Bass.

Aristotle, A. and Everson, S. (1988) *The politics*. Cambridge University Press.

Armstrong, J. S. (2001a) *Combining Forecasts*. In *Principles of Forecasting*, ed. Armstrong, J. S. Boston: Kluwer Academic.

Armstrong, J. S. (2001b) *Evaluating Forecasting Methods*. In *Principles of Forecasting*, ed. Armstrong, J. S. Boston: Kluwer Academic.

Armstrong, J. S. (2008) 'Combining Forecasts', *Social Science Research Network*, pp. 1-19. Available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1164582 (Accessed September 15, 2021).

Aruga, K. and Managi, S. (2011) 'Price linkages in the copper futures, primary, and scrap markets', *Resources, Conservation and Recycling*, 56, pp. 43-47.

Badaracco, J.L. (1992) 'Business ethics: Four spheres of executive responsibility', *California Management Review*, 34(3), pp. 64-79.

Baker, K. M. (1976) *Condorcet, Selected Writings*. Indianapolis: Bobbs-Merrill.

Baron, R. S., Hoppe, S. I., Kao, C. F., Brunsman, B., Linneweh, B. and Rogers, D. (1996) 'Social Corroboration and Opinion Extremity', *Journal of Experimental Social Psychology*, 32(6), pp. 537-560.

Batchelor, R. and Dua, P. (1995) 'Forecaster Diversity and the Benefits of Combing Forecasts'. *Management Science*, 14(1), pp. 68-75.

Bazerman, M.H., and Moore, D.A. (2008) *Judgment in Managerial Decision-making*. 7th ed. Chichester: Wiley.

Beber, A. and Pianay, J. (2016) 'Expectations, Fundamentals, and Asset Returns Evidence from the Commodity Markets', *Social Science Research Network*, pp. 1-72. Available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2800172 (Accessed September 24, 2018).

Bek, C. (2013) 'Can Technical Analysis Predict Commodity Price Developments', (M.Sc. Economics & Business Administration: Finance & Strategic Management Masters), Copenhagen Business School, Copenhagen. Available from

https://research.cbs.dk/en/studentProjects/417ff9e2-a02f-4518-b513-3eb197b645b7 (Accessed June 6, 2018).

Berg, J. E., Nelson, F. D. and Rietz, T. A. (2008) 'Prediction market accuracy in the long run', *International Journal of Forecasting*, 24(2), pp. 285-300.

Blatter, J. K. (2008) Case Study. In *The Sage Encyclopedia of Qualitative Research Methods*. Eds Given, L., Los Angeles, Calif.: Sage Publications.

Bliemel, F. (1973) 'Theil's Forecast Accuracy Coefficient: A Clarification', *Journal of Marketing Research*, X, pp. 444-6.

Blinder, A. S. and Morgan, J. (2000) 'Are Two Heads Better Than One? An Experimental Analysis of Group vs. Individual Decision making', *National Bureau of Economic Research*, Working Paper No. 7909, pp. 6-47. Available from http://www.nber.org/papers/w7909 (Accessed October 19, 2019).

Bolger, F., Stranieri, A., Wright, G. and Yearwood, J. (2011) 'Does the Delphi process lead to increased accuracy in group-based judgmental forecasts or does it simply induce consensus amongst judgmental forecasters?', *Technological Forecasting & Social Change*, 78(9), pp. 1671-1680.

Bolger, F. and Wright, G. (2011) 'Improving the Delphi process: Lessons from social psychological research', *Technological Forecasting & Social Change*, 78(9), pp. 1500-1513.

Borin, A., Cristadoro, R., Golinelli, R. and Parigi, G. (2012) 'Forecasting World Output the Rising Importance of Emerging Economies', *Social Science Research Network*, pp. 1-45. Available from http://ssrn.com/abstract=2030809 (Accessed September 5, 2018).

Bottom, W. P., Ladha, K. and Miller, G. J. (2002) 'Propagation of Individual Bias through Group Judgment: Error in the Treatment of Asymmetrically Informative Signals', *Journal of Risk and Uncertainty*, 25(2), pp. 147-163.

Boudoukh, J., Richardson, M., Shen, Y., and Whitelaw, R. F. (2007) 'Do Asset Prices Reflect Fundamentals Freshly Squeezed Evidence from the OJ Market', *Journal of Financial Economics*, 83(2), pp. 397-412.

Brabham, D. C. (2013) *Crowdsourcing*. The MIT Press Essential Knowledge series. MIT Press. Kindle Edition.

Bradbury, H., and Mainmelis, C. (2001) 'Learning history and organizational praxis', *Journal of Management Inquiry*, 10(4), pp. 340–357.

Bredin, D. and Poti, V. (2018) 'Commodity Pricing Evidence from Rational and Behavioral Models', *Social Science Research Network*, pp. 1-62. Available from http://ssrn.com/abstract=3175275 (Accessed May 23, 2018).

Brier, G. W. (1950) 'Verification of forecasts expressed in terms of probability', Monthly Weather Review, 78(1), pp. 1-3.

Brooks, C., Prokopczuk, M. and Wu, Y. (2015) 'Booms and Busts in Commodity Markets Bubbles or Fundamentals', *Journal of Futures Markets*, 35(10), pp. 1-38.

Brown, B. B. (1968) 'Delphi Process: A Methodology Used for the Elicitation of Opinions of Experts', RAND Corporation, Santa Monica, CA. pp. 1-18. Available from https://www.rand.org/pubs/papers/P3925.html. (Accessed December 11, 2021).

Brown, R. (2000) *Group Processes: Dynamics within and Between Groups*. 2nd ed. Oxford: Blackwell.

Cammack, D. (2013) 'Aristotle on the Virtue of the Multitude', *Political Theory*, 41, pp. 175-202.

Cennamo, C., Berrone, P. and Gomez-Mejia, L.R. (2009) 'Does stakeholder management have a dark side?', *Journal of Business Ethics*, 89(4), pp. 491-507.

Chen, K. and Plott, C. R. (2002) 'Information Aggregation Mechanisms: Concept, Design, and Implementation for a Sales Forecasting Problem', Division of the Humanities and Social Sciences, California Institute of Technology, *Social Science Working Paper* No. 113. Available from http://www.hss.caltech.edu/SSPapers/wp1131.pdf (Accessed September 12, 2019).

Chen, T., Wu, L., and Yan, I. K.-M. (2013) 'On the Use of International Commodity Futures Spread for Forecasting Chinas Net Imports', *The World Economy*, pp. 861-879.

Cheng, I. H., and Xiong, W. (2013) 'The financialization of commodity markets', *National Bureau of Economic Research Working Paper Series* (19642), pp. 1-37.

Available from https://www.nber.org/papers/w19642 (Accessed September 28, 2018).

Chinn, M. D. and Coibion, O. (2010) 'The predictive content of commodity futures', *National Bureau of Economic Research Working Paper Series* (15830), pp. 1-37. Available from https://www.nber.org/papers/w15830 (Accessed September 28, 2018).

Choi, H., Leatham, D. J. and Sukcharoen, K. (2015) 'Oil Price Forecasting Using Crack Spread Futures and Oil Exchange Traded Funds', *Contemporary Economics*, 9(1), pp. 29-44.

Cifarelli, G. and Paesani, P. (2013) 'Speculative Cotton Pricing in the 1920s A Nonlinear Tale of Noise Traders and Fundamentalists', *Social Science Research Network*, pp. 1-30. Available from http://ssrn.com/abstract=2232980 (Accessed June 6, 2018).

CME Group. (2020) Futures & Options Trading for Risk Management - CME Group [Online] Available from https://www.cmegroup.com/ (Accessed September 5, 2020).

Coghlan, D. (2007) 'Insider action research: opportunities and challenges', *Management Research News*, 30(5), pp. 335-343.

Coghlan, D. (2010) 'Seeking common ground in the diversity and diffusion of action research and collaborative management research action modalities: Toward a general empirical method', *In Research in Organizational Change and Development*, 18, pp. 149–181.

Coghlan, D. and Brannick, T. (2014) *Doing action research in your own organization*. 4th ed. London: Sage.

Coghlan, D., Cirella, S., and Shani, A. (2012) 'Action Research and Collaborative Management Research: More than Meets the Eye?', *International Journal of Action Research*, 8(1), pp. 45-67.

Cohen, A. R. and Bradford, D. L. (1989) 'Influence without authority: the use of alliance, reciprocity, and exchange to accomplish work', *Organizational Dynamics*, 17(3), pp. 5-17.

Collins, H.M. (1983) ‘An empirical relativist programme in the sociology of scientific knowledge’, In K.D Knorr-Cetina and M. Mulkay (eds), *Science Observed: Perspectives on the Social Study of Science*. London: Sage, pp. 3–10.

Consensus Economics. (2020) Economic forecasts and indicators [Online] Available from https://www.consensuseconomics.com/ (Accessed December 18, 2020).

Cooper, D. J. and Kagel, J. H. (2005) ‘Are Two Heads Better Than One? Team versus Individual Play in Signalling Games’, *American Economic Review*, 95(3), pp. 477-509.

Cooper, J., Kelly, K.A. and Weaver, K. (2001) ‘Attitudes, norms, and social groups’, *Blackwell Handbook of social psychology: Group processes*. Blackwell Publishing, pp.259-282.

Cowgill, B. (2005) *Putting Crowd Wisdom to Work*. Available from https://googleblog.blogspot.com/2005/09/putting-crowd-wisdom-to-work.html (Accessed January 2, 2019).

Creswell, J. W. (2012) *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. 3rd Edition. VitalSource Bookshelf version.

Crowson, P. (2018) ‘Intensity of use reexamined’, *Mineral Economics*, 31, pp. 61-70.

Cuddington, J. T. (1992) ‘Long-Run Trends in 26 primary commodity prices’, *Journal of Development Economics*, 39, pp. 207-227.

Da, Z. and Huang, X. (2018) ‘Harnessing the Wisdom of Crowds’, Social Science Network, pp. 1-42. Available from https://ssrn.com/abstract=2731884 (Accessed November 16, 2022).

Dalkey, N. and Helmer, O. (1963) ‘An Experimental Application of the Delphi Method to the Use of Experts’, *Management Science*, 9(3), pp. 458-467.

Deaton, A. and Laroque, G. (2003) ‘A Model of Commodity Prices After Sir Arthur Lewis’, *Journal of Development Economics*, 71, pp. 289-310.

De Bondt, W.F. and Forbes, W.P. (1999) ‘Herding in analyst earnings forecasts: evidence from the United Kingdom’, *European Financial Management*, 5, pp. 143-163.

Demirer, R., Lee, H-T. and Lien, D. (2015) 'Commodity Financialization and Herd Behavior in Commodity Futures Markets', *International Review of Financial Analysis*, 39, pp. 1-28. Available from https://ssrn.com/abstract=2265506 (Accessed April 6, 2018).

Dietrich, F. and Spiekermann, K. (2016) 'Jury Theorems', Munich University Library, pp. 1-16. Available from https://mpra.ub.uni-muenchen.de/72951/ (Accessed September 21, 2019).

Dobra, J. and Dobra, M. (2014) 'Another look at non-renewable resource exhaustion', *Mineral Economics*, 27, pp. 33-41.

Drogen, L. and Jha, V. (2013) 'Generating Abnormal Returns Using Crowdsourced Earnings Forecasts from Estimize', Social Science Network, pp. 1-13. Available from https://ssrn.com/abstract=2337709 (Accessed November 16, 2022).

Drummond, H. (2001) *The Art of Decision-making: Mirrors of Imagination, Masks of Fate*. Chichester: Wiley.

Easterby-Smith, M., Thorpe, R. and Jackson, P. (2012) *Management Research*. Sage Publications (UK), VitalBook file.

Estellés-Arolas, E. and González-Ladrón-de-Guevara, F. (2012) 'Towards an integrated crowdsourcing definition', *Journal of Information Science*, 38(2), pp. 189-200.

Estimize. (2020) Estimize is designed to collect opinions from the widest possible range of contributors [Online] Available from https://www.estimize.com/ (Accessed April 22, 2020).

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

Fernandez, V. (2015) 'Influence in commodity markets: Measuring co-movement globally', *Resources Policy*, 45, pp. 151-164.

Figlewski, S. (1979) 'Subjective information and market efficiency in a betting market', *Journal of Political Economy*, 87(1), pp. 75-88.

Fisher, R. (1989) 'Negotiating Inside Out: What Are the Best Ways to Relate Internal Negotiations with External Ones?', *Negotiation Journal*, 5, pp. 33-41.

Fisher, R., Ury, W. and Patton, B. (2012) *Getting to yes: negotiating agreement without giving in*. London: Random House Business Books.

Flyvbjerg, B. (2004) 'Phronetic Planning Research: Theoretical and methodological reflections', *Planning theory & practice*, 5(3), pp. 283–306. Available from https://ssrn.com/abstract=2278320 (Accessed November 18, 2022).

Flyvbjerg, B. (2006) 'Making Organization Research Matter: Power, Values, and Phronesis', In Clegg, S. R., Hardy, C., Lawrence, T. B. and Nord, W. R. Eds., *The Sage Handbook of Organization Studies*, 2nd edition, Thousand Oaks, CA, Sage, pp. 370-387. Available from https://ssrn.com/abstract=2278252 (Accessed November 18, 2022).

Flyvbjerg, B., Clegg, S. R. and Haugaard, M. (2014) 'Reflections on Phronetic Social Science: A Dialogue between Stewart Clegg, Bent Flyvbjerg, and Mark Haugaard', *Journal of Political Power*, June 2014, pp. 1-64. Available from https://ssrn.com/abstract=2460651 (Accessed November 18, 2022).

Flyvbjerg, B., Landman, T. and Schram, S. F. (2012) 'Important Next Steps in Phronetic Social Science', In Flyvbjerg, B., Landman, T. and Schram, S. Eds., *Real Social Science: Applied Phronesis*, Cambridge, UK: Cambridge University Press, pp. 285-297. Available from https://ssrn.com/abstract=2278190 (Accessed November 18, 2022).

Forrest, D. and McHale, I. (2005) 'Longshot Bias: Insights from the Betting Market on Men's Professional Tennis', In *Information Efficiency in Financial and Betting Markets*. Ed. Williams, L. V. Cambridge, UK: Cambridge University Press.

Forsythe, R., Rietz, T. A. and Ross, T. W. (1999) 'Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets', *Journal of Economic Behavior & Organization*, 39(1), pp. 83-110.

Frankel, J. A. (2014) 'Effects of speculation and interest rates in a "carry trade" model of commodity prices', *Journal of International Money and Finance*, 42, pp. 88-112.

Frischmann, B. M., Marciano, A. and Ramello, G. B. (2019) 'Retrospectives: Tragedy of the Commons after 50 Years', *Journal of Economic Perspectives*, 33(4), pp. 211–228.

Galton, F. (1907a) 'Vox Populi (The voice of the people)', *Nature*, 75(March 7), pp. 450–51.

Galton, F. (1907b) 'Letters to the Editor – The Ballot-Box', *Nature*, 75(March 28), pp. 509.

Galton, F. (1908) *Memories of my life.* New York. E. P. Dutton and Company, 31 West Twenty-first Street.

Gaugler, T. (2015) 'What drives resource prices? A qualitative review with recommendations for further development of the Hotelling model', *Mineral Economics*, 28(1-2), pp. 37-51.

Gauvin, L. and Rebillard, C. (2015) 'Assessing the Global Impact of a Chinese Hard Landing Through Trade and Commodity Price Channels', *Banque de France Working Paper*, Paper 562, pp. 1-62. Available from http://ssrn.com/abstract=2625958 (Accessed September 3, 2018).

Genre, V., Kenny, G., Meyler, A. and Timmermann, A. (2013) 'Combining expert forecasts: Can anything beat the simple average?', *International Journal of Forecasting*, 29(1), pp. 108-121.

Gigone, D. and Hastie, R. (1997) 'Proper Analysis of the Accuracy of Group Judgments', *Psychological Bulletin*, 121(1), pp. 149–167.

Gilbert, C. L. (1995) 'Modelling Market Fundamentals A Model of the Aluminium Market', *Journal of Applied Econometrics*, 10(4), pp. 385-410.

Gilbert, R.O. (1987) *Statistical Methods for Environmental Pollution Monitoring*. John Wiley and Sons, New York.

Gioia, D. A. and Chittipeddi, K. (1991) 'Sensemaking and sensegiving in strategic change initiation', *Strategic Management Journal*, 12(6), pp. 433-448.

Gold. (2021) World Gold Council [Online] Available from https://www.gold.org/ (Accessed October 1, 2021).

Golding, B. and Campbell, H. (2014) ‘Sustainability of metals production: the case of copper’, *Mineral Economics*, 26, pp.85-95.

Goodpaster, K.E. (1991) ‘Business ethics and stakeholder analysis’, *Business Ethics Quarterly*, 1(1), pp. 53-73.

Gorton, G. and Rouwenhorst, K. G. (2004) ‘Facts and Fantasies about Commodity Futures’, *Financial Analysts Journal*, 62(2), pp. 47-68.

Goss, B. A. and Avsar, S. G. (2013) ‘Simultaneity, Forecasting and Profits in London Copper Futures’, *Australian Economic Papers*, 52(2), pp. 79-96.

Greenberg, E. (2012) *Introduction to Bayesian Econometrics*. Cambridge University Press.

Gustafson, D. H., Shukla, R. K., Delbecq, A. and Walster, G.W. (1973) ‘A Comparative Study of Differences in Subjective Likelihood Estimates Made by Individuals, Interacting Groups, Delphi Groups, and Nominal Groups’, *Organizational Behavior and Human Performance*, 9(2), pp. 280-291.

Guttman, L. (1945) ‘A basis for analyzing test-retest reliability’, *Psychometrika*, 10(4), pp. 255-282.

Guzmán, J. I. and Silva, E. (2018) ‘Copper price determination fundamentals versus non-fundamentals’, *Mineral Economics*, 31(3), pp. 283-300.

Habermas, J. (1998) ‘Between Facts and Norms: An Author’s Reflections’, *Denver University Law Review*, 76, pp. 937.

Hahn, R.W. and Tetlock, P.C. (2006) *Information Markets: A new way of making decisions*. Washington: AEI-Brookings Joint Center for Regulatory Studies.

Hanson, R. (1999) ‘Decision Markets’, *IEEE Intelligent Systems*, 14(3), pp. 16-19.

Haque, M. A., Topal, E. and Lilford, E. (2015) ‘Relationship between the gold price and the Australian dollar-US dollar exchange rate’, *Mineral Economics*, 28, pp. 65-78.

Hardin, G. (1968) ‘The tragedy of the commons: the population problem has no technical solution; it requires a fundamental extension in morality’, *Science*, 162(3859), pp. 1243-1248.

Hastie, R. (1983) *Review Essay: Experimental Evidence of Group Accuracy*. In *Information Pooling and Group Decision Making*, eds. Grofman, B. and Owen, G. Greenwich, CT: JAI Press, pp. 133-146.

Hayek, F. A. (1945) 'The use of knowledge in society', *The American Economic Review*, 35(4), pp. 519-530.

Heath, C. and Gonzales, R. (1995) 'Interaction with Others Increases Decision Confidence but Not Decision Quality: Evidence against Information Collection Views of Interactive Decision Making', *Organizational Behavior and Human Decision Processes*, 61(3), pp. 305–326.

Heron, J. (1996) *Co-operative inquiry: Research into the human condition*. London: Sage.

Heron, J. and Reason, P. (2001) 'The Practice of Co-operative Inquiry Research 'with' rather than 'on' People', In *Handbook of Action Research Participative Inquiry and Practice.* Reason, P. and Bradbury, H. eds. London: Sage Publications, pp. 179-188.

Heron, J., and Reason, P. (2008) 'Extending epistemology within a cooperative inquiry', In: *Handbook of Action Research*, Reason, P. and Bradbury, H. eds. London: Sage Publications, pp. 366–380.

Heylighen, F. (2016) 'Stigmergy as a universal coordination mechanism I: Definition and components', *Cognitive Systems Research*, 38, pp. 4-13.

Hong, H., Du, Q., Wang, A. G., Fan, W., and Xu, D. (2016) 'Crowd Wisdom: The Impact of Opinion Diversity and Participant Independence on Crowd Performance', *Twenty-second Americas Conference on Information Systems*, San Diego, pp. 1-9.

Hong, L. and Page, S. E. (2012) 'Some Microfoundations of Collective Wisdom', In Landemore, H. and Elster, J. eds. *Collective Wisdom: Principles and Mechanisms*. Cambridge University Press. Kindle Edition.

Howe, J. (2006) 'The Rise of Crowdsourcing', *Wired*, 14(6). Available from http://www.wired.com/wired/archive/14.06/crowds.html (Accessed November 18, 2022).

Hu, C. and Xiong, W. (2013) 'Are commodity Futures prices barometer of the global economy?', *National Bureau of Economic Research Working Paper* (No. w19706), pp. 1-48. Available from https://www.nber.org/papers/w19706 (Accessed September 28, 2018).

Huchet, N. and Fam, P. G. (2016) 'The role of speculation in international futures markets on commodity prices', *Research in International Business and Finance*, 37, pp. 49-65.

Humphreys, D. (2019) 'The mining industry after the boom', *Mineral Economics*, 32(2), pp. 145-151.

Hyndman, R. J. and Koehler, A. B. (2006) 'Another look at measures of forecast accuracy', *International Journal of Forecasting*, 22(4), pp. 679-688.

IEM. (2019) Iowa Electronic Markets [Online] Available from https://iemweb.biz.uiowa.edu/ (Accessed December 11, 2019).

Issler, J. V., Rodrigues, C. and Burjack, R. (2014) 'Using common features to understand the behavior of metal-commodity prices and forecast them at different horizons', *Journal of International Money and Finance*, 42, pp. 310-335.

Janis, I. L. (1973) 'Groupthink and group dynamics: A social psychological analysis of defective policy decisions', *Policy Studies Journal*, 2(1), pp. 19-25.

Jeppesen, L. B., Lakhani, K. R., Lohse, P. A., and Panetta, J. A. (2007) 'The Value of Openness in Scientific Problem Solving', *Harvard Business School Working Paper*. Available from http://www.hbs.edu/research/pdf/07-050.pdf (Accessed February 18, 2023).

Johnson, P. and Duberley, J. (2003) 'Reflexivity in Management Research', *Journal of Management Studies*, 40(5), pp. 1279-1303.

Jolls, C. (1998) 'Behavioral Economics Analysis of Redistributive Legal Rules', *Vanderbilt Law Review*, 51(6), pp. 1658-1678.

Kerr, N. L. and Park, E. S. (2001) *Group Performance in Collaborative and Social Dilemma Tasks: Progress and Prospects*. In *Group Processes.* Hogg, M. A. and Tindale, R. S. Eds. Oxford: Blackwell.

Kerr, N. L., MacCoun, R. J. and Kramer, G. P. (1996) 'Bias in Judgment: Comparing Individuals and Groups', *Psychological Review*, 103(4), pp. 689-713.

Keyt, D. and Miller, F. D. (1991) *Aristotle's Theory of Distributive Justice.* In *A Companion to Aristotle's Politics*. Oxford: Basil Blackwell.

Klarreich, E. (2003), 'Best Guess', *Science News*, 164(16), pp. 251-253.

Knorr-Cetina, K. D. (1983) 'The ethnographic study of scientific work: Towards a constructivist interpretation of science', In: K.D. Knorr-Cetina and M. Mulkay (Eds), *Science Observed: Perspectives on the Social Study of Science*. London: Sage, pp. 115–40.

Koutsoyiannis, A. (1977) *Theory of Econometrics.* 2nd Edition. London, UK: The Macmillan Press Ltd.

Krech, D., Crutchfield, R. S. and Ballachey, E. L. (1962) *Individual in Society: A textbook of social psychology*. New York: McGraw-Hill.

Kuhn, T. S. (1970) *The Structure of Scientific Revolutions*. 2nd Edition. Chicago: University of Chicago Press.

Landemore, H. (2012b) 'Collective Wisdom: Old and new', In Landemore, H. and Elster, J. Eds. *Collective Wisdom: Principles and Mechanisms*. Cambridge University Press. Kindle Edition.

Landemore, H. E. (2012a) 'Why the many are smarter than the few and why it matters', *Journal of Deliberative Democracy*, 8(1), pp. 1-14.

LBMA. (2020) The London Bullion Market Association [Online] Available from https://www.lbma.org.uk/ (Accessed August 14, 2020).

Lee, C. M., Ma, P., and Wang, C. (2016) 'The Search for Peer Firms: When Do Crowds Provide Wisdom?', Harvard Business School Accounting & Management Unit Working Paper, (15-032), pp.14-46. Available from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2517496 (Accessed November 18, 2022).

Lescaroux, F. (2009) 'On the excess co-movement of commodity prices — A note about the role of fundamental factors in short-run dynamics,' *Energy Policy*, 37(10), pp. 3906-3913.

Levmore, S. (2002) 'Simply Efficient Markets and the Role of Regulation: Lessons from the Iowa Electronic Markets and the Hollywood Stock Exchange', *Journal of Corporate Law*, 28(4), pp. 589-606.

Levy, D. M. and Peart, S. (2002) 'Galton's two papers on voting as robust estimation', In *Public Choice III*. Mueller, D. C. Cambridge, UK: Cambridge University Press. pp. 357–365.

Lincoln, Y. S., and Guba, E. G. (1985) *Naturalistic inquiry*. Beverly Hills, CA: Sage.

LinkedIn. (2020) Welcome to your professional community [Online] Available from https://www.linkedin.com/ (Accessed July 31, 2020).

List, C. (2008) 'Collective Wisdom Lessons from the Theory of Judgment Aggregation', In Landemore, H. and Elster, J. Eds. *Collective Wisdom: Principles and Mechanisms*. Cambridge University Press. Kindle Edition. pp. 338-367.

List, C. and Goodin, R. E. (2001) 'Epistemic Democracy: Generalizing the Condorcet Jury Theorem', *Journal of Political Philosophy*, 9, pp. 283-297.

LME. (2019). London Metal Exchange [Online] Available from http://www.lme.co.uk/ (Accessed December 11, 2019).

Lord, C. G., Ross, L., and Lepper, M. R. (1979) 'Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence', *Journal of Personality and Social Psychology*, 37(11), pp. 2098-2109.

Lorge, I., Fox, D., Davitz, J., and Brenner, M. (1958) 'A Survey of Studies Contrasting the Quality of Group Performance and Individual Performance', 1920–1957'. *Psychological Bulletin*, 55(6), pp. 337-372.

Loury, G. C. (1994) 'Self-censorship in public discourse: A theory of "political correctness" and related phenomena', *Rationality and Society*, 6(4), pp. 428-461.

Malla, A. (2016) 'Testing the wisdom of crowds: Investigating whether the smartest group actually performs better than other groups', (Degree Programme in Business Information Technology), Oulu University of Applied Sciences. Available from

https://www.theseus.fi/bitstream/10024/118647/3/MALLA_AVINASH.pdf (Accessed October 28, 2018).

Mann, H. B. (1945) 'Nonparametric Tests Against Trend', *Econometrica: Journal of the econometric society*, 13(3), pp. 245-259.

Manski, C. F. (2006) 'Interpreting the Predictions of Prediction Markets', *Economics Letters* 91(3), pp. 425-29.

Marques, J. M., Abrams, D., Paez, D. and Hogg, M.A. (2002) 'Social Categorization, Social Identification, and Rejection of Deviant Group Members', In Hogg, M. A. and Tindale, R. S. eds. *Group Processes.* Oxford: Blackwell.

Mazur, J. P. (2015) 'The Financialization of Commodities', (Master Grande Ecole - Capital Markets & International Banking). *Social Science Research Network*, pp. 1-73. Available from http://ssrn.com/abstract=2620639 (Accessed September 3, 2018).

McKay, J. and Marshall, P. (2001) 'The dual imperatives of action research', *Information Technology & People*. 14(1). pp. 46-59.

Microsoft Excel. (2021) Microsoft Corporation Excel. Retrieved from https://office.microsoft.com/excel

Mill, J. S. (1956) *On Liberty*: Edited with an Introduction by Currin V. Shields. Indianapolis, IN: Bobbs Merrill.

Millard Fernández, C. A. (2016) 'Commodity price forecasts, futures prices, and pricing models' [Online], (Degree of Master of Science in Engineering). Available from https://repositorio.uc.cl/xmlui/bitstream/handle/11534/21176/millard_cristobal.pdf (Accessed May 12, 2018).

Möller, E. (2003) Wikipedia: *WikiLove* [Online] Available from https://en.wikipedia.org/w/index.php?title=Wikipedia:WikiLove&oldid=909010 (Accessed February 25, 2020).

Moore, G. H. and Wallis, W. A. (1943) 'Time Series Significance Tests Based on Signs of Differences', *Journal of the American Statistical Association*, 38(222), pp. 153-164.

Mueller, D. C. (2003) *Public Choice III*. Cambridge, UK: Cambridge University Press.

Orwell, G. (2013) *Animal Farm*. London, England: Penguin Books.

Ostrom, E. (1998) 'A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997', *The American Political Science Review*, 92(1), pp. 1-22.

Ostrom, E. (2000) 'Collective Action and the Evolution of Social Norms', *Journal of Economic Perspectives*, 14(3), pp. 137-158.

Ostrom, V. (1973) 'Can Federalism Make a Difference?', *Publius*, 3(2), pp. 197–237.

Oxford Economics. (2020) Commodity Price Forecasts [Online] Available from https://www.oxfordeconomics.com/ (Accessed December 18, 2020).

Page, S. E. (2007) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. New Jersey, USA: Princeton University Press.

Pedler, M. (2012) *Action Learning for Managers*. Ashgate Publishing Ltd. Kindle Edition.

Perri-Coste, F. H. (1907) 'Letters to the Editor – The Ballot-Box', *Nature*, 75(March 28), pp. 509.

Posner, R. A. (2004) *Catastrophe: Risk and Response*. New York: Oxford University Press.

PwC. (2021) PwC's global mining and metals practice [Online] Available from https://www.pwc.com/gx/en/industries/energy-utilities-resources/mining-metals.html (Accessed June 11, 2021).

Raelin, J. (2009) 'Seeking conceptual clarity in the action modalities', *Action Learning: Research and Practice*, 6(1), pp. 17-24.

Rasmussen. (2021) The Most Comprehensive Public Opinion Data Anywhere [Online] Available from https://www.rasmussenreports.com/ (Accessed June 11, 2021).

Rawls, J. (1971) *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Raymond, E. (2001) *The Cathedral and the Bazaar*. 2nd Edition. Sebastopol, CA: O'Reilly.

Ready-Campbell, N. and Hill, S. (2011) 'Expert Stock Picker: The Wisdom of (Experts in) Crowds', *International Journal of Electronic Commerce*, 15(3), pp. 73-102.

Reagle, J. M. (2010) *Good Faith Collaboration: The Culture of Wikipedia*. Cambridge. London: MIT Press.

Revans, R.W. (1971) *Developing effective managers: A new approach to business education*. London: Longmans.

Revans, R.W. (1998) *ABC of action learning*. London: Lemos & Crane.

Rheingold, H. (2002) *Smart Mobs: The Next Social Revolution*. Cambridge: Basic Books.

Rijshouwer, E.A. (2019) *Organizing Democracy: Power concentration and self-organization in the evolution of Wikipedia* Ph.D. thesis [Online]. Erasmus University Rotterdam. Available from http://hdl.handle.net/1765/113937 (Accessed February 20, 2020).

Roberts, M. C. (2009) 'Duration and characteristics of metal price cycles', *Resources Policy*, 34(3), pp. 87-102.

Roll, R. (1984) 'Orange Juice and Weather', *The American Economic Review*, 74(5), pp. 861-880.

Rowe, G. and Wright, G. (1999) 'The Delphi technique as a forecasting tool: issues and analysis', *International Journal of Forecasting*, 15(4), pp. 353-375.

Rowe, G. and Wright, G. (2001) 'Expert opinions in forecasting: The role of the Delphi Technique', In *Principles of Forecasting: A Handbook for Researchers and Practitioners* (International Series in Operations Research & Management Science). Kindle Edition.

Rowe, G. and Wright, G. (2011) 'The Delphi technique: Past, present, and future prospects — Introduction to the special issue', *Technological Forecasting & Social Change*, 78(9), pp. 1487–1490.

Rülke, J-C., Silgoner, M. and Wörzb, J. (2016) 'Herding behavior of business cycle forecasters', *International Journal of Forecasting*, 32, pp. 23-33.

Sankey, H. (2000) 'Kuhn's Ontological Relativism', *Science & Education*, 9, pp. 59-75.

Sauer, R. D. (1998) 'The Economics of Wagering Markets', *Journal of Economic Literature*, 36(4), pp. 2021-2064.

Schein, E.H. (1987) *The clinical perspective in fieldwork*. Thousand Oaks: Sage.

Schein, E.H. (1995) 'Process consultation, action research and clinical inquiry: Are they the same?', *Journal of Managerial Psychology*, 10(6), pp. 14-19.

Schein, E.H. (2008) 'Clinical inquiry/research', In: Reason, P. and Bradbury, H. (Eds), *Handbook of Action Research* (2nd Edition), London: Sage. pp. 266–279.

Schkade, D., Sunstein, C. R. and Hastie, R. (2007) 'What Really Happened on Deliberation Day?', *California Law Review*, 95, pp. 915-940.

Schram, S., Flyvbjerg, B. and Landman, T. (2013) 'Political Political Science: A Phronetic Approach', *New Political Science*, 35(3), pp. 359-372. Available from https://ssrn.com/abstract=2309929 (Accessed November 18, 2022).

Scott, M. (2020) WordSmith Tools version 8, Stroud: Lexical Analysis Software.

Servan-Schreiber, E., Wolfers, J., Pennock, D. M. and. Galebach, B. (2004) 'Prediction Markets: Does Money Matter?', *Electronic Markets*, 14(3), pp. 243-251.

Shani, A. B., Mohrman, S. A., Pasmore, W., Stymne, B., and Adler, N. (2008) *Handbook of collaborative management research*. Thousand Oaks, CA: Sage.

Shani, A.B., and Pasmore, W.A. (1985) 'Organization inquiry: Towards a new model of the action research process', Reproduced in Coghlan, D. & Shani (Rami), A.B. Editors. *Fundamentals of organization development*, 1, London: Sage. pp. 249-260.

Shefrin, H. (2001) 'Behavioral Corporate Finance', *Journal of Applied Corporate Finance*, 14(3), pp. 113-126.

Shiller, R. (2005) *Irrational Exuberance*. 2nd Edition. Princeton, NJ: Princeton University Press.

Shleifer, A. (2000) *Inefficient Markets: An Introduction to Behavioral Finance*. Oxford: Oxford University Press.

Simmel, G. (1950) *The Sociology of Georg Simmel*, translated and edited by Kurt H. Wolff, Glencoe, Ill.: Free Press, pp. 387.

Snowberg, E., Wolfers, J. and Zitzewitz, E. (2005) *Information (In)efficiency in Prediction Markets*. In *Information Efficiency in Financial and Betting Markets*. Editor Williams, L. V. Cambridge, UK: Cambridge University Press.

Spider Financial Corp. (2021) NumXL Pro: Time Series and Statistics Software (Version 1.67). Available from https://numxl.com/ (Accessed September 16, 2021).

Srivastva, S. and Cooperrider, D. L. (1987) 'Appreciative inquiry in organizational life', *Research in Organizational Change and Development*, 1, pp. 129-169.

Stake, R. (1995) *The art of case study research*. Thousand Oaks, CA: Sage

Stake, R. (2003) Case studies. In Denzin, N. K. and Lincoln, Y. S. Eds. *Strategies of qualitative inquiry*, 2nd Edition. Thousand Oaks, CA: Sage. pp.134-164.

Stasser, G. and Dietz-Uhler, B. (2001) 'Collective Choice, Judgment, and Problem Solving', In *Blackwell Handbook of Group Psychology: Group Processes*, Hogg, M. A. and Tindale, R. S. eds. Oxford: Blackwell.

Stewart, T. R. and Lusk, C. M. (1994) 'Seven Components of Judgmental Forecasting Skill: Implications for Research and the Improvement of Forecasts', *Journal of Forecasting*, 13(7), pp. 579-599.

Stuermer, M. (2015) '150 Years of Boom and Bust What Drives Mineral Commodity Prices', *Macroeconomic Dynamics*, 22(3), pp.702-717.

Stuermer, M. (2017) 'Demand Shocks Fuel Commodity Price Booms and Busts', *Economic Letter*, 12(14), pp.1-4.

Sunstein, C. R. (2006) *Infotopia: How Many Minds Produce Knowledge*. Oxford University Press New York.

Surowiecki, J. (2004) *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business, Economies, Societies, and Nations*. New York: Doubleday.

Susemihl, F. and Hicks, R. D. (1976) *The Politics of Aristotle*. New York: Arno Press.

Tapia Cortez, C. A., Coulton, J., Sammut, C., and Saydam, S. (2018) 'Determining the chaotic behaviour of copper prices in the long-term using annual price data', *Palgrave Communications*, 4(1), pp. 1-13.

Tetlock, P. and Gardner, D. (2016) *Super-Forecasting: The Art and Science of Prediction*. London: Penguin Random House.

Thaler, R. H. and Ziemba, W. T. (1988) 'Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries', *Journal of Economic Perspectives*, 2(2), pp. 161-174.

Thaler, R. H. (2005) *Advances in Behavioral Finance (Vol. 2)*. Thaler, R. H. Ed. New York: Russell Sage Foundation.

Thaler, R. H. and De Bondt, W. (1993) *Advances in Behavioral Finance*. New York: Russell Sage Foundation.

The Economist Intelligence Unit. (2020) Global Insights & Market Intelligence [Online] Available from http://www.eiu.com/home.aspx (Accessed December 18, 2020).

Theil, H. (1961) *Economic Forecasts and Policy*. 2nd Ed. Amsterdam: North-Holland.

Theil, H. (1966) *Applied Economic Forecasting*. Chicago: Rand McNally.

Thomson Reuters GFMS. (2020) Thomson Reuters GFMS Financial Risk Solutions [Online] Available from http://financial-risk-solutions.thomsonreuters.info/GFMS (Accessed December 18, 2020).

Tilton, J. E. (2018) 'Assessing the market power of mineral commodity producers', *Mineral Economics*, 31, pp. 71-76.

Torbert, W. R. (1991) *The power of balance*. Thousand Oaks, CA: Sage.

Torbert, W. R. and Taylor, S. S. (2008) 'Action inquiry: Interviewing multiple qualities in attention for timely inquiry', In: Reason, P. and Bradbury, H. (Eds), *Handbook of action research* (2nd Edition) London: Sage. pp. 239–251.

Trueman, B. (1994) 'Analyst forecasts and herding behavior', *Review of Financial Studies*, 7, pp. 97-124.

Ury, W. (2013) 'The Five Ps of Persuasion: Roger Fisher's Approach to Influence', *Negotiation Journal*, pp. 133-140.

Waldron, J. (1995) 'The Wisdom of the Multitude: Some Reflections on Book 3, Chapter 11 of Aristotle's Politics', *Political Theory*, 23(4), pp. 563-584.

Wales, J. (2004) *Founder Letter*. Wikipedia. Available online: https://foundation.wikimedia.org/wiki/Founder_letter/Founder_letter_Sept_2004 (Accessed: 25 February 2020).

Wallis, K. F. (2014) 'Revisiting Francis Galton's Forecasting Competition', *Statistical Science*, 29(3), pp. 420-424.

Weick, K.E. (1988) 'Enacted sensemaking in crisis situations', *Journal of Management Studies*, 25(4). pp. 305–317.

Weir, J. P. (2005) 'Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM', *Journal of Strength and Conditioning Research*, 19(1), pp. 231–240.

Welch, I. (2000) 'Herding among security analysts', *Journal of Financial Economics*, 58, pp. 369-396.

Wellmer, F. W. and Scholz, R. W. (2017) 'Peak minerals What can we learn from the history of mineral economics and the cases of gold and phosphorus', *Mineral Economics*, 30, pp. 73-93.

Went, P., Jirasakuldech, B. and Emekter, R. (2009) 'Bubbles in Commodities Markets', *Social Science Research Network*, pp. 1-42. Available from http://ssrn.com/abstract=1342768 (Accessed April 6, 2018).

Wikimedia. (2017) Wikimedia: Arbitration Committee [Online] Available from https://meta.wikimedia.org/wiki/Arbitration_Committee (Accessed February 25, 2020).

Wikipedia. (2001) Wikipedia: Neutral Point of View [Online] Available from https://en.wikipedia.org/w/index.php?title=Wikipedia:Neutral_point_of_view&oldid=270453 (Accessed February 25, 2020).

Wikipedia. (2019) Wikipedia: Endoxa [Online] Available from https://en.wikipedia.org/wiki/Endoxa (Accessed July 7, 2019).

Wikipedia. (2020) Wikipedia: About [Online] Available from https://en.wikipedia.org/wiki/Wikipedia:About (Accessed February 25, 2020).

Wolcott, H. F. (1990) 'On seeking and rejecting validity in qualitative research'. In Eisner E. W. and Peshkin A. Eds., *Qualitative inquiry in education: The continuing debate,* pp. 121–152. New York: Teachers College Press.

Wolfers, J. and Zitzewitz, E. (2004) 'Prediction Markets', *Journal of Economic Perspectives*, 18(2), pp. 107-126.

Wood Mackenzie. (2020) Metals and mining: Research for Global Metals and coal markets [Online] Available from https://www.woodmac.com/research/products/metals-and-mining/ (Accessed December 18, 2020).

Yin, L., Yang, Q., and Su, Z. (2017) 'Predictability of Financialization and Co-Movement in Commodity Market What Is the Role of Technical Indicators', *Quantitative Finance,* 17(5), pp. 795-812.

Zhang, K., Nieto, A., and Kleit, A. N. (2015) 'The real option value of mining operations using mean-reverting commodity prices', Mineral Economics, 28, pp. 11-22.

Annexure A: CM panel Participant Research Project Background

Introduction

Over the last decade, the total impairments for only the top forty mining organizations were around two hundred and forty-eight billion U.S. dollars (PWC, 2020). A perennial challenge facing mining organizations is how they can reliably forecast metal prices to support their planning and improve their sustainability. From the inception of a project when cut-off grades are calculated, and feasibility studies are undertaken, forecasted metal price assumptions play a crucial role. For existing operations, annual plans to allocate cash flows to Opex, Capex, taxation, debt servicing, dividends, et cetera, are similarly based on assumptions about forecasted metal prices. Within the mining industry, the use of consensus metal price forecasts for M&A transactions is a recognized approach; however, the consensus price derivation usually occurs independently of any deliberation between the contributing forecasters.

It should be feasible for mining organizations to consider interacting through existing industry associations to establish a confidential forum to make consensus metal price forecasts that include the views of organizational members in conjunction with the inputs of outside specialist participants. Such an industry co-joint metal consensus forecasting panel could combine existing consensus forecasts provided by third parties, e.g., Consensus Economics (2020), with the insights held by mining industry experts. Within the context of publicly traded commodity markets, e.g., COMEX (2020), LBMA (2020), and LME (2020), such an industry consensus metal price forecast would be independent of the markets as the intention behind the collaboration was not aimed at trading on the forecasts. The consensus forecast would be more akin to future price expectations views regularly expressed by industry experts, either as part of their outlook guidance or as the basis for calculating their declared mineral resources and reserves.

For specific organizations, when they are making substantial decisions, considering another perspective which involves minimal additional resources and does not preclude adhering to the inhouse view, would seem prudent. As to the confidentiality of competitive insights, participation on any industry consensus panel could be structured to ensure anonymity, and the resulting consensus metal price forecast restricted to only those participating in the derivation. Given mining organizations have incurred substantial impairments because of metal price fluctuations, should all avenues not be considered to mitigate the related risk?

Background

Aristotle (Politics, 350 B.C.E) "*... many are better judges than a single man ...; for some understand one part, and some another, and among them they understand the whole* " (Everson, 1988, p.66). Condorcet formulated the Law of Averages in 1785 (Baker, 1976) in his "Jury Theorem." In 1906 Sir Galton found from an experiment conducted at an English fair that the "popular vote" could prove more accurate than most individual predictions. Surowiecki (2004), "*The Wisdom of Crowds,*" cites examples of how acting collectively can result in an outcome surpassing the capabilities of the individual group participants. Sunstein (2006), "*Infotopia: How Many Minds Produce Knowledge,*" explores the value of collective deliberation and the possibilities for open source collaboration. Page (2007), "*The Difference: How the Power of Diversity Creates Better Groups,*" explores why diversity is more important than sheer numbers in achieving better results. Tetlock & Gardner (2016), "*Super-Forecasting: The Art and Science of Prediction,*" examine the potential of collective intelligence and collaboration for obtaining the maximum value from shared insights. The mathematics of getting better forecasts through combing different backgrounds and forecasting methodologies has been widely researched, for example, Bachelor and Dua (1995) and Armstrong (2001).

The idea that through co-operating to share and deliberate insights, a benefit could accrue to those participating in the collaboration in the form of a more reliable outcome has been around since Aristotle. For many consensus forecasts, the process and aggregation mechanism are similar to the popular voting approach used by Sir Galton with either the mean or median representing the consensus estimate. Hayek (1945) described a key feature of markets as their ability to pool diverse information and, in the process, communicate that information to other market participants allowing them to make better-informed decisions. Sunstein (2006) examined how prediction markets could emulate the workings of Hayek's markets to communicate information. Several organizations offer consensus earnings estimates, with Estimize (2020) with almost three thousand indicators covered, and with over one hundred thousand estimators, perhaps the most diverse consensus estimation service. Of the eighty-five economic indicators Estimize covers, commodity prices are not included, opening the possibility for the mining industry collectively to explore the opportunity to tap into the collective industry insights.

Research project

You have been invited to take part in a doctoral research project, on a bespoke website. An underlying premise in extending the invitation to participate in the research project is that you either prepare or use metal price forecasts in your working environment or you follow commodities and have insights into emerging market dynamics. Participants in the research project includes industry experts from mining organizations, commodity trading entities, financial institutions, and market research organizations. The research project aims to specifically co-opt mining industry professionals to participate jointly in making metal price forecasts alongside their industry peers. The expected outcome of the research project is to establish if a consensus approach for forecasting expected future metal prices can prove more reliable for industry stakeholders. The interaction amongst participants is on the web-based consensus research panel - http://www.comsensusmetals.org.

The research project's focus is on forecasted metal prices for use by industry experts within their professional environment. Participation is not expected to require disclosing any confidential information or to serve as a basis for any trading or hedging decisions. Participation will be anonymous using pseudonyms or usernames on the website, to avoid revealing the identity of any participants. Should any participant choose to disclose their identity on the consensus research panel, it will not alter the understanding that participating is only in a private capacity. Consequently, any views expressed on the consensus research panel will be regarded solely as those of the participant, and not those of any associated organization. Regardless of any disclosures made on the consensus research panel, in the any research reports, all references to participation will be anonymized and codified, e.g., Industry Expert A, B, C et cetera.

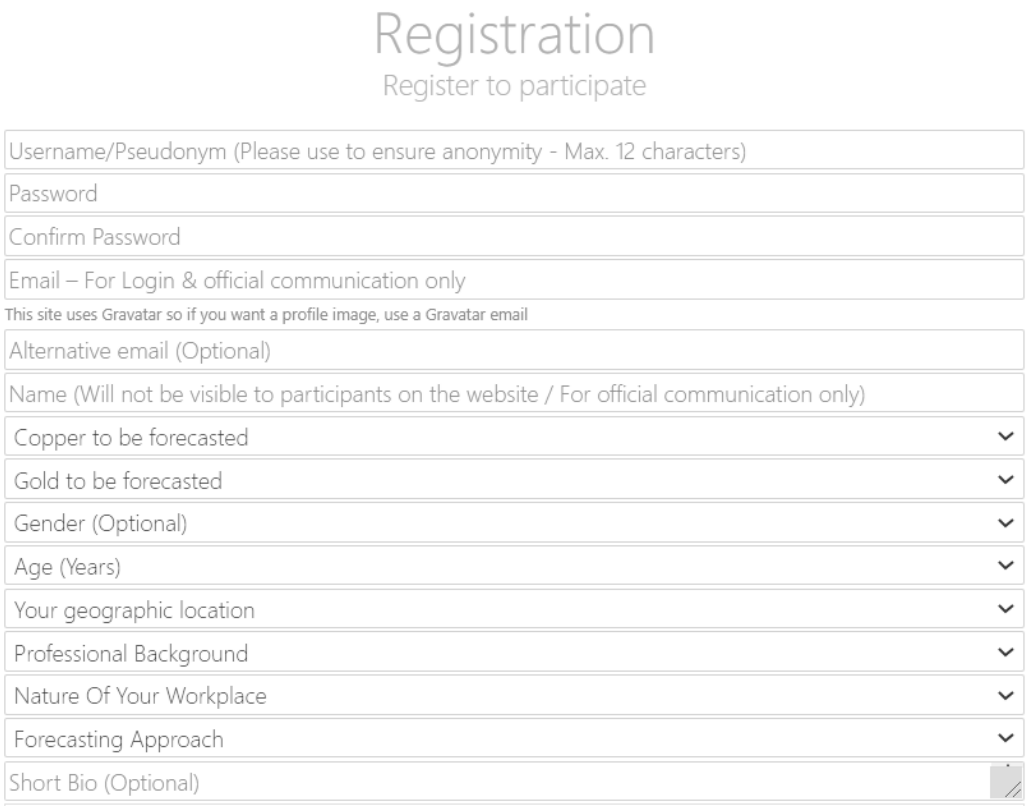
Taking part in the research project is not anticipated to cause participants any adverse consequences or emotions. The potential psychological effects of involvement in the research projects for participants should be similar to their experience in performing their usual professional responsibilities of either making or using metal price forecasts. No financial compensation, reimbursement, or reward will be paid to participants for their participation in the research project. Participants will have the opportunity to receive an electronic copy of the final research report, should they so choose. Participants will be free to stop participating at any time should they want and have all their past contributions deleted upon request.

Panel participants will be expected to contribute metal price forecasts regularly over four months (September – December 2020), and participation is likely to require approximately thirty minutes per month. In making their metal price forecasts, participants will be asked to provide a brief (maximum 300 characters) indication of the basis for their forecasted metal prices. Participants are encouraged to discuss their forecasts and those submitted by other participants, as a prompt to surfacing shared insights, aimed at improving the reliability of the consensus forecasted metal prices. Should any participant choose to post any link to relevant research on the consensus research panel website, it is on the understanding that the participant has the necessary right to share the research. As the consensus metal price forecasts are the joint effort of all participants, access to the contibuting forecasts of all participants and the evolving consensus metal price forecasts will continually be available to everyone participating throughout the research project.

Conclusion

The research project is about the collective interaction of participants through iterative cycles of forecasting, sharing insights, and aggregating the participants' expectations in open view of all participants, to make consensus metal price forecasts. The research project aims to explore the possible benefits of using a diverse consensus panel for forecasting metal prices in the natural resource industry, and through specifically encouraging interactions to share and deliberate insights, test the reliability of the approach for forecasting metal prices. If the research project should prove positive, it could assist in structuring longer term consensus forecasts by industry stakeholders, through the auspices of an industry association in future. Ultimately, should the idea of tapping into the "wisdom **IN** the industry" be successful, it may help mining organizations reduce their uncertainty around future metal prices and, consequently, impairments.

Annexure B: CM panel Participant Registration And Consent Form



PARTICIPANT CONSENT QUESTIONS	
Consent Questions (All questions need to be answered "Yes" in order to register and participate in the research project)	Yes/No
I confirm part of my skill set is preparing, using, or tracking metal price forecasts. It is an expectation to participate in the research project that you have the necessary skill set. Please check the appropriate box alongside. If you do not meet the skill requirement, please decline the opportunity to participate in the research project.	○ Yes ○ No
I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason, and without my rights being affected. In addition, should I not wish to answer any questions or submit any forecasts, I am free to decline.	○ Yes ○ No
I understand that, under the UK Data Protection Act, I can at any time ask for access to the information I provided, and I can also request the destruction of that information if I wish, and withdraw from the research project.	○ Yes ○ No
I understand that confidentiality and anonymity will be maintained, and it will not be possible to identify me in any publications.	○ Yes ○ No
I agree for the data collected from me to be used in future research and understand that any such use of identifiable data would be reviewed and approved by a university research ethics committee.	○ Yes ○ No
I understand and agree that my participation will be recorded/captured. I am aware of and consent to your use of this recorded information for the purpose of data analysis.	○ Yes ○ No
I understand that my responses will be kept strictly confidential. I give permission for members of the research team to have access to my anonymized responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research project.	○ Yes ○ No
I agree that the data collected from me may be used in other relevant future research on the same confidential basis as specified for this research project.	○ Yes ○ No
I understand the purpose of the research project is to investigate metal price forecasting and specifically acknowledge the intention of the research is not to provide any trading or financial recommendation on future metal prices. Any inferences drawn or made about future metal prices because of participating in the research are solely mine, and I take full responsibility for any actions I may take stemming from such expectations.	○ Yes ○ No
I agree to take part in the the research project.	○ Yes ○ No

Submit

Additional Information

Should you have any further questions, please read the Additional Information webpage. If you still have any questions or concerns thereafter, please use the contact details provided to contact us for further information or clarification.

Availability of dissertation

When the research project is completed and the final dissertation is published, a copy will be made available on the website, should any participants wish to have access to a copy.

Annexure C: CM panel Participant's Additional Research Information

Research Addition Information

Title of Research Project:

How will a consensus panel approach for forecasting copper and gold prices, which specifically includes mining industry experts' participation, improve the reliability of forecast price estimates for planning purposes?

Short Title:

How can consensus copper and gold price forecasting be more effectively used in the mining industry for planning?

Researcher: John Lamprecht
University Email:
Telephone:

Invitation

You are being invited to take part in a research project. Before you decide whether to participate, it is important for you to understand why the research is taking place and what it will involve. Please take the time to read the following information carefully and feel free to ask us if you would like more information or if there is anything that you do not understand. We would like to stress that you do not have to accept the invitation and should only agree to take part if you want to. An underlying premise in extending the invitation to participate in the research project is that you either prepare or use metal price forecasts in your working environment or you follow commodities and have insights into emerging market dynamics. If this assumption is not correct, please indicate "No" on the Registration Page and decline the opportunity to participate in the research project.

What is the purpose of the research project?

A perennial problem facing mining organizations is how they can more reliably forecast metal prices to consistently support their planning and improve their sustainability. The intention of the research project aims to explore if co-opting mining industry professionals to participate in in making metal price forecasts alongside their industry peers, and jointly surface insights about expected metal prices to reduce the price uncertainty for all participants involved in the research project."

Why have I been chosen to take part?

Participants in the research project includes industry experts from mining organizations, commodity trading entities, financial institutions, and market research organizations. You have been chosen because of your experience within or to the natural resource industry. On the basis that the nature of the research project closely mirrors the participant's professional responsibilities in making or using metal price forecasts, it is expected participation in the research panel will require minimal additional time from participants (approximately 30 mintutes per month). An assumption is that amongst the actions you regularly undertake is preparing or using metal price forecasts, should this not be part of your experiences, please indicate "No" on the Registration Page and decline participation in the research project."

Do I have to take part?

It is entirely up to you to decide whether you want to participate in the research project. If you do choose to proceed, you should indicate your consent on the Registration Page. You are free to withdraw at any time from the research project without providing a reason or risk incurring any repercussions. Participation will be anonymous using a pseudonyms or usernames you choose so as not to reveal your identity. Should you however choose to disclose your identity on the consensus panel, it will be on the understanding that you are specifically participating in your personal capacity and the views expressed are your own and not those of any organization.

What will happen if I take part?

You will be asked to participate in a metal price forecasts and discussion concerning regarding your expectations of metal prices over the coming months. In making such predictions you will be asked to provide a brief (maximum 300 characters) indication of the basis for your forecasts. Panel participants will be permitted and encouraged to discuss all forecasts and views submitted. Should you want to post any relevant metal price forecast research on the consensus panel this will be permitted, but you should ensure you have the right and necessary authority to share the document. In the final research report, all consensus panel participation references will be anonymized, and codified, e.g., Industry Expert A, B, C et cetera. There are no other commitments or restrictions associated with participating. In compliance with the University of Liverpool's research data management policy, the Research project data will be stored for five years. Under freedom of information legislation, you are entitled to request access the information you have provided at any time.

Expenses and/or compensation

Participants should not incur any additional expenses as a result of taking part in this research project. No compensation will be paid to panel participants or any incidental expenses refunded.

Are there any risks in taking part?

Taking part in this research project is not anticipated to cause you any adverse consequences or emotions. The potential physical or psychological effects of involvement in the research projects for participants should be the same as you experience in performing your usual professional task of making or using metal price forecasts. However, should you experience any discomfort from taking part in this research project, please contact the researcher or supervisor to address the reason immediately.
Researcher: John Lamprecht Email: John.Lamprecht@online.liverpool.ac.uk
Supervisor: Jim Hanly Email: Jim.Hanly@online.liverpool.ac.uk

Are there any benefits in taking part?

While there are not necessarily immediate benefits for those participating in the research project, it is hoped that this research project will have a beneficial impact on the reliability of forecasted metal prices used by industry experts in the natural resource sector. As the research project is an iterative cycle of forecasting, deliberating and reforecasting metal price estimates, with the ongoing results continuously shared with all panel participants, it is hoped the research project will help inform the professional practice of the participants in future. The research project could also serve as a precursor to a similar project by an industry association on behalf of its members, including mining organizations, in future.

If you have any concerns, complaints or if there are any problems or questions you feel have not been adequately addressed?

If you are unhappy, or if you have any complaints about the research project in the first instance, please feel free to let us know by contacting the researcher or supervisor, and we will try to help. If you remain unhappy or have a complaint which you feel you cannot come to us with, then, you should contact:
Research Governance Officer at Email: ethics@liv.ac.uk, or
University of Liverpool Email: Liverpoolethics@ohecampus.com.
When contacting the Research Governance Officer or Liverpool Research Participant Advocate, please provide details of the name or description of the research project (so that it can be identified), the researcher(s) involved, and the particulars of the complaint you wish to make.
For participants from the United States of America, the following contact details can be utilized:
USA number +1-612-312-1210 or email address: Liverpoolethics@ohecampus.com

Will my participation be kept confidential?

All the information that we collect about you during the research project will be kept strictly confidential. You will not be able to be identified or identifiable in any reports or publications. Your institution will also not be identified or identifiable. Any data collected about you during discussions will be stored online in a password-protected form. Data collected may be shared in an anonymised form to allow reuse by other research teams. As stated in point 12 below, you may terminate your participation at any time and stipulate how your past participation is to be handled. As outline above, all references in the research report will be anonymised to prevent any individuals or their institutions from being identified.

What will happen to the results of the research project?

The results of the research project will be presented in a DBA thesis, and as per University of Liverpool's research data policy, the data produced during this research will be made openly available to the broader academic community. You or your organization will not be identified in any report or publication. If you wish to receive an electronic copy of any reports resulting from the research, please indicate so on the Registration Page and you will be added to the circulation list.

What will happen if I want to stop taking part?

You can withdraw from participating in this research project at any time, without explanation. Results up to the period of withdrawal may be used if you are happy for this to be done. Otherwise, you may request that their destruction and no further use will be made of them."

Who can I contact if I have further questions?

For any further question, please contact the researcher or the research supervisor.
Researcher: John Lamprecht Email: John.Lamprecht@online.liverpool.ac.uk
Supervisor: Dr. Jim Hanly Email: Jim.Hanly@online.liverpool.ac.uk

Thank you for taking part in this research project.

Annexure D: PwC Top 40 Mines Report 2011-2020

PWC Top 40 Mines Report 2011-2020 ($ billion)	Total / Average	2020	2019	2018	2017	2016	2015	2014	2013	2012	2011
Bloomberg Commodity Index (Year end - December)	**100.0**	78.1	80.9	76.7	88.2	87.5	78.6	104.3	125.8	139.1	140.7
10 Year CAGR / Annual percentage change	**(6.3%)**	(3.5%)	5.4%	(13.0%)	0.7%	11.4%	(24.7%)	(17.0%)	(9.6%)	(1.1%)	(13.4%)
Bloomberg Commodity Index (Annual average)	**105.5**	69.9	79.4	86.2	84.9	83.0	94.5	126.7	131.0	141.6	158.0
10 Year CAGR / Annual percentage change	**(8.7%)**	(12.0%)	(7.8%)	1.5%	2.2%	(12.1%)	(25.4%)	(3.3%)	(7.5%)	(10.3%)	15.7%
Aggregate market capitalisation	**944**	**1,467**	**898**	**757**	**926**	**714**	**494**	**791**	**958**	**1,234**	**1,202**
10 Year CAGR / Annual percentage change	**2.2%**	63.4%	18.6%	(18.3%)	29.7%	44.5%	(37.5%)	(17.4%)	(22.4%)	2.7%	(25.1%)
Aggregated income statement											
Revenue	**6,522**	656	692	683	600	496	539	690	719	731	716
Operating expenses	**(4,941)**	(482)	(524)	(518)	(454)	(390)	(448)	(531)	(554)	(553)	(487)
EBITDA	**1,581**	**174**	**168**	**165**	**146**	**106**	**91**	**159**	**165**	**178**	**229**
Impairment charges	**(258)**	(11)	(14)	(12)	(4)	(19)	(53)	(27)	(57)	(45)	(16)
Amortisation, depreciation and impairment	**(424)**	(50)	(50)	(47)	(41)	(44)	(42)	(48)	(42)	(34)	(26)
PBIT	**899**	**113**	**104**	**106**	**101**	**43**	**(4)**	**84**	**66**	**99**	**187**
Net finance cost	**(119)**	(10)	(14)	(13)	(11)	(9)	(19)	(15)	(16)	(6)	(6)
PBT	**780**	**103**	**90**	**93**	**90**	**34**	**(23)**	**69**	**50**	**93**	**181**
Income tax expense	**(263)**	(32)	(29)	(27)	(29)	(15)	(4)	(24)	(30)	(25)	(48)
Net profit	**517**	**71**	**61**	**66**	**61**	**19**	**(27)**	**45**	**20**	**68**	**133**
Adjusted net profit (Excl. Impairment)	**774**	**82**	**75**	**78**	**64**	**38**	**26**	**72**	**77**	**113**	**149**
Impairment charge % of NP (Excl.Impairment)	**33.3%**	13.4%	18.7%	15.4%	6.3%	50.0%	203.8%	37.5%	74.0%	39.8%	10.7%
Year on year increase/ (decrease) in revenue	**(1.0%)**	(5.2%)	1.3%	13.8%	21.0%	(8.0%)	(22.0%)	(4.0%)	(2.0%)	2.0%	65.0%
Year on year increase/ (decrease) in EBITDA	**(3.0%)**	3.6%	1.8%	13.0%	37.7%	16.0%	(43.0%)	(4.0%)	(7.0%)	(22.0%)	21.0%
Year on year increase/ (decrease) in net profit	**(6.7%)**	16.4%	(7.6%)	8.2%	221.1%	(170.0%)	(160.0%)	125.0%	(71.0%)	(49.0%)	21.0%
EBITDA margin	**24.2%**	26.5%	24.3%	24.2%	24.3%	21.4%	16.9%	23.0%	22.9%	24.4%	32.0%
Aggregated cash flow statement											
Operating activities	**1,268**	142	130	134	119	89	92	127	124	137	174
Investing activities	**(872)**	(56)	(69)	(63)	(45)	(40)	(69)	(93)	(125)	(169)	(142)
Financing activities	**(366)**	(51)	(66)	(70)	(63)	(44)	(31)	(31)	(3)	21	(28)
Dividends paid	**(367)**	(37)	(55)	(43)	(36)	(16)	(28)	(40)	(41)	(38)	(33)
Share buy backs	**(82)**	(1)	(7)	(15)	(7)	(4)	(7)	(6)	(4)	(5)	(26)
Free cash flow	**466**	**81**	**69**	**77**	**71**	**40**	**23**	**24**	**(6)**	**11**	**76**
Impairment charge % of Investing Activities	**29.6%**	19.6%	20.3%	19.0%	8.7%	47.5%	76.8%	29.0%	45.6%	26.6%	11.3%
Aggregated balance sheet											
Cash	**105**	123	88	101	102	86	82	83	168	104	113
Property, plant and equipment	**653**	653	649	610	663	616	579	745	712	701	601
Total assets	**1,149**	1,163	1,139	1,080	1,129	1,063	1,047	1,231	1,256	1,245	1,139
Total liabilities	**(571)**	(588)	(576)	(540)	(573)	(563)	(569)	(630)	(624)	(563)	(482)
Total equity	**578**	**575**	**563**	**540**	**556**	**500**	**478**	**601**	**632**	**682**	**657**
Impairment charge % of Fixed assets	**3.9%**	1.7%	2.2%	2.0%	0.6%	3.1%	9.2%	3.6%	8.0%	6.4%	2.7%

Annexure E: LBMA-PMS 2001

Forecast 2001

Golden Bars & the Three Bears

In the end the race was too close to call. So we sent everything to Tallahassee. (Just kidding). No – we went to the Supremes (the gowned group from Motown, not D.C.), who chorused, "Some day, we'll be toge-eh-eh-ther."

And so we were, in the persons of Philip Klapwijk of GFMS and PeterWard of Lehman Bros., who harmonised on the annual average of $280 (actual: $279.102). But with a tie, it takes the high and low to win the gold – and on that basis, the prize (a one-ounce Johnson Matthey bar) goes to Philip, whose $255 and $315 better matched the tenor of the times, striking a chord with the actuals of $262.80 and $312.70 versus Peter's more bear-like baritone at $235 and $310.

Well, there's always next year, for which many of our forecasters have grown even more bearish after 2000's performance, when gold did its Groundhog Day thing – only poking its head above $300 for a two-week stretch during February. For 2001, the average of averages slid from $295.942 to $275.91. Silver was equally tarnished – last year, $5.546, this year, $4.768.

Speaking of which, in Silver News, the Undimpled Chad Award goes to Andy "Bear Minimum" Smith of Mitsui. At $5.00, he raised penny pinching to screamingly high art, missing the actual average ($4.9525) by a mere $0.0475. After having nailed the gold last year, Bullseye's in a hairy position for such a world-renowned bear – doubly long gold.

We end with a quote from Shakespeare (Henry IV) contributed by Hennie Rijnbeek of Rabobank: "Is it not strange that desire should so many years outlive performance?"

Meaning? Your guess is as good as ours, or maybe better. Find out right here, next year.

Gold

	Name	Company	City	High	Low	Average
1	Arnold, Ted	Prudential-Bache Int. Ltd.	London	$283	$250	$268
2	Bianchini, Egizio	Nesbitt Burns	Toronto	$315	$260	$280
3	Christian, Jeffrey	CPM Group	New York	$314	$255	$294
4	Crisp, Kevin	Credit Suisse First Boston	London	$300	$250	$275
5	Cross, Jessica	Mutual Metals	London	$310	$250	$283
6	Fewings, Martin	NM Rothschild & Sons	London	$300	$250	$280
7	Goode, Keith	Bell Securities	Sydney	$290	$270	$277
8	Hawkes, Neil	CRU International	London	$290	$250	$265
9	Klapwijk, Philip	GFMS	London	$290	$252	$271
10	Lasserre, Frédéric	SG	Paris	$300	$250	$270
11	Lester, Howard	Rue Stern & Co	New York	$320	$260	$289
12	Malladieu, David	Scotia Capital Inc.	Toronto	$300	$250	$286
13	McConvey, Daniel	Goldman Sachs	New York	$295	$250	$275
14	Murenbeeld, Martin	M. Murenbeeld & Assoc.	Vancouver	$324	$256	$297
15	Naqvi, Kamal	Macquarie Bank	London	$300	$260	$275
16	Norrish, Kevin	Barclays Capital	London	$295	$250	$267
17	O'Connell, Rhona	Canaccord Capital (Europe)	London	$325	$265	$285
18	Pastrutti, Frédéric	MKS Finance SA	Geneva	$320	$260	$282
19	Reade, John	UBSWarburg	London	$310	$260	$282.5
20	Rhodes, Jeffrey	Standard Bank London	Dubai	$295.65	$253.50	$272.40
21	Rijnbeek, Hennie	Rabobank International	London	$295	$255	$268
22	Smith, Andy	Mitsui & Co.	London	$290	$210	$250
23	Takai, Bob	Sumitomo Corp	Tokyo	$285	$255	$265
24	Ward, Peter	Lehman Bros.	New York	$295	$240	$265
	AVERAGES			$301.74	$251.31	$275.91

Silver

	Name	Company	City	High	Low	Average
1	Arnold, Ted	Prudential-Bache Int. Ltd.	London	$4.85	$4.45	$4.67
2	Butler, Steven	Nesbitt Burns	Toronto	$5.15	$4.00	$5.00
3	Christian, Jeffrey	CPM Group	New York	$5.40	$4.55	$5.67
4	Crisp, Kevin	Credit Suisse First Boston	London	$5.00	$4.20	$4.70
5	Cross, Jessica	Mutual Metals	London	$5.10	$3.60	$4.85
6	Fewings, Martin	NM Rothschild & Sons	London	$4.90	$4.40	$4.80
7	Goode, Keith	Bell Securities	Sydney	***	***	$5.05
8	Hawkes, Neil	CRU International	London	$4.95	$4.40	$4.80
9	Klapwijk, Philip	GFMS	London	$4.98	$4.38	$4.72
10	Lasserre, Frédéric	SG	Paris	$5.50	$4.00	$4.40
11	Lester, Howard	Rue Stern & Co.	New York	$5.60	$4.40	$5.30
12	Naqvi, Kamal	Macquarie Bank	London	$5.25	$4.25	$4.75
13	Norrish, Kevin	Barclays Capital	London	$5.20	$4.40	$4.85
14	O'Connell, Rhona	Canaccord Capital (Europe)	London	$6.00	$4.50	$5.25
15	Pastrutti, Frédéric	MKS Finance SA	Geneva	$5.50	$3.90	$4.60
16	Reade, John	UBSWarburg	London	$5.20	$4.40	$4.85
17	Rhodes, Jeffrey	Standard Bank London	Dubai	$5.17	$4.27	$4.76
18	Rijnbeek, Hennie	Rabobank International	London	$5.20	$4.40	$4.65
19	Smith, Andy	Mitsui & Co.	London	$5.25	$3.50	$4.25
20	Takai, Bob	Sumitomo Corp	Tokyo	$5.00	$4.00	$4.55
21	Ward, Peter	Lehman Bros.	New York	$5.10	$4.40	$4.65
	AVERAGES			$5.260	$4.210	$4.768

Ted Arnold

Prudential Bache International Ltd, London

● Gold

Range: $250 – $283
Average: $268

It is difficult to see gold able to achieve or maintain high levels during 2001. The Washington Accord is likely to begin to fray a bit over the course of the year. The largest demand-driven areas – India, Asia – have become increasingly price sensitive.

● Silver

Range: $4.45 – $4.85
Average: $4.67

The same principles apply to silver. There is an enormous amount of it available, and we don't see any reasons for cutbacks in secondary supply. Nor do we see any reasons for sharp increases in offtake. The large demand centres won't necessarily increase their purchases at lower price levels – they have enough already, and they won't buy extra just because it's cheap.

Egizio Bianchini

Nesbitt Burns, Toronto

● Gold

Range: $260 - $315
Average: $US 280

With an economic slowdown beginning to engulf the US, its likely that interest rates will fall markedly and perhaps precipitously. If the rules of economics prevail, the US dollar should decline against other major currencies including gold. Additionally, this reflationary act should also work in favour of the yellow metal. However, we would want to make it clear that the magnitude of the US dollar decline will have to be sufficient to begin to seriously erode the almost religious confidence that investors world wide have in all things American, including the currency. Perhaps a 1.15 to 1.30 dollar/euro exchange would do the trick. Also weighing on bullion prices is the current lack of investment demand. The solution to this is simple. A rising gold price will attract investors, especially if a crisis of confidence were to develop.

Steven Butler

Nesbitt Burns, Toronto

● Silver

Range: $4.00 - $5.15
Average: $5.00

Our supply-demand outlook for silver is based on relatively flat supply and a modest increase in demand through 2001. We estimate a primary supply deficit (before official sector sales and hedging) of 167 million ounces in 2000 and 177 million ounces in 2001, up from 156 million ounces in 1999 (as per World Silver Survey 2000). While this should provide a fundamental basis for improved silver prices, we believe the metal will continue to be pegged with our cautious outlook for the gold price ($280 per ounce in 2001). Based on a gold:silver ratio of 56:1, we estimate an average silver price of $5.00 per ounce in 2001.

Jeffrey Christian

CPM Group, NewYork

● Gold

Range: $255 - $314
Average: $294

The gold price is expected to remain under the influence of two main groups of participants. One group are the technically oriented short-term proprietary traders. The other group consists of investors. The former group has been the single most important factor driving gold prices down in recent years. While we witnessed a sharp reduction in their short-selling activity in 2000, they remain the largest volume participants in the gold market, and continue to have the ability to drive prices down if they feel that investors are not likely to be aggressive buyers. Investors for their part have remained apart from gold, in part reflecting a reluctance to leave the U.S. dollar, U.S. debt, and equity markets. The equity market worldwide performed poorly in 2000, and may continue to flounder in 2001. This is helping to start investors thinking about gold as an alternative asset, although there has not been a great deal of actual buying up to the end of the year 2000. That is because investors remain interested in U.S. dollar denominated debt. If this romance fades in the first half of 2001, investors may convert their idle glances at gold into substantive purchases. Given a relatively tight and increasingly illiquid gold market, any increase in real investment demand should be reflected in measurable increases in gold prices.

● Silver

Range: $4.55 - $6.40
Average: $5.67

The silver market meanwhile remains on hold. Most market participants view the continual reduction in unreported stocks in Europe as a sign of gathering tightness. The lack of physical silver is expected to be reflected in a sharp increase in silver prices at some point. This point may be reached in 2001, although there may be a period of continued weak prices first earlier in the year.

Kevin Crisp

Credit Suisse First Boston London

● Gold

Range: $250 – $300
Average: $275

Key macro issues for gold in 2001 will be the currency markets and the falling US dollar and the slow down in global economic activity. The greater stability in gold supply (from all sources) witnessed over the past year will continue, leaving the demand-side as the principal driver of gold market sentiment. A broadly weaker US dollar will provide some support to gold demand but currency moves alone may not be enough to offset the negative impact on consumer spending and jewellery buying as economic growth slows and consumer sentiment deteriorates. Unless dollar gold rallies sharply (which would likely trigger increased supply), a softer US dollar will expose many gold producers to lower gold prices in domestic currency terms. Longer term this will further curtail exploration and development and short term may lead to some renewed interest in price risk management.

● Silver

Range: $4.20 – $5.00
Average: $4.70

The softer tone to the silver market and the contraction in silver trading activity over the past year (as witnessed by the LBMA's monthly clearing data) is unlikely to show much improvement before the second half of the year. Silver's strong credentials as an industrial metal and the weaker global economic picture offer little scope for any demand-side growth in 2001. The photographic sector is also exhibiting a softer tone. At the same time the availability of silver supply to the market shows little sign of tightening and significantly, China appears set to remain a source of metal following the granting of new export quotas for the coming year.

Jessica Cross

Virtual Metals Research & Consulting, London

● Gold

Range: $250 – $310
Average: $283

Critical to the metals' price performance will be the fate of local currencies against the US Dollar. On the supply side, hedging decisions will be dictated by the relative value of the Australian Dollar and the South African Rand; any weakening could trigger further forward selling from those producers. On the demand side the performance and health of the Indian Rupee will influence metals flows into and out of the Subcontinent. Should the market see the continued sales from Non-Washington Agreement countries as characterised 2000, the supply/demand balance will favour the lower end of the trading range.

● Silver

Range: $3.60 – $5.10
Average: $4.05

If 2000 – especially the closing months – is anything to go by, we must expect a very poor year for silver. With the investor conspicuous in his absence, the price is likely to continue to sag. China's decision late last year to increase the number of export licences from 2 to between 8-10, with explicit intention of reducing local surpluses, does not augur well for the international market.

Martin Fewings

NM Rothschild & Sons, London

● Gold

Range: $250-$300
Average: $280

Having languished in the bottom half of the $US262-295/oz range it traded within for much of last year, we see little change for gold over the first few months of 2001. Given its recent break from the US dollar, gold's fundamentals are likely to hold more sway near-term – mediocre demand and ample supply will offer little stimulus. Moreover, the two major influences on gold – the activities of central banks and producers – have become more predictable for the present. The resultant reduction in (near-term) uncertainty has also led to a significant contraction in liquidity. However, in the longer term, the US dollar should come into play again. A forecast depreciation of the US dollar (in line with lower

interest rates) should encourage higher prices from the end of the second quarter.

● Silver

Range: $4.40-$4.90

Average: $4.80

Given the sharp decline in price, traded volumes and market interest, it is becoming increasingly difficult to argue a positive case for silver. While a forecast weakening in the US dollar should benefit prices, the mystery of above ground stocks remains the key to silver's future. Further sizeable disposals from Chinese stocks can be expected, with the threat of even larger sales acting to constrain any price rises. At least equally important will be the continuation of disinvestment from the private sector. Indeed just how private holders of silver inventories react to the languishing market is the largest uncertainty facing silver.

Keith Goode

Bell Securities, Sydney

● Gold

Range: $260 - $290

Average: $277

In the past year, there has been an incredibly close correlation between the gold price and the Euro. So far with the recent appreciation of the Euro, we have seen an improvement in the gold price too. The general expectation is that the Euro will improve against the US$ mainly from January, when it becomes a physical currency and benefits from the grey market along with other influences, although it has surprised many by trading so close to US$0.92 already. Gold is also gaining short-term from weak equity markets. However, such switches have tended to be short-lived, sometimes lasting only a few days. At heart, it really comes down to what the speculators selling gold short elect to do, and how many times they are forced to scramble for cover.

So gold should be capable of recovering back to the price at which it entered the Euro about two years ago – US$290/oz – and I think that there are enough factors out there to limit the downside. It could get above US$290/oz and retest higher levels, however, we can talk about that once it manages to stay above US$290/oz.

● Silver

Range: $4.50 – 5.25

Average: $5.05

As for silver, I think that the traditional links between the precious metals have been broken in the past year, platinum rising dramatically and silver falling (in both cases partly due to China) relative to gold. Possibly silver can average about US$5.05/oz and maintain a ratio in the order of 55 x relative to gold

Neil Hawkes

CRU International, London

● Gold

Range: $250-$290

Average: $265

We expect 2001 to be another downbeat year for gold. There is some scope for brief rallies on the back of a weaker US picture, with a weaker US$ helping demand and hindering producer selling. However, we see no reason for any investor flight to gold to be anything other than brief. Gold's fundamental picture remains bleak, with net central bank sales continuing apace, mine output yet to fall sharply and producer hedging more subdued. But perhaps most worrying is lacklustre demand. In addition to investors' waning interest, physical offtake proved to be disappointing in 2000, with India the notable laggard. Faltering US demand this year will only add to the slack picture in Asia, although a weaker US$ should provide some support, particularly from European buyers. Indeed, it is our assumption that this buying support from Europe (and Asia) will prevent prices from plunging to new (20-year) depths.

● Silver

Range: $4.40-4.95

Average: $4.80

The picture for silver looks gloomy. Prices are likely to remain under downward pressure for the moment from a combination of too much supply, both from stocks (e.g. China) and fresh output (e.g. Mexico), dwindling investor interest and jaded Indian demand. Despite all this, we believe the downside is limited. Such low prices should spur some price-sensitive Asian demand, providing a floor not too far below current values. Moreover, the flood out of China may start to ebb, after the initial surge, as stocks slide. Also, silver may be saved from further humiliation by being sidelined by investors, no longer interested in any participation, on the short as well as the long side.

Philip Klapwijk

Gold Fields Mineral Services London

● Gold

Range: $252 – $290

Average: $271

Gold's prospects are difficult to assess because of changing economic circumstances and a probable further decline in the dollar. Slower world economic growth and a weaker American dollar will tend to pull the (dollar) gold price in opposite directions. A continued dollar decline should support the price, although unlikely to push gold beyond $290. Over the long term, the negative correlation between the dollar's exchange rate and the gold price is not particularly strong.

If dollar weakness is a sign of impending recession in the US, the gold price may come under pressure – any severe slowdown in the American economy would result in lower jewellery consumption not just in the US but also in those parts of the world that are particularly dependent on their merchandise exports to the United States (e.g. East Asia). Much will depend upon the extent to which America's economic difficulties are exported elsewhere. In itself, the US does not loom large enough in global demand (13% of 2000 global jewellery consumption) to present that much of a danger to the price.

We expect movements in gold's supply/demand variables to have a somewhat larger bearing on the price this year. Producer hedging should be more of a factor, although we do not envisage a major rise in positions in the first half. A historically high level of official sector sales will continue to restrain gold's upside. Similarly, it is doubtful whether any spike in the price could be sustained in the face of the additional supply from scrap and investor dishoarding that would result. (Significantly so far there has been no safe haven buying of bullion by funds or private individuals seeking shelter from tumbling stockmarkets). We do not foresee any decline in global mine production this year.

On the downside, our base case assumes that the US economy will make a soft landing and that the rest of the world will only be moderately affected. (For example, in our recent Update 2 publication we forecast first half jewellery demand to be flat rather than down year-on-year). This, together with a softer dollar should prevent gold from testing new lows.

● Silver

Range: $4.30 - $4.98

Average: $4.72

The silver price has recently been under tremendous pressure from inventory liquidations. A key issue is whether these sales have largely finished or will continue notwithstanding a silver price far closer to $4.50 than $5.00. In our view, the amount of metal that is available from private investor and Chinese government stocks will diminish at these lower price levels. However, the continued existence of substantial inventories will limit silver's upside, even if the metal has lately been a little "oversold". Furthermore, the support provided by strong physical demand in 2000 will not be as pronounced this year. Already, there are signs raw material demand from the industrial and photographic sectors has begun to ease. Of course, much will depend upon the extent to which price sensitive consumers, above all in India, absorb the slack in the market that will probably be created by lower fabrication demand in industrial countries and "excess" supply from inventories. An important factor here will be the value of the Indian rupee, which ought to benefit in 2001 from a weaker dollar and lower oil prices. Overall, though, as largely an industrial metal, it is difficult to see silver recovering greatly in what may turn out to be a year of lower growth in the world economy.

Frédéric Lasserre

SG Paris

● Gold

Range: $250 - $300

Average: $270

For the gold market, 2001 should bear a strong resemblance to 2000. The market will remain dominated by the fear of new central banks sales (mainly from Latin America) while only a few purchases (mainly from eastern Europe) would reassure it. Nevertheless, the three main factors that have influenced gold prices in 2000 (namely the US dollar, the Dow Jones and oil prices, in that particular order) should be even more helpful in 2001. The weakening of the USD, due to the clear slowdown of the US economy, should stimulate physical demand and reduce producers' hedging incentive. Gold leasing rates will remain low and the contango will be reduced further by the extension of the lending horizon by the major central banks. The only event that could boost gold prices would be a producers' MoU (Memorandum of Understanding) to limit or interrupt hedging programmes. But the probability of seeing this sort of producers' "Washington Agreement" remains fairly weak due to the diversity of individual situations and the divergence of interests.

● Silver

Range: $4.00 - $5.50

Average: $4.40

The silver market has well and truly given up hope that its official "deficit" position, which lasted for years, should one day underpin prices. The existence of massive hidden stocks explains why the silver price is on a downtrend while the market is supposed to be in deficit for ten years. In 2001, China will probably take hedge funds role as responsible for all suspicious changes in prices or leasing rates. Silver physical demand is currently losing its closest ally, the

photographic industry, as digital technology is now in position to take a sizeable market share. On the investors' side, silver has been definitely dethroned by PGMs (especially palladium) as metals funniest playground.

Howard Levine

Bear Stearns Commodity Sales, NewYork

● Gold

Range: $260 – $320

Average $289

Freshman Economics teaches us that an extended decline in price will tend to reduce supply and increase demand. Belatedly, in 2001, we should begin to see evidence of this remedial process in gold's price performance.

Over the past five years, the sharp decline in dollar-denominated gold prices has failed to either crimp supply or invigorate demand. Primary production levels have been supported by more adequate local currency returns in many producing areas and by producers' successful forward hedge programs. Meanwhile, investment offtake of gold for portfolio diversification purposes has been muted by enthusiastic demand for equities and a comfortingly strong US dollar. And, of course, central banks' large reductions in their monetary bullion holdings have not been helpful.

The deflation of the "so-called bubble" in high tech equities and the evident peaking of the dollar against the euro, however, should forestall further slippage in gold prices. Investment demand for gold, though being dissuaded at the moment by recessionary anxiety, should improve as portfolio managers adopt a more defensive stance. Primary supply should be undercut as advantageous forward hedge transactions roll off and as local currency values rebound against the dollar. The wind-down of present central bank gold sales programs also is in sight.

● Silver

Range: $4.40 – $5.60

Average $5.10

Silver, as an industrial raw material, has tended to suffer price erosion during episodes of global economic slowdown, but this go-round the downside risk appears to be rather limited.

Unless truly a "hard landing" befalls the US economy, it is difficult to envision from where the selling would derive to overwhelm demand. Recently acquired large speculative holdings are believed either to have either been liquidated or tucked away for the long haul. The prospect for stagnant commercial usage and slightly augmented primary supply seems already to have been priced into the market. Even if a small production surplus were to emerge, its market impact would be inconsequential after a decade of huge inventory drawdowns.

There also is a positive flipside for silver of slower economic growth – a more conducive investment environment for precious metals as a defensive holding. Although a continuing reallocation away from equities and from the US dollar likely would be more supportive to gold, dollar-denominated silver prices should benefit directly from an ongoing recovery in the euro against the dollar. Moreover, with the gold/silver ratio now back out to 60:1, silver's business cycle disadvantage putatively has already been at least partially priced in.

David Mallalieu

Scotia Capital Inc, Toronto

● Gold

Range: $250 - $300

Average: $286

We believe that ongoing uncertainty in the currency markets will provide for some periods of enthusiasm in the bullion price; however, unabated low cost supply from bottom line focused producers combined with a steady stream of supply from the European Central Banks will keep gold price volatility relatively low. Chinks in the armor of the U.S. dollar have started to appear; however, the Americans will not stand idly by while their currency begins to erode relative to its peers. We are not expecting that the U.S. dollar will lose its position as the reserve currency of choice in 2001. It will not be until deficiencies in the Euro and the Yen become obvious as alternative reserve currencies will gold start to more fully take on its financial mantle. We do not expect this to occur until at least 2002.

Daniel McConvey

Goldman Sachs, NewYork

● Gold

Range: $250 - $295

Average: $275

We are increasingly worried about gold's downside – especially if the US falls into a recession. While we no longer view landslide central bank sales as a likely threat, the dearth of global inflation and global investment buying continues to erode gold's perceived investment qualities.

We can identify four potential positive scenarios. First, we support the consensus view that a continued weaker dollar will help the gold price, although gold remains around its low of about $263 despite the recovery of the euro from $0.827 last fall to over $0.95. Our firm expects further weakening against major currencies in 2001, however, there is not the potential for a gold price move of 1985 - 1987 magnitude, when the dollar weakened massively against both East Asian and European currencies. Second, there is possible scope for less selling from smaller central banks going forward. We are forecasting 200 tonnes per year from non-signatory central banks and this number may prove high. Third, somewhat higher lending rates are possible in 2001, if hedging increases and supply to the lending market decreases from the smaller central banks that are sellers. This could squeeze contangos and de-motivate further hedging and fund shorting. Finally, if equity markets experience continued turmoil, gold may still benefit. While markets have been turbulent, the 10.1% fall in the S&P in 2000 was less than half the one day fall in 1987. Gold's qualities as an alternative investment have not been fully tested.

Three factors worry us. First, weaker global GDP growth – our Goldman Sachs Economics Group has lowered its global GDP estimate for the year from 4.2% to 3.5%. As pointed out by GFMS in its latest update, global gold demand is likely to be materially impacted by the outlook for the US economy, which has recently headed into a downturn. Second, producer hedging could return in force. With producer worries/hopes of a repeat of the 1999 gold price spike fading, and concern over more falls in contango rates (down from about 6% to 4% over the last few months), they are likely to hedge more. Third, the material decline in liquidity levels has helped accelerate the decline in investor interest. Future increases in the gold price are likely to come solely from the physical markets outside of the Western World.

Martin Murenbeeld

M. Murenbeeld & Assoc., Vancouver

● Gold

Range: $256 - $324

Average: $297

There is a "bullish" bias in our baseline forecast for two reasons. The first is the US Dollar, which we expect to decline further in 2001, and the second is the stock market(s), which we expect to remain depressed. The Fed will be cutting interest rates again and again in order to "reflate" the US and world economies. While our baseline forecast does not assume any unusual supply pressures, furthermore, it does allow for some weakness in gold jewelry demand on account of weak economic conditions.

Kamal Naqvi

Macquarie Bank, London

● Gold

Range: $240 - $300

Average: $275

The great hope for gold in 2001 is a weaker US dollar. The resilient strength of the greenback, particularly against the euro, was the key feature of 2000 with other notables being the record levels of official sector sales, modest physical demand, reduced producer selling activity and lower trading liquidity. Unfortunately, we expect more of the same in 2001. A slowdown in the global economy is unlikely to be positive for physical demand, while supply is likely to remain around current levels with mine supply now stagnating but being offset by the official sector. Investor demand remains largely absent, although this could be encouraged by a more significant weakening of the US$, lower interest rates and inflationary pressure.

● Silver

Range: $4.25 - $5.25

Average: $4.75

2000 was a year to forget for silver and 2001 may see little improvement. Jewellery demand in Asia (particularly India) has been disappointing, while China continues to export silver at record levels; resulting in speculative longs finally abandoning their long-held positions. Things are not getting any better for silver, with a slowdown in the global economy likely to hit both industrial and photographic demand. Only a reduction in supply seems possible to arrest a fall below $4.50/oz in 2001, particularly with the Berkshire Hathaway stockpile still (possibly?) overhanging the market.

Kevin Norrish

Barclays Capital, London

● Gold

Range: $250 – $295

Average: $267

More sellers than buyers? The old chestnut applies more appropriately to gold than to any other commodity market at present (and we use the term commodity with purpose, for that is what the gold market has become and a fairly inactive one at that).

Latin American central banks turned sellers in 2000 and there is more to come from this quarter in 2001. If the experience of the silver market is anything to go by, Chinese gold market liberalisation could well lead to further net flows of central bank gold into the market. And this on top of the 400t already guaranteed by the Washington Accord.

After last year's inactivity, producers will need at the very least to top up their hedge books with more sales simply to maintain existing positions. Industry consolidation is far from complete and the acquisition of non-hedgers by hedgers could add to the flow of gold sales from this source.

But where will the demand come from? The prospect of a weaker US$ might once have provided some crumb of comfort for gold bulls, but with key Asian currencies now looking even shakier than the greenback, Asian demand is likely to remain subdued in 2001. In the jewellery sector gold continues to lose market share to platinum, demand for which appears to be unharmed by its rapid appreciation in price. A short covering rally perhaps? Speculators enter 2001 in fairly neutral mode and will probably want to sell before they need to buy. If so, gold's 20-year price of $252/oz could be under threat.

● Silver

Range: $4.40 – $5.20

Average $4.85

Prospects for silver look equally poor. The shift in net speculative positions in Comex silver from long to short toward the end of 2001 suggests that funds are finally getting tired of waiting for the long-run supply/demand deficit to push prices higher. And who can blame them? Surging mine output (as by-product), official sales (China) and high rates of scrap recovery meant that supply grew at almost twice the pace of demand in 2000 and the supply gap eased to its narrowest for many years. Expect more of the same in 2001 and for silver prices to struggle to get much above $5/oz.

Rhona O'Connell

Canaccord Capital (Europe) Ltd, London

● Gold

Range: $265 - $325

Average: $285

Any call on the gold price is effectively a call on the outlook for the US dollar. The outcome of the US Presidential election is probably of less significance than might at first be thought, given the continued presence of Greenspan at the Fed, but the awareness of the need to address the current account deficit is important. George W Bush has already spoken of this as a priority, and conceded under questioning that this would involve some weakening of the dollar. Whether such weakening can be kept to a narrow range is perhaps debatable.

What is reasonably impressive is that, over the past few months, gold has withstood pressure from the dollar's strength and benefited from physical support in the $260-265 band. The underlying fundamentals are not strong enough to have built up much pressure, however, and a change of range would require a burst of short covering.

Given that large-scale speculators have been running a cumulative short position on COMEX of 178t as of mid-January, this is perfectly conceivable. The likelihood of fresh mine selling into any price strength is relatively low, but the demand side of the market is still very price sensitive and any rally of more than a few dollars sidelines buyers.

Currently the safest way to forecast the gold price is to suggest that the price in Euros will move sideways, and let the reader take his own view on the currency! On the view that the dollar has peaked, but with few expectations of any over-bearing stresses in the financial system, it is arguable that gold will trade in higher ranges in future than it has over the past two years. Typically when gold changes range it does so by roughly 20%, but the fundamentals in place at present do not suggest that it has the inherent strength to do so this time around.

● Gold

Range: $4.50 - $6.00

Average: $5.25

The silver market is, in terms of market dynamics, currently dominated by China and India. Although India is not the world's largest silver consumer, it is one of the most price elastic, and recent local price weakness has generated steady physical demand. There is also good demand for investment grade material destined for parts of the Far East.

Much of this demand is, however, being met by exports from China and there has been little stress on supplies from London or Zurich. Chinese inventories are high (we estimate a minimum of 600M ounces), stemming from the days when China was on the silver standard, and also reflecting an excess of local mine production over demand. Any rally in 2000 has seen silver teased out from this source. For the longer term the western market is expected to continue to sustain deficits in the region of 100M ounces annually, but the availability of hidden inventory (leftover from the late 1970s) and the underlying cost structure of the primary supply side argues against a sustainable bull market.

Professional activity in the market remains low, but trading exposure is biased towards the short side. We firmly believe that rumours of recent sales from Berkshire Hathaway can be discounted in favour of short-side trading plus the appearance of Oriental physical material.

The recent weakness in the price is partly a function of this plentiful supply but is also allied to the weakness in the gold price. Steady physical demand will maintain prices above $4.50 and a rally to levels in excess of $5 is easily foreseeable.

Frédéric Panizzutti

MKS Finance SA, Geneva

● Gold

Range: $260 - $320

Average: $282

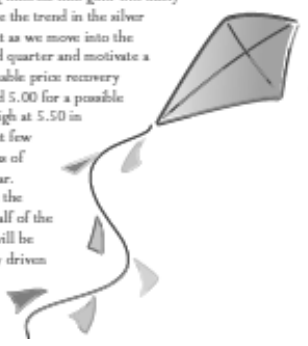
We expect gold to face renewed price pressure in the first quarter of 2001 on the back of weaker physical activity. A consistent price recovery is expected in the second quarter of the year on the back of a major global currency trend change leading to a weakening USD. Interest rates cuts by the Fed in the USA to ensure a soft landing scenario should attract fresh interest into gold. The physical activity as well as the investment and speculative buying will likely take off toward summer time and enable gold to break over its 300.00 main psychological target level and likely move as high as 320.00 toward year end. In a higher demand environment dominated by "liquidity-trading" the ECB "Washington Accord", announced in September 1999 would finally have a positive market impact and enable the yellow metal to rally. The year 2001 will remain sharply dominated by currency markets and general long-term assets reallocation.

● Silver

Range: $3.90 - $5.50

Average: $4.60

Silver will continue to face punctual downside pressure within the first half of the year 2001. Renewed long position liquidation could likely further depress the white metal price as low as 3.90 toward first quarter end.
More USD weakness and fresh buying interest into gold will likely reverse the trend in the silver market as we move into the second quarter and motivate a reasonable price recovery toward 5.00 for a possible year high at 5.50 in the last few months of the year.
While the first half of the year will be mainly driven by specific long liquidation, the second part of the year will be mainly driven by major currency plays but also by the stronger gold market. Trading volumes will increase and markets will be more volatile again.



John Reade

UBS Warburg, London

● Gold

Range: $260 - $310

Average: $282.50

One of the major contributing factors to the poor performance of the gold price in 2000 was the strength of the US dollar. Given that our economists are forecasting that the recent dollar weakness will continue, we are expecting a reasonably good performance from gold this year.

We expect gold supply to remain constrained in 2000. The mining sector will report similar or only slightly higher gold production. Central bank sales will continue at the high levels seen in 2000, but we do not expect any acceleration. Producers are likely to add to their aggregate hedge positions, but not at the pace of the peak hedging years in 1997 and 1999. Finally, one of the key elements in supply last year, namely the dishoarding of privately held European gold, is not likely to be an issue this year due to the sharp increase in the value of the Euro.

The weaker US dollar will assist gold demand as well. Already we have seen a very promising start to physical demand in the first few days of the New Year. The combination of restrained supply and higher gold demand will allow gold to move steadily higher through the year.

● Silver

Range: $4.40 - $5.20

Average: $4.85

Silver performed very poorly in 2000 and has fallen to fresh three-year lows in the first few days of the New Year. We believe that silver is currently over-sold and that a bounce can be expected in the short term. Through the balance of the year we expect silver to trade higher in sympathy with gold and the weak US dollar.

Jeffrey Rhodes

Standard Bank London – Dubai Office

● Gold

Range: $253.50 - $295.65
Average: $272.40

The gold price earthquake of October 1999 when gold reached $340 per ounce and the aftershock in February last year, which saw gold trade above $310, had a profound impact on the market in Y2K. Speculative interest waned with the reduction in trading volume highlighted by the sharp fall in clearing turnover in London. The physical markets were also shaken by the extreme price volatility, and it has taken a long time for traders in the gold souk to recover from the infamous Washington Central Bank statement.

However, there were encouraging signs in Q4 2000 that the physical market is returning to normal and demand was buoyant during the Ramadan and Christmas periods. The physical sector, which thrives on the kind of price stability seen in the second half of last year, does not usually drive gold higher, but it can underpin the market. My view is that while the fund selling seen on the COMEX since the turn of the year is likely to push the price below $260 in Q1 2001, physical demand will provide strong support between $258 and $252. Falling US interest rates and a weaker US dollar are predicted this year, which will be plus factors for the yellow metal and I would expect to see a stronger gold price in the last quarter of the year, closer to $300 than $250.

● Silver

Range: $4.27 - $5.17
Average: $4.76

The big loser in 2000, falling by 15%. The blame for this has been attributed to a flow of metal out of China, although low silver lease rates/high contangos and a strong US dollar were an attractive package for silver producers. The predicted slowdown in the US economy is not good news for this industrial metal and a test of $4.40 in Q1 is probable, while further weakness to $4.25 is a distinct possibility. However, silver has a habit of doing what is least expected and a recovery above $5 in the second half of 2001 would be no great surprise.

Hennie Rijnbeek

Rabobank International, London

● Gold

Range: $255 – $295
Average: $268

As in 1999 (and the years before), we have dreamed, prayed and hoped that in 2000 the market would turn, but as usual we were disappointed. For 2001 we do not expect any dramatic change in this situation – no ECB announcements and no massive change in the investor's perception. There will be buying in India, Turkey and in the Far East, just like there will be selling by producers and central banks. Of course we will see spikes and dips, but we expect the spikes to flatten and the dips to deepen.

The anticipated slowdown of the global economy will take its toll in jewellery demand. Losses in the stock markets will not lead to flight to quality, at least not to gold.

All in all not too much different from 2000, except that this year we start some 20 dollars lower.

● Silver

Range: $4.40 – $5.20
Average: $4.65

Silver will take its direction from gold again. The weakening of the global economy will depress industrial and jewellery demand. Output will be robust and investors are not expected to show any substantial interest.

Andy Smith

Mitsui London

● Gold

Range: $210 – $290
Average: $250

The void in business can be measured by the length of a conspiracy theorist's Email. Is gold too small to count, to enter or to dally in before exit?

After a dead chad bounce, macro/political hopes ride with a wild bunch (Saddam, Castro, the EU) seeking to 'stand up' to the dollar and yearning for a US financial bloodbath of Perkinpah proportions. As macro hedge funds retire from a sport they found too hard, perhaps best not to bet the portfolio – or mine – on picking the time and consequences of a dollar turn? For speculators, time will heal the phobia of all carry trades, contracted in August 1998, which has deterred aggressive short selling all 2000. For central banks, the legacy position gets more painful: disappearing market depth, barricades (choice!) in a supposed open meadow of absorbent physical demand; and potentially explosive mine hedging. Those holding their breath for an end to central bank sales (a new Guinness Book of Records category?) will have noted a silver lining: the dominant issue in silver since 1998 (a century after most official silver was sold) is…official (Chinese) selling. Smaller official gold holders (a collectively very big 7000+tonnes) are jumping ship while the SS Duisenberg ploughs a predictable, icy course. Unless any 'Duisenberg 2' statement includes the word buy – fade it.

2001 will probably see the end to the longest strike in gold mining – the down tools on new gold borrowing from last February. Catching up for the 2000 hiatus would mean about 700 tonnes new hedging in 2001 (GFMS calculate average annual hedging 340 tonne 1995-99). The bad news – having discovered that Godot is more timely than macro miracles; having seen hedge trimming turn into a Texas chainsaw massacre of opportunity loss; having bought calls and received no thanks and less return – consolidation is the last strategic card for miners to play (before closure). If the hedging appetite of acquisitors were grafted onto under-hedged targets, expect 1000+ tonnes of new hedging. Market too thin to take it? Expect it sooner.

● Silver

Range: $3.50 – $5.25
Average: $4.25

"You'd better hope he's really dead" was the epitaph ex-Philadelphia mayor Frank 'Big Bambino' Rizzo wished for. Silver better hope Mr Buffett is out. Its flight to the origin of volatility and volume since his entry is gold-like; leaving a Roach Motel for remaining longs, windows nailed down, doors locked. Sub $4.70 silver and sub-gold lease rates suggest any supposed silver 'gap' has been inundated. Spring for cyclical metals will be late in 2001. Good news – China probably sold less in 2000 (India may have run out of Chinese metal by Q4). Bad news – China will probably use all instruments (except auctions) to offload into price strength. They will have to compete with producers with itchier fingers on hedge triggers. And both will have to wrestle in a contracting ring of forward/option volumes. Low spot and high contango means buyers may pick their punches – waiting to buy spot on (big) dips rather than roping themselves into forward buying.

Bob Takai

Sumitomo Corp, Tokyo

● Gold

Range: $255 - $285
Average: $265

I don't expect gold to move much in 2001. The dollar's strength will be diminished, and that will limit the downside by traditional bargain hunting in Asia, while the usual sellers – producers and central banks – will cap the upside. An unexciting market will also discourage speculation. All in all, gold will be the quietest commodity of all.

● Silver

Range: $4.00 - $5.00
Average: $4.55

I expect silver will edge lower during 2001. Exports from China, increasing production in Mexico, less demand in India and more fund liquidation will point the price southward. The quiet gold market will also discourage investment in the white metal.

Peter Ward

Lehman Bros. New York

● Gold

Range: $240 - $295
Average: $265

I continue to think that gold is in the process of commoditisation, or, as I've also called it, "reverse alchemy."

There are few reasons to anticipate a sustained higher price for gold and many in favour of a declining one. The reasons for central banks to hold massive amounts of gold have clearly declined and they will continue to sell their reserves into the market. In addition, private investors have shown no inclination toward gold despite the recent turbulence in global equity markets. Therefore, we believe private investors will also be sellers.

Some point to the deficit between consumption and mine production as a reason for the price to recover. But, this is misleading as that gap is filled by the ongoing liquidation of above-ground stocks. Our fear is that the rate of liquidation will accelerate in the future. It would then take an increasingly large gap to be able to absorb it – which would require lower prices to encourage more jewellery consumption and curtail mine production.

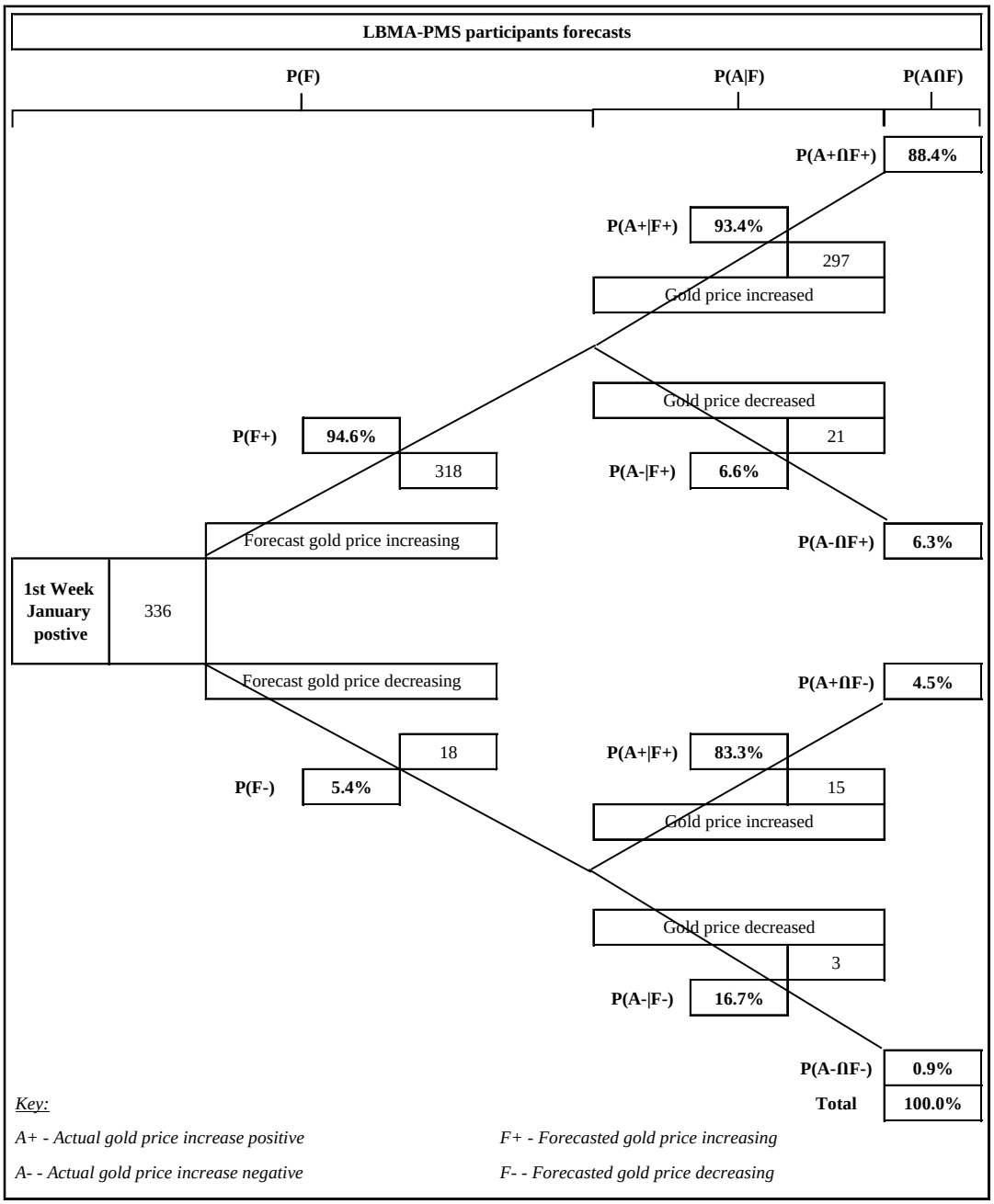
● Silver

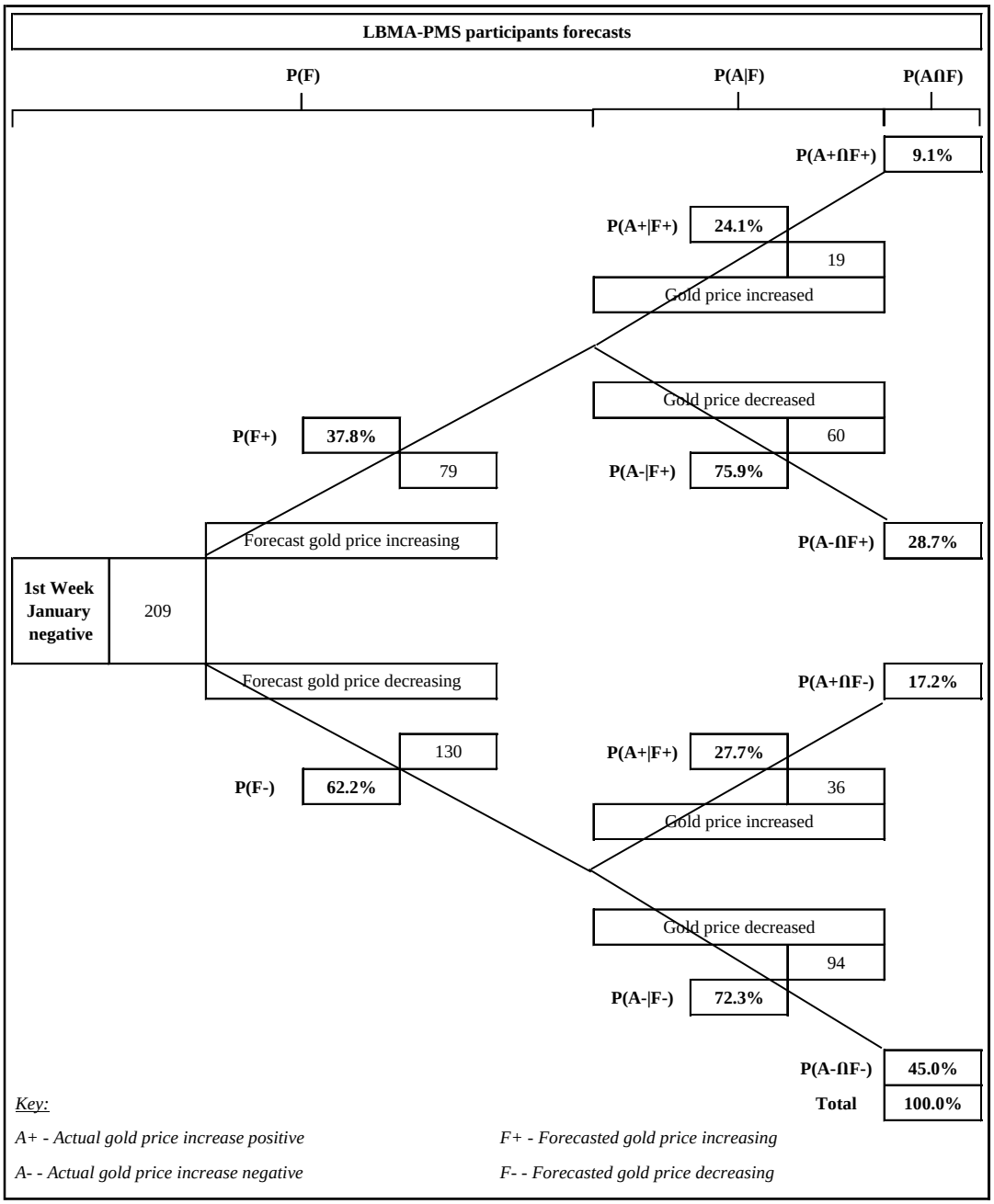
Range: $4.40 - $5.00
Average: $4.65

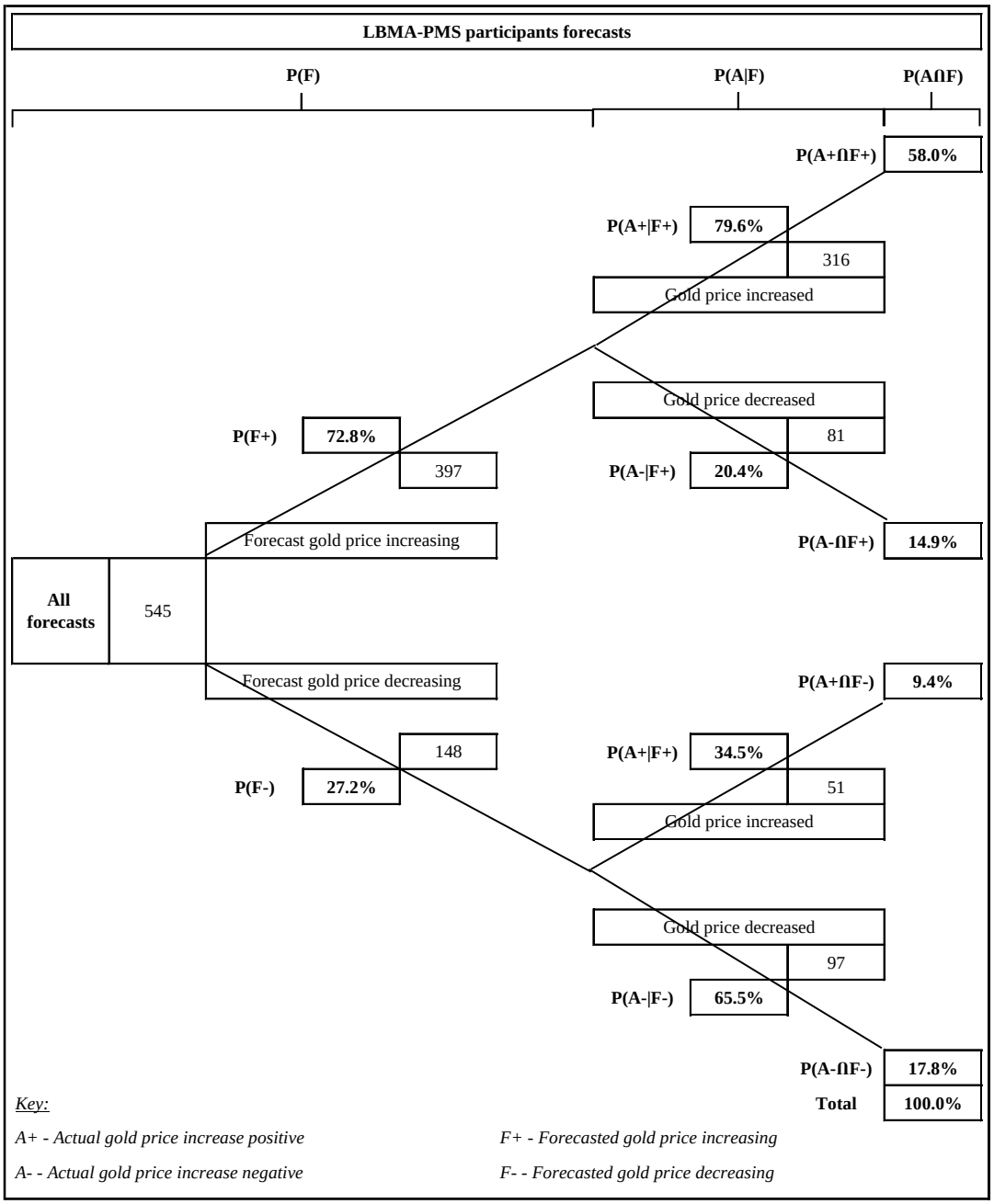
Weak end markets and too much price-insensitive supply don't paint a bright picture for rising prices over the next year.

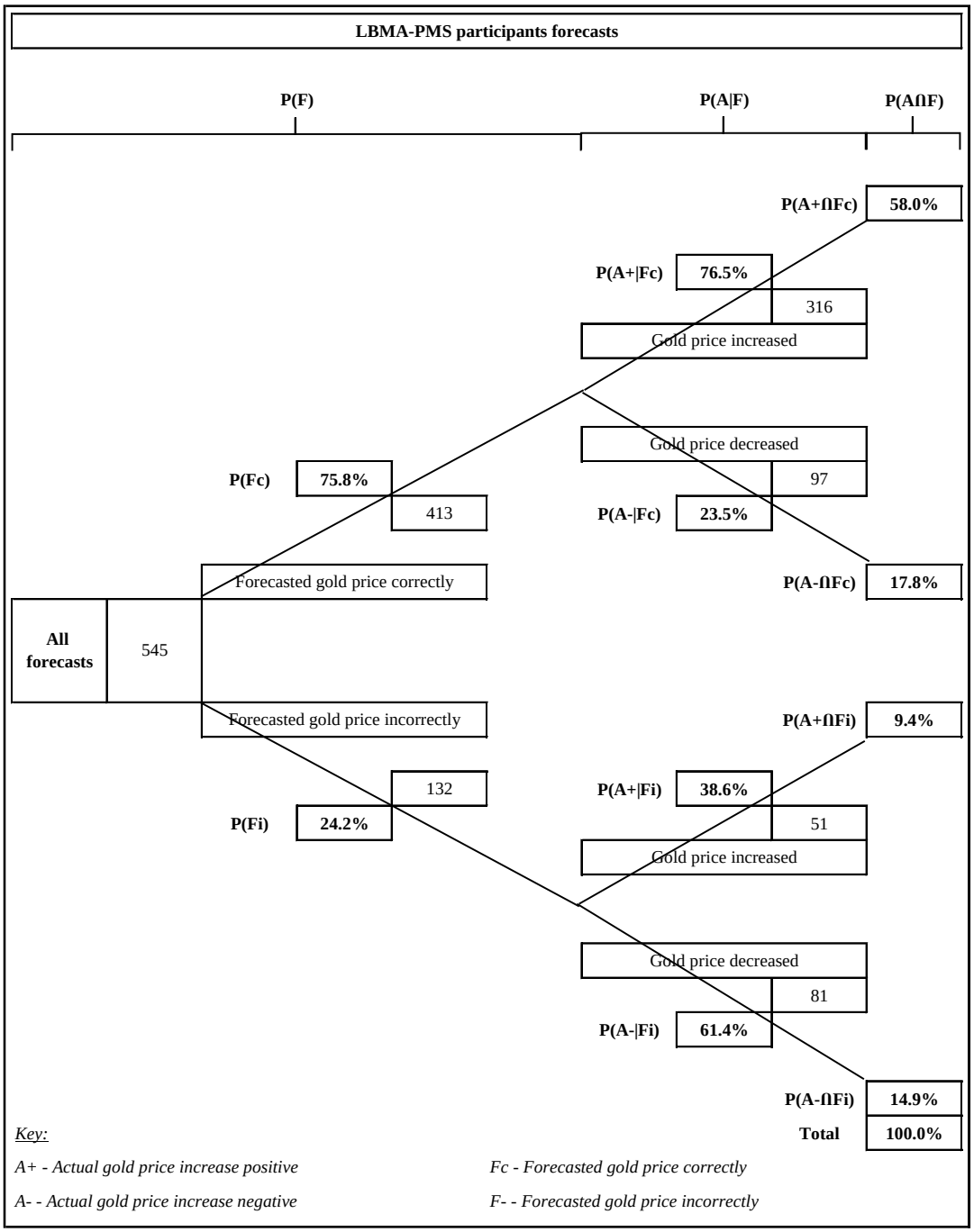
***Forecast 2000* is published by the London Bullion Market Association**
For further information please contact
Susanne M. Capano, Editor, 6 Frederick's Place London EC2R 8BT
Telephone: 020 7796 3067
Fax: 020 7796 4345
Email: alchemist@lbma.org.uk

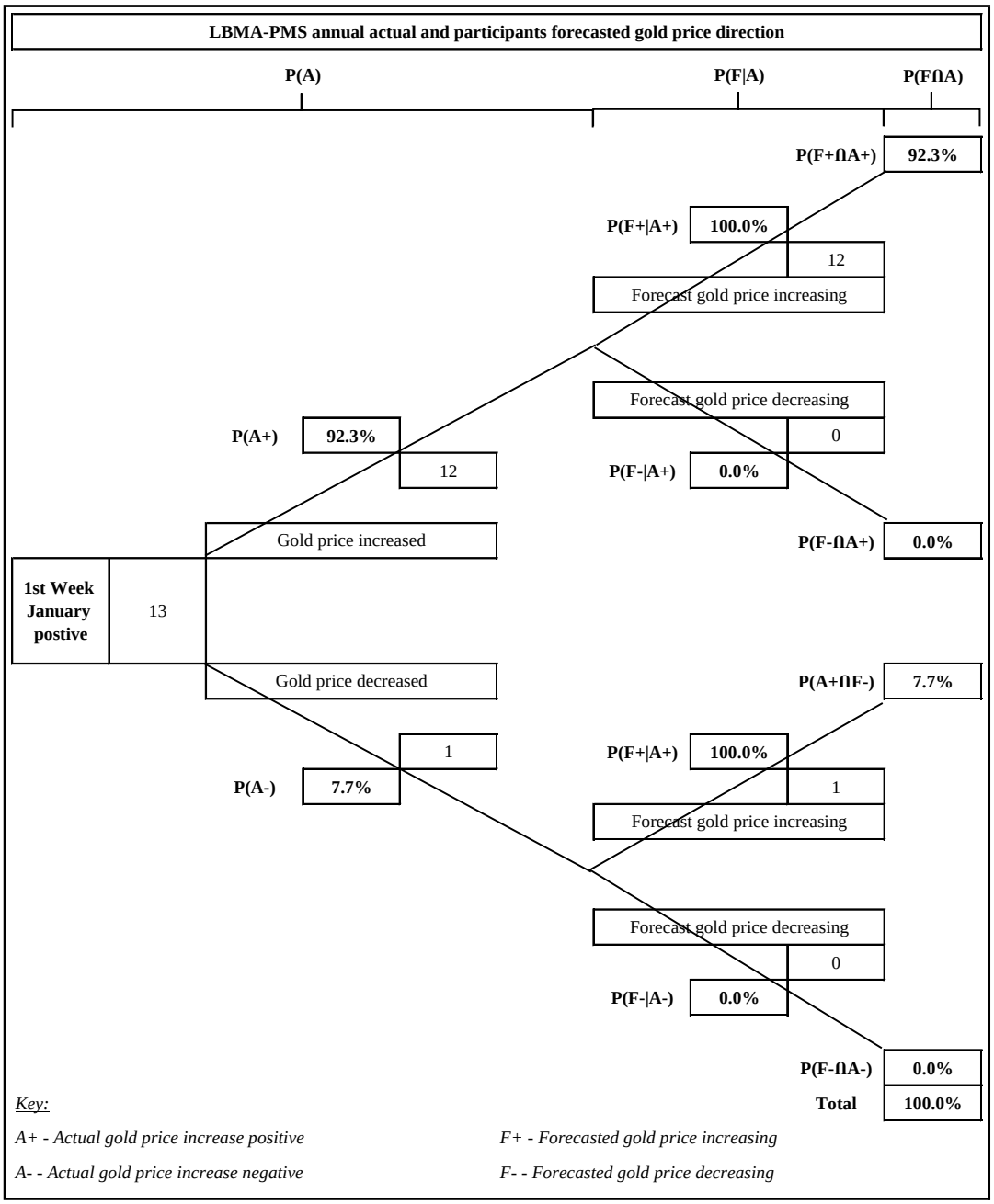
Annexure F: LBMA-PMS Gold Forecasts Probabilities

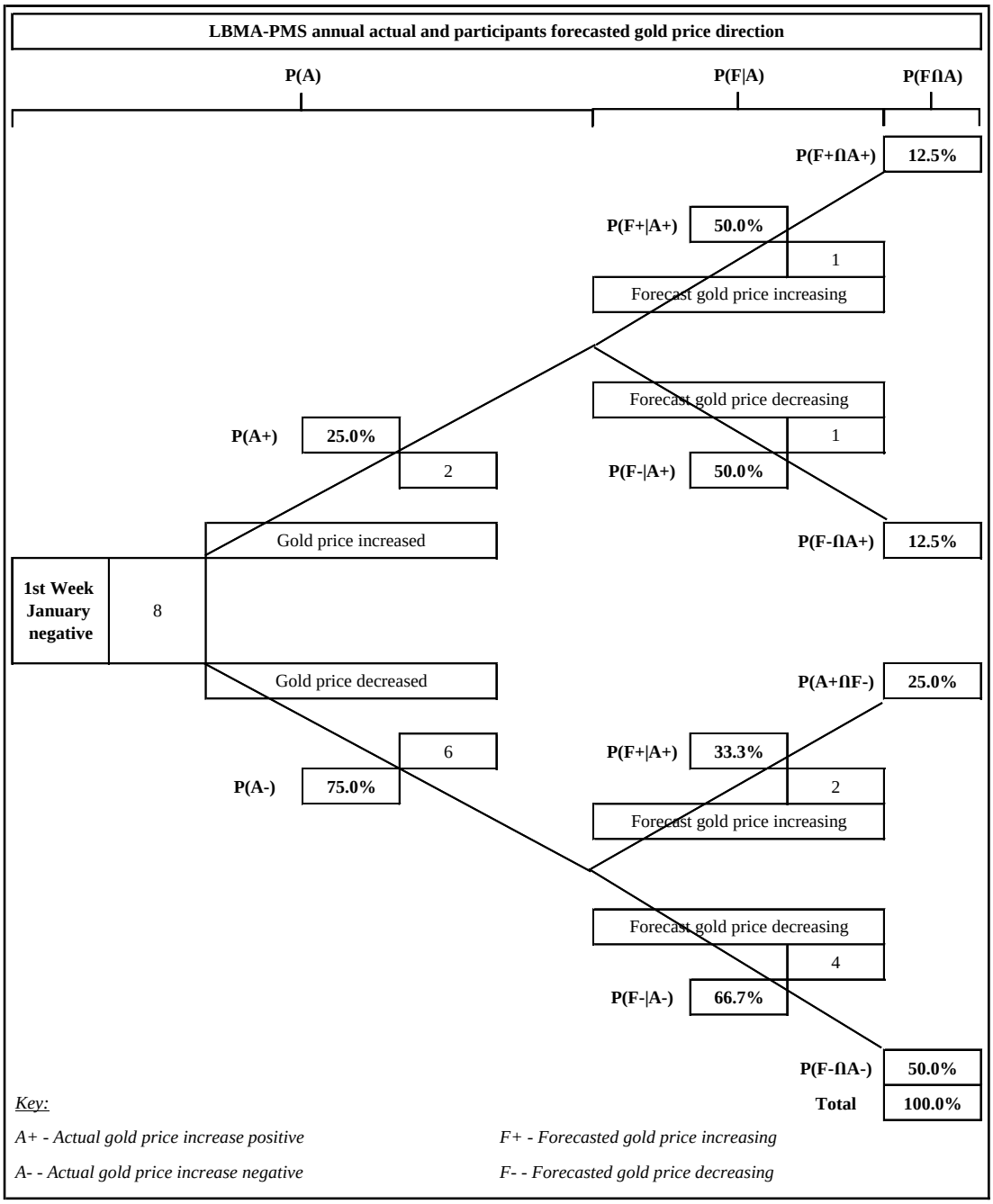


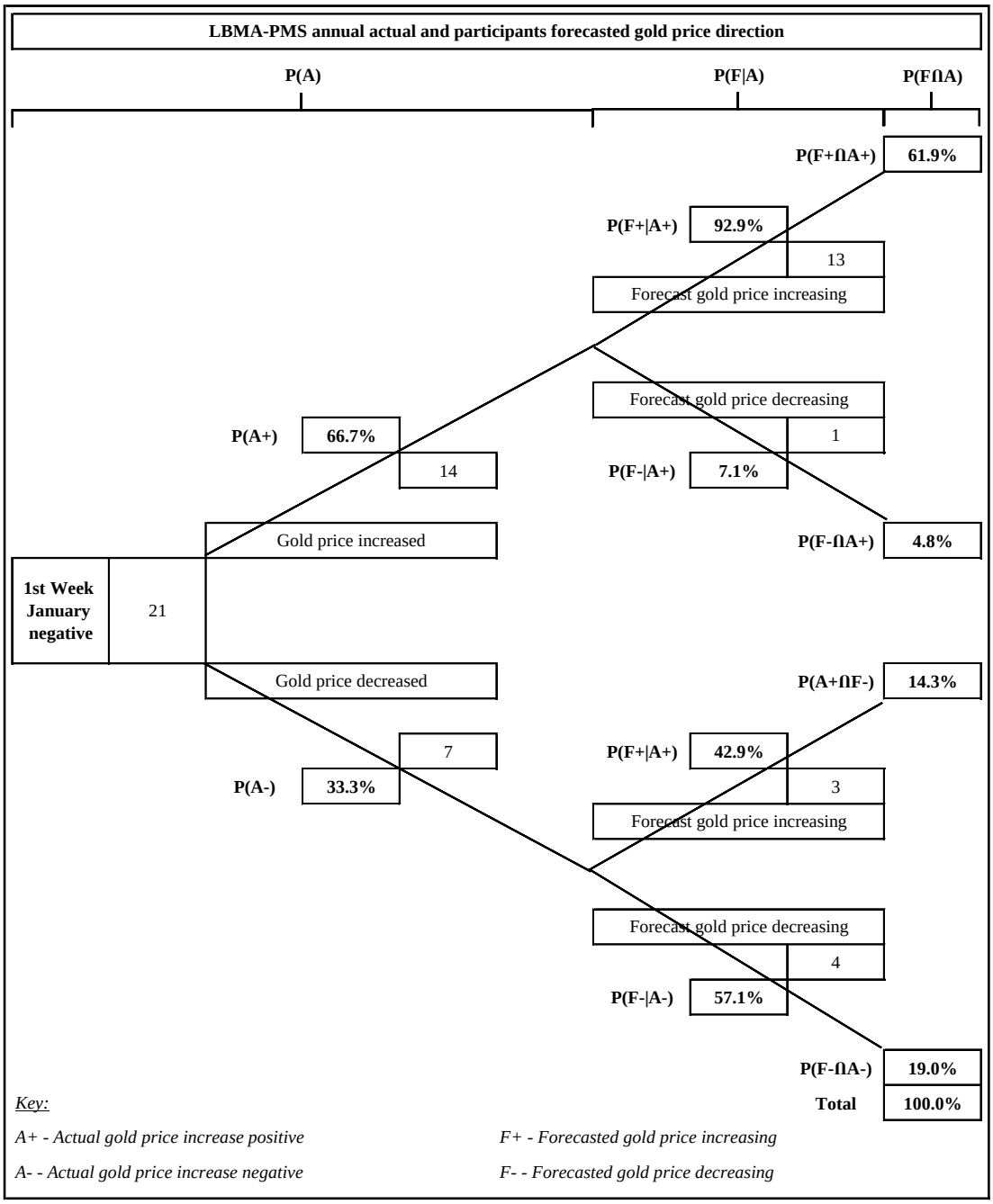

LBMA-PMS participants forecasts
P(F)
P(A|F)
P(A∩F)
P(A+∩F+) 9.1%
P(A+|F+) 24.1%
19
Gold price increased
Gold price decreased
P(F+) 37.8%
60
79
P(A-|F+) 75.9%
Forecast gold price increasing
P(A-∩F+) 28.7%
1st Week January negative
209
Forecast gold price decreasing
P(A+∩F-) 17.2%
130
P(A+|F+) 27.7%
P(F-) 62.2%
36
Gold price increased
Gold price decreased
94
P(A-|F-) 72.3%
P(A-∩F-) 45.0%
Total 100.0%
Key:
A+ - Actual gold price increase positive
F+ - Forecasted gold price increasing
A- - Actual gold price increase negative
F- - Forecasted gold price decreasing

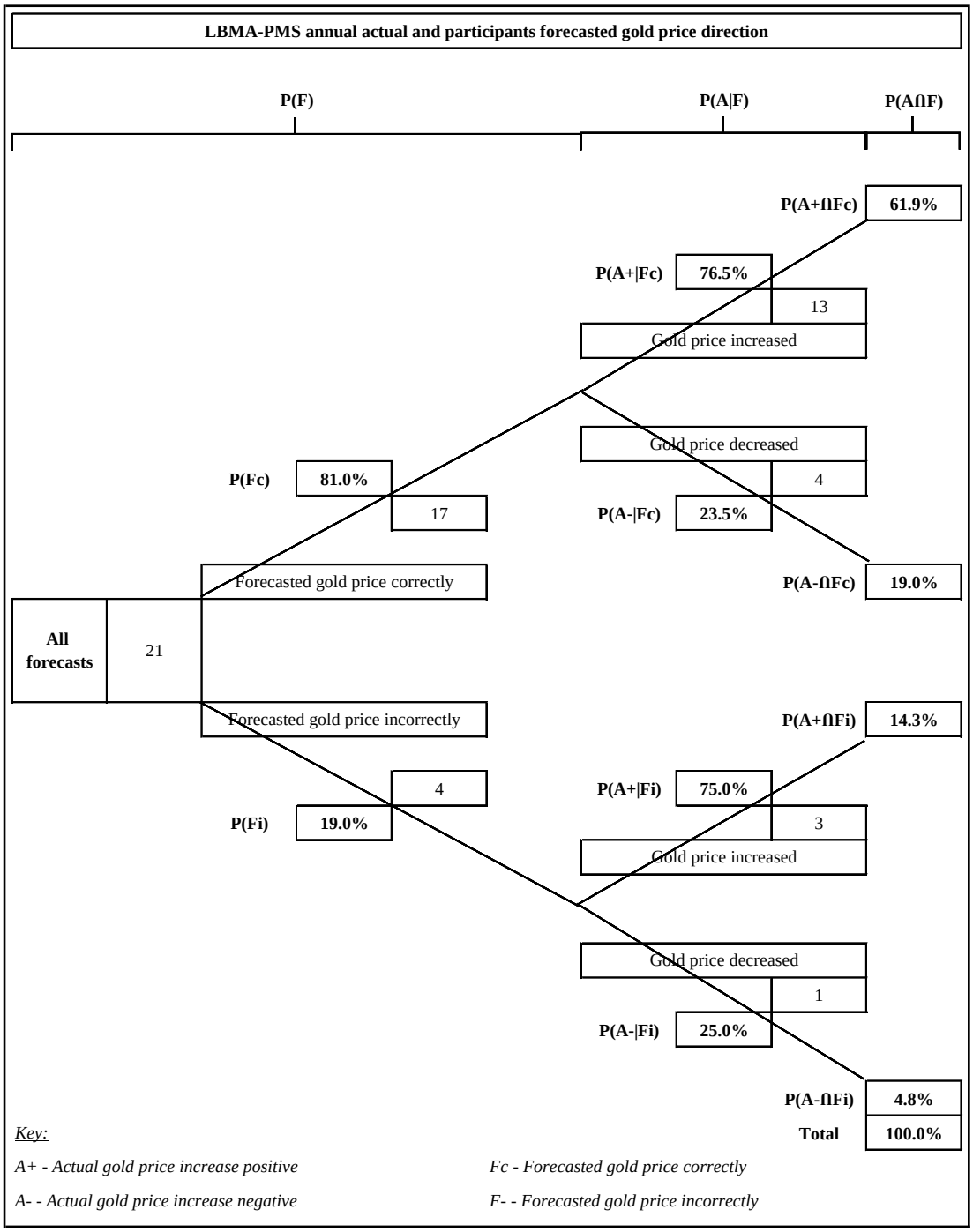

LBMA-PMS participants forecasts
P(F)
P(A|F)
P(A∩F)
P(A+∩F+) 58.0%
P(A+|F+) 79.6%
316
Gold price increased
Gold price decreased
81
P(A-|F+) 20.4%
P(F+) 72.8%
397
Forecast gold price increasing
P(A-∩F+) 14.9%
All forecasts
545
Forecast gold price decreasing
P(A+∩F-) 9.4%
148
P(A+|F+) 34.5%
P(F-) 27.2%
51
Gold price increased
Gold price decreased
97
P(A-|F-) 65.5%
P(A-∩F-) 17.8%
Total 100.0%
Key:
A+ - Actual gold price increase positive
F+ - Forecasted gold price increasing
A- - Actual gold price increase negative
F- - Forecasted gold price decreasing


LBMA-PMS participants forecasts
P(F)
P(A|F)
P(A∩F)
P(A+∩Fc) 58.0%
P(A+|Fc) 76.5%
316
Gold price increased
Gold price decreased
P(Fc) 75.8%
97
413
P(A-|Fc) 23.5%
Forecasted gold price correctly
P(A-∩Fc) 17.8%
All forecasts
545
Forecasted gold price incorrectly
P(A+∩Fi) 9.4%
132
P(A+|Fi) 38.6%
P(Fi) 24.2%
51
Gold price increased
Gold price decreased
81
P(A-|Fi) 61.4%
P(A-∩Fi) 14.9%
Total 100.0%
Key:
A+ - Actual gold price increase positive
Fc - Forecasted gold price correctly
A- - Actual gold price increase negative
F- - Forecasted gold price incorrectly


LBMA-PMS annual actual and participants forecasted gold price direction
P(A)
P(F|A)
P(F∩A)
P(F+∩A+)
92.3%
P(F+|A+)
100.0%
12
Forecast gold price increasing
Forecast gold price decreasing
0
P(A+)
92.3%
12
P(F-|A+)
0.0%
Gold price increased
P(F-∩A+)
0.0%
1st Week January postive
13
Gold price decreased
P(A+∩F-)
7.7%
1
P(F+|A+)
100.0%
P(A-)
7.7%
1
Forecast gold price increasing
Forecast gold price decreasing
0
P(F-|A-)
0.0%
P(F-∩A-)
0.0%
Total
100.0%
Key:
A+ - Actual gold price increase positive
F+ - Forecasted gold price increasing
A- - Actual gold price increase negative
F- - Forecasted gold price decreasing


LBMA-PMS annual actual and participants forecasted gold price direction
P(A)
P(F|A)
P(FΠA)
1st Week January negative
8
Gold price increased
Gold price decreased
P(A+)
25.0%
2
P(A-)
75.0%
6
P(F+|A+)
50.0%
1
Forecast gold price increasing
Forecast gold price decreasing
P(F-|A+)
50.0%
1
P(F+|A-)
33.3%
2
Forecast gold price increasing
Forecast gold price decreasing
P(F-|A-)
66.7%
4
P(F+ΠA+)
12.5%
P(F-ΠA+)
12.5%
P(A+ΠF-)
25.0%
P(F-ΠA-)
50.0%
Total
100.0%
Key:
A+ - Actual gold price increase positive
A- - Actual gold price increase negative
F+ - Forecasted gold price increasing
F- - Forecasted gold price decreasing


LBMA-PMS annual actual and participants forecasted gold price direction
P(A)
P(F|A)
P(FΠA)
1st Week January negative
21
P(A+)
66.7%
14
P(A-)
33.3%
7
Gold price increased
Gold price decreased
P(F+|A+)
92.9%
13
Forecast gold price increasing
Forecast gold price decreasing
1
P(F-|A+)
7.1%
P(F+|A+)
42.9%
3
Forecast gold price increasing
Forecast gold price decreasing
4
P(F-|A-)
57.1%
P(F+ΠA+)
61.9%
P(F-ΠA+)
4.8%
P(A+ΠF-)
14.3%
P(F-ΠA-)
19.0%
Total
100.0%
Key:
A+ - Actual gold price increase positive
A- - Actual gold price increase negative
F+ - Forecasted gold price increasing
F- - Forecasted gold price decreasing


LBMA-PMS annual actual and participants forecasted gold price direction
P(F)
P(A|F)
P(AΩF)
P(A+ΩFc)
61.9%
P(A+|Fc)
76.5%
13
Gold price increased
Gold price decreased
P(Fc)
81.0%
4
17
P(A-|Fc)
23.5%
Forecasted gold price correctly
P(A-ΩFc)
19.0%
All forecasts
21
Forecasted gold price incorrectly
P(A+ΩFi)
14.3%
4
P(A+|Fi)
75.0%
P(Fi)
19.0%
3
Gold price increased
Gold price decreased
1
P(A-|Fi)
25.0%
P(A-ΩFi)
4.8%
Total
100.0%
Key:
A+ - Actual gold price increase positive
Fc - Forecasted gold price correctly
A- - Actual gold price increase negative
F- - Forecasted gold price incorrectly

Annexure G: LBMA-PMS Condorcet Jury Theorem Probabilities

Year	1st Week market direction	Actual Market direction	Total forecasts	Up forecasts	Down forecasts	Condorcet Up probability	Condorcet Down probability
2000	0.8%	0.2%	25	19	6	99.8%	0.2%
2001	(3.9%)	(3.0%)	24	8	16	7.6%	92.4%
2002	2.7%	12.5%	25	22	3	100.0%	0.0%
2003	12.3%	14.8%	24	24	0	100.0%	0.0%
2004	15.8%	11.2%	24	24	0	100.0%	0.0%
2005	3.6%	7.9%	25	21	4	100.0%	0.0%
2006	19.7%	26.4%	25	25	0	100.0%	0.0%
2007	3.7%	13.2%	29	26	3	100.0%	0.0%
2008	23.5%	20.3%	24	24	0	100.0%	0.0%
2009	(1.8%)	10.3%	24	15	9	92.4%	7.6%
2010	15.8%	20.6%	26	26	0	100.0%	0.0%
2011	12.0%	22.1%	24	24	0	100.0%	0.0%
2012	2.3%	5.8%	26	25	1	100.0%	0.0%
2013	(0.3%)	(18.3%)	23	18	5	99.9%	0.1%
2014	(12.8%)	(11.4%)	28	0	28	0.0%	100.0%
2015	(5.1%)	(9.2%)	31	4	27	0.0%	100.0%
2016	(5.9%)	7.3%	31	4	27	0.0%	100.0%
2017	(6.5%)	0.5%	23	11	12	50.0%	50.0%
2018	4.7%	0.9%	24	21	3	100.0%	0.0%
2019	1.4%	8.9%	30	27	3	100.0%	0.0%
2020	11.9%	21.3%	30	29	1	100.0%	0.0%
Average	**4.5%**	**12.3%**	**26**	**19**	**7**	**99.53%**	**0.47%**
Gold price >1%	**8.4%**	**14.5%**	**26**	**23**	**4**	**100.00%**	**0.00%**
Gold price <1%	**-3.3%**	**-5.8%**	**25**	**12**	**14**	**50.00%**	**50.00%**

Annexure H: LBMA-PMS Gold Forecasts (High-Low) - Consensus Estimates

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LBMA average price for the prior period	LBMA average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	2000	$279/oz	$279/oz	$285/oz	$297/oz	$295/oz	$303/oz
a	Actual % Δ / APE %	**Neutral**	**0.2%**	**2.1%**	6.4%	5.7%	8.4%
	2001	$279/oz	$271/oz	$271/oz	$277/oz	$275/oz	$285/oz
a	Actual % Δ / APE %	**Down**	**(3.0%)**	**(0.1%)**	2.1%	1.5%	5.0%
	2002	$271/oz	$310/oz	$285/oz	$289/oz	$289/oz	$295/oz
a	Actual % Δ / APE %	**Up**	**12.5%**	(8.0%)	(6.6%)	(6.7%)	**(4.7%)**
	2003	$310/oz	$363/oz	$350/oz	$357/oz	$355/oz	$365/oz
a	Actual % Δ / APE %	**Up**	**14.8%**	(3.7%)	(1.7%)	(2.3%)	**0.5%**
	2004	$363/oz	$409/oz	$410/oz	$422/oz	$423/oz	$438/oz
a	Actual % Δ / APE %	**Up**	**11.2%**	**0.2%**	3.3%	3.3%	6.9%
	2005	$409/oz	$444/oz	$425/oz	$437/oz	$440/oz	$449/oz
a	Actual % Δ / APE %	**Up**	**7.9%**	(4.4%)	(1.8%)	(1.0%)	**0.9%**
	2006	$444/oz	$604/oz	$525/oz	$542/oz	$540/oz	$555/oz
a	Actual % Δ / APE %	**Up**	**26.4%**	(13.1%)	(10.3%)	(10.6%)	**(8.1%)**
	2007	$604/oz	$695/oz	$638/oz	$654/oz	$658/oz	$670/oz
a	Actual % Δ / APE %	**Up**	**13.2%**	(8.3%)	(5.9%)	(5.5%)	**(3.7%)**
	2008	$695/oz	$872/oz	$840/oz	$877/oz	$859/oz	$901/oz
a	Actual % Δ / APE %	**Up**	**20.3%**	(3.7%)	**0.5%**	(1.5%)	3.3%
	2009	$872/oz	$972/oz	$859/oz	$898/oz	$904/oz	$951/oz
a	Actual % Δ / APE %	**Up**	**10.3%**	(11.6%)	(7.7%)	(7.1%)	**(2.2%)**
	2010	$972/oz	$1,225/oz	$1,150/oz	$1,190/oz	$1,184/oz	$1,214/oz
a	Actual % Δ / APE %	**Up**	**20.6%**	(6.1%)	(2.9%)	(3.3%)	**(0.8%)**
	2011	$1,225/oz	$1,572/oz	$1,415/oz	$1,450/oz	$1,443/oz	$1,479/oz
a	Actual % Δ / APE %	**Up**	**22.1%**	(10.0%)	(7.7%)	(8.2%)	**(5.9%)**
	2012	$1,572/oz	$1,669/oz	$1,718/oz	$1,749/oz	$1,765/oz	$1,800/oz
a	Actual % Δ / APE %	**Up**	**5.8%**	**2.9%**	4.8%	5.8%	7.9%
	2013	$1,669/oz	$1,411/oz	$1,670/oz	$1,721/oz	$1,730/oz	$1,763/oz
a	Actual % Δ / APE %	**Down**	**(18.3%)**	**18.3%**	22.0%	22.6%	24.9%
	2014	$1,411/oz	$1,266/oz	$1,194/oz	$1,223/oz	$1,225/oz	$1,260/oz
a	Actual % Δ / APE %	**Down**	**(11.4%)**	(5.7%)	(3.4%)	(3.3%)	**(0.5%)**
	2015	$1,266/oz	$1,160/oz	$1,193/oz	$1,221/oz	$1,225/oz	$1,258/oz
a	Actual % Δ / APE %	**Down**	**(9.2%)**	**2.8%**	5.2%	5.6%	8.5%
	2016	$1,160/oz	$1,251/oz	$1,078/oz	$1,105/oz	$1,101/oz	$1,138/oz
a	Actual % Δ / APE %	**Up**	**7.3%**	(13.9%)	(11.7%)	(12.0%)	**(9.1%)**
	2017	$1,251/oz	$1,257/oz	$1,213/oz	$1,240/oz	$1,250/oz	$1,270/oz
a	Actual % Δ / APE %	**Neutral**	**0.5%**	(3.6%)	(1.4%)	**(0.6%)**	1.0%
	2018	$1,257/oz	$1,268/oz	$1,299/oz	$1,318/oz	$1,325/oz	$1,346/oz
a	Actual % Δ / APE %	**Neutral**	**0.9%**	**2.4%**	3.9%	4.5%	6.1%
	2019	$1,268/oz	$1,393/oz	$1,300/oz	$1,313/oz	$1,313/oz	$1,329/oz
a	Actual % Δ / APE %	**Up**	**8.9%**	(6.7%)	(5.7%)	(5.8%)	**(4.6%)**
	2020	$1,393/oz	$1,770/oz	$1,550/oz	$1,566/oz	$1,565/oz	$1,583/oz
a	Actual % Δ / APE %	**Up**	**21.3%**	(12.4%)	(11.5%)	(11.6%)	**(10.5%)**
	2000-2020	$903/oz	$974/oz	$936/oz	$959/oz	$960/oz	$983/oz
b	Actual % Δ / APE %	**Up**	**12.3%**	7.6%	6.8%	7.0%	**6.6%**
c	Root Mean Square error	160		102	100	**102**	103
d	Theil U(II) – Median	156.7%		99.9%	**97.2%**	100.0%	100.6%
e	Theil U(II) – Prior			63.8%	62.0%	**63.8%**	64.2%
f	ReLMAE - Prior			99.3%	78.7%	81.1%	**66.3%**

Annexure I: LBMA-PMS Silver Forecasts - Consensus Estimates

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LBMA average price for the prior period	LBMA average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	2000	$5.22/oz	$4.95/oz	$5.26/oz	$5.55/oz	$5.50/oz	$5.73/oz
a	Actual % Δ / APE %	**Down**	**(5.5%)**	**6.3%**	12.1%	11.1%	15.7%
	2001	$4.95/oz	$4.37/oz	$4.63/oz	$4.77/oz	$4.75/oz	$5.22/oz
a	Actual % Δ / APE %	**Down**	**(13.3%)**	**5.8%**	9.1%	8.7%	19.5%
	2002	$4.37/oz	$4.60/oz	$4.33/oz	$4.47/oz	$4.45/oz	$4.65/oz
a	Actual % Δ / APE %	**Up**	**5.0%**	(5.9%)	(2.8%)	(3.3%)	**1.1%**
	2003	$4.60/oz	$4.88/oz	$4.65/oz	$4.87/oz	$4.80/oz	$5.05/oz
a	Actual % Δ / APE %	**Up**	**5.7%**	(4.7%)	**(0.2%)**	(1.6%)	3.5%
	2004	$4.88/oz	$6.66/oz	$5.84/oz	$6.00/oz	$5.95/oz	$6.19/oz
a	Actual % Δ / APE %	**Up**	**26.7%**	(12.4%)	(9.9%)	(10.7%)	**(7.1%)**
	2005	$6.66/oz	$7.31/oz	$6.25/oz	$6.56/oz	$6.51/oz	$6.95/oz
a	Actual % Δ / APE %	**Up**	**8.9%**	(14.5%)	(10.2%)	(10.9%)	**(5.0%)**
	2006	$7.31/oz	$11.55/oz	$8.31/oz	$8.66/oz	$8.70/oz	$9.04/oz
a	Actual % Δ / APE %	**Up**	**36.7%**	(28.1%)	(25.0%)	(24.7%)	**(21.7%)**
	2007	$11.55/oz	$13.38/oz	$11.98/oz	$12.57/oz	$12.50/oz	$13.08/oz
a	Actual % Δ / APE %	**Up**	**13.7%**	(10.5%)	(6.1%)	(6.6%)	**(2.2%)**
	2008	$13.38/oz	$14.99/oz	$14.13/oz	$15.17/oz	$15.10/oz	$16.00/oz
a	Actual % Δ / APE %	**Up**	**10.7%**	(5.8%)	1.2%	**0.7%**	6.7%
	2009	$14.99/oz	$14.67/oz	$10.35/oz	$11.58/oz	$11.45/oz	$13.00/oz
a	Actual % Δ / APE %	**Down**	**(2.2%)**	(29.5%)	(21.1%)	(22.0%)	**(11.4%)**
	2010	$14.67/oz	$20.19/oz	$17.07/oz	$19.02/oz	$18.65/oz	$19.95/oz
a	Actual % Δ / APE %	**Up**	**27.3%**	(15.5%)	(5.8%)	(7.6%)	**(1.2%)**
	2011	$20.19/oz	$35.12/oz	$26.38/oz	$29.88/oz	$30.00/oz	$32.53/oz
a	Actual % Δ / APE %	**Up**	**42.5%**	(24.9%)	(14.9%)	(14.6%)	**(7.4%)**
	2012	$35.12/oz	$31.15/oz	$31.78/oz	$33.98/oz	$34.00/oz	$36.13/oz
a	Actual % Δ / APE %	**Down**	**(12.7%)**	**2.0%**	9.1%	9.2%	16.0%
	2013	$31.15/oz	$23.79/oz	$31.08/oz	$33.21/oz	$32.50/oz	$35.93/oz
a	Actual % Δ / APE %	**Down**	**(30.9%)**	**30.6%**	39.6%	36.6%	51.0%
	2014	$23.79/oz	$19.08/oz	$19.00/oz	$19.95/oz	$19.90/oz	$21.00/oz
a	Actual % Δ / APE %	**Down**	**(24.7%)**	**(0.4%)**	4.5%	4.3%	10.1%
	2015	$19.08/oz	$15.68/oz	$16.10/oz	$16.76/oz	$16.88/oz	$17.65/oz
a	Actual % Δ / APE %	**Down**	**(21.7%)**	**2.7%**	6.9%	7.7%	12.6%
	2016	$15.68/oz	$17.14/oz	$13.73/oz	$14.74/oz	$14.80/oz	$15.53/oz
a	Actual % Δ / APE %	**Up**	**8.5%**	(19.9%)	(14.0%)	(13.7%)	**(9.4%)**
	2017	$17.14/oz	$17.05/oz	$17.18/oz	$17.77/oz	$17.90/oz	$18.46/oz
a	Actual % Δ / APE %	**Neutral**	**(0.5%)**	**0.7%**	4.3%	5.0%	8.2%
	2018	$17.05/oz	$15.71/oz	$17.25/oz	$17.81/oz	$17.90/oz	$18.49/oz
a	Actual % Δ / APE %	**Down**	**(8.5%)**	**9.8%**	13.4%	13.9%	17.7%
	2019	$15.71/oz	$16.21/oz	$15.88/oz	$16.28/oz	$16.50/oz	$16.61/oz
a	Actual % Δ / APE %	**Up**	**3.1%**	(2.0%)	**0.4%**	1.8%	2.5%
	2020	$16.21/oz	$20.55/oz	$17.69/oz	$18.21/oz	$18.40/oz	$18.95/oz
a	Actual % Δ / APE %	**Up**	**21.1%**	(13.9%)	(11.4%)	(10.5%)	**(7.8%)**
	2000-2020	$14.46/oz	$15.19/oz	$14.23/oz	$15.13/oz	$15.10/oz	$16.01/oz
b	Actual % Δ / APE %	**Neutral**	**18.6%**	13.0%	11.8%	11.8%	12.5%
c	Root Mean Square error	4.41		3.06	2.79	2.68	3.19
d	Theil U(II) – Median	164.3%		113.9%	104.0%	100.0%	118.7%
e	Theil U(II) – Prior	100.0%		69.3%	63.3%	60.9%	72.3%
f	ReLMAE - Prior			128.2%	99.4%	101.1%	88.2%

Annexure J: LBMA-PMS Platinum Forecasts - Consensus Estimates

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LBMA average price for the prior period	LBMA average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	2002	$529/oz	$540/oz	$435/oz	$463/oz	$450/oz	$499/oz
a	Actual % Δ / APE %	**Up**	**2.1%**	(19.5%)	(14.4%)	(16.7%)	**(7.6%)**
	2003	$540/oz	$693/oz	$592/oz	$609/oz	$615/oz	$634/oz
a	Actual % Δ / APE %	**Up**	**22.0%**	(14.6%)	(12.0%)	(11.2%)	**(8.5%)**
	2004	$693/oz	$846/oz	$775/oz	$811/oz	$815/oz	$840/oz
a	Actual % Δ / APE %	**Up**	**18.1%**	(8.4%)	(4.1%)	(3.7%)	**(0.7%)**
	2005	$846/oz	$897/oz	$774/oz	$814/oz	$813/oz	$860/oz
a	Actual % Δ / APE %	**Up**	**5.7%**	(13.8%)	(9.3%)	(9.5%)	**(4.2%)**
	2006	$897/oz	$1,143/oz	$926/oz	$984/oz	$981/oz	$1,039/oz
a	Actual % Δ / APE %	**Up**	**21.5%**	(19.0%)	(13.9%)	(14.2%)	**(9.1%)**
	2007	$1,143/oz	$1,305/oz	$1,100/oz	$1,172/oz	$1,173/oz	$1,225/oz
a	Actual % Δ / APE %	**Up**	**12.4%**	(15.7%)	(10.2%)	(10.1%)	**(6.1%)**
	2008	$1,305/oz	$1,575/oz	$1,494/oz	$1,557/oz	$1,545/oz	$1,620/oz
a	Actual % Δ / APE %	**Up**	**17.2%**	(5.2%)	**(1.1%)**	(1.9%)	2.9%
	2009	$1,575/oz	$1,207/oz	$944/oz	$996/oz	$1,005/oz	$1,052/oz
a	Actual % Δ / APE %	**Down**	**(30.4%)**	(21.8%)	(17.5%)	(16.8%)	**(12.9%)**
	2010	$1,207/oz	$1,611/oz	$1,470/oz	$1,558/oz	$1,565/oz	$1,625/oz
a	Actual % Δ / APE %	**Up**	**25.1%**	(8.8%)	(3.3%)	(2.9%)	**0.9%**
	2011	$1,611/oz	$1,721/oz	$1,775/oz	$1,813/oz	$1,810/oz	$1,875/oz
a	Actual % Δ / APE %	**Up**	**6.4%**	**3.2%**	5.4%	5.2%	9.0%
	2012	$1,721/oz	$1,552/oz	$1,525/oz	$1,624/oz	$1,630/oz	$1,719/oz
a	Actual % Δ / APE %	**Down**	**(10.9%)**	**(1.7%)**	4.7%	5.0%	10.8%
	2013	$1,552/oz	$1,486/oz	$1,648/oz	$1,682/oz	$1,690/oz	$1,711/oz
a	Actual % Δ / APE %	**Down**	**(4.4%)**	**10.9%**	13.2%	13.7%	15.1%
	2014	$1,486/oz	$1,384/oz	$1,468/oz	$1,490/oz	$1,483/oz	$1,537/oz
a	Actual % Δ / APE %	**Down**	**(7.4%)**	**6.0%**	7.7%	7.1%	11.0%
	2015	$1,384/oz	$1,053/oz	$1,265/oz	$1,294/oz	$1,284/oz	$1,338/oz
a	Actual % Δ / APE %	**Down**	**(31.4%)**	**20.1%**	22.8%	21.9%	27.0%
	2016	$1,053/oz	$989/oz	$846/oz	$911/oz	$901/oz	$980/oz
a	Actual % Δ / APE %	**Down**	**(6.6%)**	(14.5%)	(7.8%)	(8.9%)	**(0.9%)**
	2017	$989/oz	$949/oz	$990/oz	$1,014/oz	$1,015/oz	$1,035/oz
a	Actual % Δ / APE %	**Down**	**(4.2%)**	**4.4%**	6.9%	7.0%	9.1%
	2018	$949/oz	$880/oz	$960/oz	$1,000/oz	$1,012/oz	$1,053/oz
a	Actual % Δ / APE %	**Down**	**(7.9%)**	**9.2%**	13.7%	15.0%	19.7%
	2019	$880/oz	$864/oz	$830/oz	$851/oz	$855/oz	$890/oz
a	Actual % Δ / APE %	**Down**	**(1.8%)**	(3.9%)	(1.5%)	**(1.0%)**	3.0%
	2020	$864/oz	$886/oz	$948/oz	$1,005/oz	$993/oz	$1,084/oz
a	Actual % Δ / APE %	**Up**	**2.5%**	**7.0%**	13.5%	12.1%	22.3%
	2002-2020	$1,117/oz	$1,136/oz	$1,093/oz	$1,139/oz	$1,139/oz	$1,190/oz
b	Actual % Δ / APE %	**Neutral**	**13.0%**	10.2%	9.1%	9.1%	9.3%
c	Root Mean Square error	189.61		134.02	119.71	119.07	132.55
d	Theil U(II) – Median	159.2%		112.6%	100.5%	100.0%	111.3%
e	Theil U(II) – Prior	100.0%		70.7%	63.1%	62.8%	69.9%
f	ReLMAE - Prior			91.9%	78.4%	77.8%	73.6%

Annexure K: LBMA-PMS Palladium Forecasts - Consensus Estimates

Forecasting period / % Change (Δ) compared to prior period / Consensus measure APE %		LBMA average price for the prior period	LBMA average price for the period	Lower Quartile	Average	Median	Upper Quartile
Row	**Column No.**	**1**	**2**	**3**	**4**	**5**	**6**
	2001	$279/oz	$279/oz	$285/oz	$298/oz	$295/oz	$305/oz
a	Actual % Δ / APE %	**Neutral**	**0.2%**	**2.1%**	6.7%	5.7%	9.3%
	2001	$279/oz	$271/oz	$268/oz	$276/oz	$275/oz	$283/oz
a	Actual % Δ / APE %	**Down**	**(3.0%)**	**1.1%**	1.8%	1.5%	4.3%
	2002	$271/oz	$310/oz	$280/oz	$287/oz	$290/oz	$293/oz
a	Actual % Δ / APE %	**Up**	**12.5%**	9.6%	7.3%	6.4%	**5.6%**
	2003	$310/oz	$363/oz	$335/oz	$345/oz	$345/oz	$351/oz
a	Actual % Δ / APE %	**Up**	**14.8%**	7.8%	5.1%	5.0%	**3.5%**
	2004	$363/oz	$409/oz	$403/oz	$417/oz	$419/oz	$429/oz
a	Actual % Δ / APE %	**Up**	**11.2%**	**1.5%**	2.0%	2.4%	4.7%
	2005	$409/oz	$444/oz	$422/oz	$434/oz	$438/oz	$450/oz
a	Actual % Δ / APE %	**Up**	**7.9%**	5.1%	2.4%	1.5%	**1.3%**
	2006	$444/oz	$604/oz	$520/oz	$535/oz	$525/oz	$554/oz
a	Actual % Δ / APE %	**Up**	**26.4%**	13.9%	11.4%	13.1%	**8.2%**
	2007	$604/oz	$695/oz	$620/oz	$652/oz	$650/oz	$675/oz
a	Actual % Δ / APE %	**Up**	**13.2%**	10.8%	6.2%	6.5%	**3.0%**
	2008	$695/oz	$872/oz	$825/oz	$862/oz	$850/oz	$893/oz
a	Actual % Δ / APE %	**Up**	**20.3%**	5.4%	**1.1%**	2.5%	2.4%
	2009	$872/oz	$972/oz	$836/oz	$881/oz	$901/oz	$946/oz
a	Actual % Δ / APE %	**Up**	**10.3%**	14.1%	9.4%	7.4%	**2.7%**
	2010	$972/oz	$1,225/oz	$1,165/oz	$1,199/oz	$1,199/oz	$1,233/oz
a	Actual % Δ / APE %	**Up**	**20.6%**	4.8%	2.1%	2.1%	**0.7%**
	2011	$1,225/oz	$1,572/oz	$1,449/oz	$1,457/oz	$1,464/oz	$1,491/oz
a	Actual % Δ / APE %	**Up**	**22.1%**	7.8%	7.3%	6.8%	**5.1%**
	2012	$1,572/oz	$1,669/oz	$1,728/oz	$1,766/oz	$1,770/oz	$1,833/oz
a	Actual % Δ / APE %	**Up**	**5.8%**	**3.5%**	5.8%	6.1%	9.8%
	2013	$1,669/oz	$1,411/oz	$1,714/oz	$1,753/oz	$1,753/oz	$1,782/oz
a	Actual % Δ / APE %	**Down**	**(18.3%)**	**21.4%**	24.3%	24.2%	26.2%
	2014	$1,411/oz	$1,266/oz	$1,176/oz	$1,219/oz	$1,230/oz	$1,265/oz
a	Actual % Δ / APE %	**Down**	**(11.4%)**	7.2%	3.7%	2.9%	**0.1%**
	2015	$1,266/oz	$1,160/oz	$1,188/oz	$1,211/oz	$1,230/oz	$1,255/oz
a	Actual % Δ / APE %	**Down**	**(9.2%)**	**2.4%**	4.4%	6.0%	8.2%
	2016	$1,160/oz	$1,251/oz	$1,058/oz	$1,103/oz	$1,120/oz	$1,145/oz
a	Actual % Δ / APE %	**Up**	**7.3%**	15.5%	11.8%	10.5%	**8.5%**
	2017	$1,251/oz	$1,257/oz	$1,209/oz	$1,244/oz	$1,260/oz	$1,285/oz
a	Actual % Δ / APE %	**Neutral**	**0.5%**	3.8%	1.0%	**0.2%**	2.2%
	2018	$1,257/oz	$1,268/oz	$1,287/oz	$1,318/oz	$1,321/oz	$1,359/oz
a	Actual % Δ / APE %	**Neutral**	**0.9%**	**1.5%**	3.9%	4.1%	7.1%
	2019	$1,268/oz	$1,393/oz	$1,300/oz	$1,312/oz	$1,315/oz	$1,326/oz
a	Actual % Δ / APE %	**Up**	**8.9%**	6.7%	5.8%	5.6%	**4.8%**
	2020	$1,393/oz	$1,770/oz	$1,521/oz	$1,559/oz	$1,559/oz	$1,593/oz
a	Actual % Δ / APE %	**Up**	**21.3%**	14.0%	11.9%	11.9%	**10.0%**
	2000-2020	$903/oz	$974/oz	$933/oz	$959/oz	$962/oz	$988/oz
b	Actual % Δ / APE %	**Up**	**12.3%**	8.3%	7.2%	7.1%	**6.8%**
c	Root Mean Square error	160		113	106	**105**	107
d	Theil U(II) – Median	153.1%		107.9%	**101.4%**	100.0%	102.2%
e	Theil U(II) – Prior			70.5%	66.2%	**65.3%**	66.7%
f	ReLMAE - Prior			104.4%	82.6%	82.3%	**65.9%**