EBVM

The European Business Valuation Magazine

Articles

Data

News

Editors:

EACVA
EUROPEAN ASSOCIATION OF CERTIFIED VALUATORS AND ANALYSTS
SINCE 2005

IVSC

Sponsors:

SMART ZEBRA

In this issue

Prof. Dr. Stefan O. Grbenic, StB, CVA

Professor of Management Control, Accounting and Finance at Graz University of Technology and Webster University St. Louis/Vienna and Guest/Visiting Professor at University of Maribor and Istanbul Medeniyet University as well as former Visiting Professor at University of Twente. He acts as a referee for academic journals and as a member of scientific com-mittees at numerous international scientific conferences. He has published papers in major scientific journals, books and actively participated at scientific conferences around the world. Furthermore, he is Certified Tax Consultant (financial advisory), Expert Witness and Certified Valuation Analyst (EACVA/NACVA).

Prof. ddr. Timotej Jagrič, CQRM

Head of the Institute of Finance and Artificial Intelligence and Head of the international Master Data Science in Business at University of Maribor and Professor at University of Primorska as well as Guest/Visiting Professor at Graz University of Tech-nology and former Visiting Researcher at Humboldt University Berlin. He has published numerous papers in major scientific journals, acts as a referee for academic journals and actively participated at scientific conferences around the world. During his professional career, he worked as chairmen of the super-visory board of the largest re-insurance company in Slovenia, as member of the Strategic council of the government of the Republic of Slovenia, and as advisor for many government and other public institutions in the CEE region. He was elected for the vice governor of the Slovenian Central Bank. Currently, he is member of the steering committee of the Slovenian Risk Association and acts as an expert for WHO.

Artificial Intelligence in Business Valuation

This article examines the application of Artificial Intelligence in business valuation. It is divided into two parts: This part one gives an overview on AI (especially Machine Learning) models and algorithms, model performance optimization, learning paradigms, data inference, algorithm tasks as well as types of variables captured, and de-scribes the basics of different types of common algorithms categorized according to functions and similarities. Built upon these basics, the upcoming part two provides an overview on the primary tasks AI (Machine Learning) al-gorithms have been applied in business valuation research currently emphasiz-ing predictive (rather than generative) algorithms, including their capability to predict company values directly (stand-alone or automated valuation approach), their capability to identi-fy explanatory variables in predicting company values, their performance in selecting peers forming a peer group, and their performance to develop information extraction systems of financial data for busi-ness valuation purposes. In general, the research results allow for the overall conclusion that AI (Machine Learning) techniques are a promising way to improve the business valuation profession.

Valuation of Intangible Assets – An Integrated Relief-from-Royalty Method and Monte Carlo Simulation Approach

This article explores the synergies between Relief-from-Royalty Method (RFR) and Mon-te Carlo Simulation (MCS), highlighting the advantages of this combined approach in the valuation of intangible assets and illustrating how it can enhance the strategic deci-sion-making process in the face of uncertainty. For illustration purposes, a full calcula-tion example for a technology-related intangible asset in Germany is presented after a detailed explanation of the methodological approach.

Dr. Christian Reichert, CVA, MBA
Würzburg, Germany

From the Editors

AI, Governance, and Expertise: Navigating the New Frontier in Business Valuation

The rapid evolution of artificial intelligence (AI) and digital transformation continues to reshape the business valuation profession. As valuers, we are challenged to adapt our methodologies, safeguard professional judgment, and embrace innovation responsibly. This issue places a strong focus on these developments and their implications for practice.

In July 2025, the International Valuation Standards Council (IVSC) published its perspective paper "Navigating the Rise of Artificial Intelligence in Valuation: Opportunities, Risks, and Standards." This paper sets a clear framework: AI offers significant potential for efficiency, deeper analysis, and broader access to data, but cannot replace the independent professional judgment of the valuer. The new IVS, effective January 2025, reinforce the principles of governance, transparency, and accountability—reminding us that automated valuation models must remain subject to expert oversight to ensure compliance and trust in outcomes.

Building upon this broader regulatory perspective, the article by Prof. Stefan O. Grbenic and Prof. Timotej Jagrič, "Artificial Intelligence in Business Valuation – Part I: Algorithms, Models, and Performance," provides a profound introduction into the mechanics of AI, with particular focus on machine learning algorithms and their applicability to valuation practice. Their contribution highlights the opportunities for greater efficiency and predictive power, while stressing the continued necessity of human expertise to interpret, guide, and ethically apply AI-generated insights.

Complementing this focus on innovation, Dr. Christian Reichert's article "Valuation of Intangible Assets – An Integrated Relief-from-Royalty Method and Monte Carlo Simulation Approach" addresses one of the most challenging areas of modern valuation practice: intangible assets. By combining the well-established Relief-from-Royalty method with Monte Carlo simulation, his work illustrates how probabilistic approaches can enhance the robustness of valuation outcomes and better support decision-making under uncertainty. The case study presented demonstrates the practical relevance of this methodology, especially in contexts with limited or uncertain data.

Both contributions underline the dual imperative facing our profession: to integrate new analytical tools responsibly while safeguarding the core values of independence, transparency, and sound judgment. I warmly encourage you to explore these articles in depth. They not only expand our technical toolkit but also inspire us to reflect critically on the future of valuation in an increasingly data-driven environment.

I invite you to engage with these thought-provoking contributions, discuss them with colleagues, and consider how their insights might inform and enrich your own professional practice. Together, by combining innovation with prudence, we can continue to strengthen the credibility and relevance of business valuation in a rapidly changing world.

Wolfgang Kniest, CVA
Managing Director and co-founder of EACVA

Sponsor:

SMART ZEBRA

Article

Artificial Intelligence in Business Valuation

Part I: Algorithms, Models, and Performance

This article examines the application of Artificial Intelligence in business valuation. It is divided into two parts: This part one gives an overview on AI (especially Machine Learning) models and algorithms, model performance optimization, learning paradigms, data inference, algorithm tasks as well as types of variables captured, and describes the basics of different types of common algorithms categorized according to functions and similarities. Built upon these basics, the upcoming part two provides an overview on the primary tasks AI (Machine Learning) algorithms have been applied in business valuation research currently emphasizing predictive (rather than generative) algorithms, including their capability to predict company values directly (stand-alone or automated valuation approach), their capability to identify explanatory variables in predicting company values, their performance in selecting peers forming a peer group, and their performance to develop information extraction systems of financial data for business valuation purposes. In general, the research results allow for the overall conclusion that AI (Machine Learning) techniques are a promising way to improve the business valuation profession.

Professor Dr. Stefan O. Grbenic, StB, CVA

Professor of Management Control, Accounting and Finance at Graz University of Technology and Webster University St. Louis/Vienna and Guest/Visiting Professor at University of Maribor and Istanbul Medeniyet University as well as former Visiting Professor at University of Twente. He acts as a referee for academic journals and as a member of scientific committees at numerous international scientific conferences. He has published papers in major scientific journals, books and actively participated at scientific conferences around the world. Furthermore, he is Certified Tax Consultant (financial advisory), Expert Witness and Certified Valuation Analyst (EACVA/NACVA).

Professor ddr. Timotej Jagrič, CQRM

Head of the Institute of Finance and Artificial Intelligence and Head of the international Master Data Science in Business at University of Maribor and Professor at University of Primorska as well as Guest/Visiting Professor at Graz University of Technology and former Visiting Researcher at Humboldt University Berlin. He has published numerous papers in major scientific journals, acts as a referee for academic journals and actively participated at scientific conferences around the world.

Contact: ebvm@eacva.de

I. Introduction

Welcome to the AI revolution![1] It will challenge us to the extreme. It is time for professionals to acknowledge, apply and adjust to that transformation, determining the need to upskill and reskill. Artificial Intelligence will significantly impact virtually every industry, and the business valuation profession will be no exception. Although it is too early to predict precisely how AI will shape the business valuation profession, one thing is certain: "*If you do not think AI will transform our world, you are seriously wrong. And for those who do not jump on this train, you will be "left in the dust".*[2] AI may lead to some job losses, particularly those that involve routine repetitive tasks. It may support valuation analysts (and, thus, may change the valuation profession) by easing repetitive tasks concerning data collection and exploitation prior to any decision making or human judgement with process automation techniques, and guiding them when making decisions by providing a better understanding of the key drivers of a valuation opinion employing Machine Learning technologies and other statistical analyses. Therefore, from the current point of view, AI technology will (at least in the close future) not replace human judgment, but it will help and augment human judgment by providing efficiencies and further analysis depth.

AI may impact the business valuation profession directly in multiple ways, e. g. it will increase efficiency in data analysis, it will improve accuracy (reducing human error and increasing the precision of valuations by identifying patterns and correlations that may not be immediately apparent to human analysts), it will strengthen predictive analytics (leveraging predictive analytics to forecast future cashflows and earnings more accurately), it will lead to a standardization of procedures (standardizing valuation methodologies, leading to more consistent and reliable valuations and contributing to a reduction in subjectivity and bias in the valuation process), it will enable dynamic valuations (AI systems can continually update valuations in real time as new information becomes available), it will enable customization and scalability (helping valuation professionals customize their analysis for different industries and business models, and scale their services to handle a larger number of clients without a corresponding increase in errors or loss of quality), it will reduce cost of valuations (by automating routine tasks, allowing valuation professionals to focus on more complex aspects of the valuation process), it will cause a shift in skill sets (valuation professionals will need to become proficient with AI tools and understand how to interpret and complement AI-generated analyses, shifting the required skill sets within the profession, prioritizing data science and analytical skills alongside traditional finance and accounting knowledge), it will improve market analysis (analysing market sentiment and macroeconomic trends), and it will improve compliance and due diligence (AI systems can assist in ensuring that valuations are compliant with relevant standards and regulations by automatically checking for inconsistencies or red flags that require closer inspection).[3]

AI is a promising source of value creation for valuation analysts, since it reduces the cost of routine-tasks (data collection, data analysis, data processing, validation checks, modelling, model review), shifts human expertise to the more complex and, correspondingly, higher value-added tasks, allows for a better management of time, and, allows for a better use and leverage of information.[4] The key will be to strike a balance between AI-driven insights and the unique, irreplaceable qualities of human intuition, creativity, and ethical judgment.[5] Furthermore, valuation professionals will face new challenges in valuing AI itself as part of the business valuation process.

II. AI, ML, Models, and Algorithms Defined

There are many definitions of Artificial Intelligence (AI). Most of the definitions cluster around building machines that can perform tasks that are typically performed by humans (requiring intelligence when performed) acting like humans. AI is a field that combines computer science and robust datasets to enable problem solving. Demonstrated by machines, intelligence – perceiving, synthesizing, and inferring information – encompasses the ability to learn and to reason, to generalize, and to infer meaning. Machines mimic cognitive functions associated with human minds. Thereby, cognitive functions include all aspects of learning, perceiving, problem solving, and reasoning. An intelligent entity pursues and achieves goals in uncertain situations, ideally learning and constantly improving, accumulating experience, and learning from mistakes, and, thus, improving with each try. AI is a technology that can pursue goals in uncertain environments.

AI systems are designed to operate with varying levels of autonomy. Consequently, different levels of AI are categorized. Weak AI (Narrow AI, Artificial Narrow Intelligence – ANI) is trained and focused to perform specific tasks in a specific narrow knowledge domain. Weak AI drives most of the common AI applications that surround us today.[6] Strong AI is made up of Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI). Artificial General Intelligence (General AI) is a theoretical form of AI where a machine would have an intelligence equaled to humans, having a self-aware consciousness that can solve problems, learn, and plan into the future. It can perform human expert tasks in multiple domains and is able to transfer knowledge across domains and adopt autonomously. Artificial Super Intelligence (Super Intelligence) would surpass the intelligence and ability of the human brain.

---

1. Although, it is evident that AI has been around since the middle of the twentieth century. The term "Artificial Intelligence" was first mentioned in 1955 by McCarthy/Minsky/Rochester/Shannon, A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. The field of AI was formally created in 1956.
2. Black, AI will Quickly Transform our World – Ignore it at Your Peril, NACVA Association News, 2023. This fully coincides with the opinion of the authors.

3. According to ChatGPT; see Shiffrin, Artificial Intelligence Promises to Transform the BVFLS Profession, The Value Examiner, no. 6 (2023): 4-6.
4. See correspondingly Magnan/Parija, How Will Technology Change the Way Business Valuations Are Performed?, The European Business Valuation Magazine, vol. 2, no. 2 (2022): 24-28.
5. Kreuter/Castillo, AI Odyssey: Beyond the Hype, NACVA Quickread, 2023.
6. E. g. advanced web search engines, recommendation engines (discovering data trends), generative or creative tools (e. g. ChatGPT, Google Gemini, DeepSeek, Claude, MS Copilot, Copilot 365 and Copilot Studio, MS Azure AI Studio, Elicit, Consensus, SciSpace, AI art), speech recognition engines (processing human speech into a written format), customer service engines (online virtual agents), automated stock trading engines, computer vision (deriving meaningful information from visual inputs), automated decision-making tools, strategic game systems, and self-driving cars.

Machine Learning (ML) is a major approach to realize AI by learning from, and making data-driven predictions based on learned experiences. It is devoted to understanding and building methods that let machines learn to improve computer performance on some set of predefined tasks. ML methods are about extracting models from data and fitting the models to the data. ML is teaching machines to develop the ability to recognize patterns by either identifying variables and/or identifying relationships between variables in a dataset and then use that knowledge to generalize beyond the dataset. ML algorithms build a model based on sample data to make predictions or decisions without being explicitly programmed doing so. They work on the assumption that strategies, algorithms, and inferences that worked well in the past are likely to continue working well in the future.

Models and algorithms can be separated into learned representation and process of learning. Hence, the model represents the specific representation that satisfies some goal learned from data, and the algorithm defines the process of learning (i. e., the methods by which machines learn to "think"). ML algorithms learn a target function f that best maps input variables X (features) to an output variable Y (label, target variable) including an error ε that is independent from the input data:[7]

$$Y=f(X)+\varepsilon$$

The relationship in the data (that is, the relationship between the features – properties or attributes – among themselves and with the target variable) is called statistical inference. Different ML algorithms make different assumptions about the shape and structure of the mapping function and how best to optimize the representation to approximate. Since all algorithms have its strengths and weaknesses, no single algorithm works best on all learning problems. Therefore, a set of different algorithms must be tried to each learning problem, since the superior algorithm is not known a priori.[8]

III. Model Performance Optimization

Generalization refers to how well the concept learned by the ML model generalizes to unseen data and is capable making predictions on it. The cause of poor generalization (that is, the model fit is not appropriate) is either overfitting or underfitting the training data.

Overfitting refers to the ML model learning (modeling) the training data too well at the expense of poorly generalizing to new data. The model learns the detail and noise in the training data as concepts to the extent that it negatively impacts its performance. Consequently, an overfitted model uses a more complicated approach than is ultimately optimal and contains more parameters than are justified by the data, extracting

some of the residual variation (i. e., the noise) as if that variation represented underlying model structure. The model memorizes training data rather than learning to generalize from a trend and may produce outputs that are virtually identical to instances from the training data set.[9]

To limit overfitting, various techniques are available, e. g. (i) explicitly penalizing overly complex models; (ii) holding back a validation dataset (which is assumed to approximate the typical unseen data that a model will encounter) and testing the ability of the model to generalize by evaluating its performance on that data not used for training, or (iii) resampling the data set and cross-validating the model on the different subsets of the training data to estimate model accuracy.[10]

Underfitting refers to a ML model that can neither model the training data nor generalize to new data, that is, failing to learn the problem from the training data sufficiently. The underlying structure of the data is not adequately captured with parameters or terms that would appear in a correctly specified model.

To limit underfitting, (i) the complexity of the model may be increased by adding more features, increasing the number of parameters, or utilizing a more flexible model; (ii) the amount of training data may be increased, allowing the model to better capture the underlying patterns in the data; (iii) the model may be regularized by adding a penalty term to the loss function that discourages large parameter values; (iv) ensemble methods may be used, combining various models to work together; (v) feature engineering may be used, creating new (more relevant) model features.

IV. Categorizing Machine Learning Algorithms

ML algorithms may be categorized, among others, according to the dimensions learning paradigm utilized, data inference assumed, task performed, as well as type of variables employed:

Figure 1: Categorization of ML algorithms

Learning Paradigms	Supervised learning, unsupervised learning, semi-supervised learn-ing, re-inforcement learning, deep learning, etc.
Data Inference	Parametric (linear) algorithms, nonparametric (nonlinear) algorithms
Tasks	Regression, classification, clustering, dimensionality reduction
Variable Type	Continuous variables, categorical variables

---

7 Algorithms have certain properties: (i) It is a mechanism to discover mathematical functions that express relationships between variables, (ii) it is goal oriented, (iii) it has a performance benchmark, since it is discovering the "best" function, (iv) it starts with random feature values, and, finally (v) it must work efficiently.

8 This is called the „No Free Lunch (NFL) Theorem"; Wolpert/Macready, No Free Lunch Theorems for Optimization, IEEE Transactions on Evolutionary Computation, vol. 1, no. 1 (1997): 67-82.

---

9 Overfitting is the use of models or procedures that violate "Occam's razor" or the "law of parsimony", implying that any complex function is a priori less probable than any simple function.

10 Related techniques are e. g. model comparison, k-fold cross-validation, regularization, early stopping, pruning, Bayesian priors, or simply dropout.

1. Learning Paradigms

Depending on the nature of the signal (feedback) available to the learning system, various learning paradigms (learning types) are categorized, and, consequently, machines may learn in various ways.

In **Supervised Learning**, the algorithm collects a series of input data (predictor variables) and output responses (labels) and builds a mathematical model aiming to learn a general rule (mapping function, target function) that maps inputs to their desired outputs (supervisory signals) given by a "teacher". Consequently, the machine learns by teacher-provided examples. The mapping function should be capable of predicting the corresponding output variables to new input data. The process of an algorithm learning from the training dataset can be thought of as a teacher supervising the learning process. Since all data is labeled, that is, the correct answers are known (for each set of inputs the output is designated), the algorithm iteratively makes predictions on the training data and is corrected by the teacher. Learning stops when the algorithm achieves an acceptable level of performance.

In the mathematical model, each training example is represented by a feature vector, and the training data is represented by a matrix. Through iterative optimization, the algorithm learns a function that can be used to predict the output responses (labels) associated with new inputs. An optimal function will allow the algorithm to correctly determine the output for previously unknown inputs (that were not a part of the training data). The cost function is related to eliminating incorrect deductions of outputs.

There are four major issues to consider in supervised learning:[11]

1. First, the bias-variance trade-off, that is, the trade-off between the strength of the assumptions made by the model (flexibility) and its sensitivity to the training data.
2. Second, the complexity of the mapping function and the amount of training data needed to train the model; if the "true" function is simple, then an inflexible learning algorithm with high bias and low variance can learn from a small amount of data; in contrast, if the function is highly complex, then the algorithm will only be able to learn using a large amount of training data paired with a flexible learning algorithm with low bias and high variance.
3. Third, the dimensionality of the input space, that is, if the input feature vectors have large dimensions, learning the function can be difficult even if the true function only depends on a small number of those features (since too many dimensions can confuse the learning algorithm).

4. Finally, fourth, the noise in the output values, that is, if the desired output values are often incorrect, the learning algorithm may not be able to find a function that exactly matches the training examples, increasing the risk of overfitting.

Consequently, generalization is probabilistic and not guaranteed, especially for complex and/or noisy datasets.

In **Unsupervised Learning**, the algorithm searches for similarity patterns in the input data, without the need to give the algorithm any output (response) data. Hence, unlabeled input data is given to the learning algorithm (that is, there is no correct answers and there is no teacher), leaving it on its own to find structure in the input data. The goal is to model the underlying structure or distribution in the data to learn the respective patterns.

Semi-Supervised Learning falls between supervised learning (with completely labeled training data) and unsupervised learning (without any labeled training data). It combines a small amount of labeled input data with a large amount of unlabeled data during training. To make use of unlabeled data, some relationship to the underlying distribution of the data must exist. Therefore, semi-supervised learning algorithms make use of at least one of the following assumptions: (i) The Continuity (Smoothness) assumption assumes that data points being close to each other are more likely to share a label; this yields a preference for decision boundaries in low-density regions. (ii) The Cluster assumption assumes that, since data tend to form discrete clusters, data points in the same cluster are more likely to share a label. (iii) The Manifold assumption assumes that data lie approximately on a manifold of lower dimension as compared to the input space; hence, learning the manifold employing both, the labeled and unlabeled data, can avoid the curse of dimensionality (the learning can proceed using distances and densities defined on the manifold).

Reinforcement Learning works with rewards and punishments where the algorithm learns through success and failure. The algorithm is guided by a performance score. An agent is trained to make decisions based on a reward system and is provided feedback that is analogous to rewards. The objective is to learn an optimal policy that maximizes the reward function or other provided reinforcement signal that accumulates from the immediate rewards. The agent interacts with its surrounding and learns via trial and error. The purpose is to learn how to optimize its rewards over time, although optimizing the rewards does not guarantee optimal learning especially in high-dimensional environments with sparse rewards. Reinforcement Learning differs from Supervised Learning in not needing labelled data as well as sub-optimal actions to be explicitly corrected. Instead, the focus is on finding a balance between exploration (of uncharted territory) and exploitation (of current knowledge).

Deep Learning aims at discovering multiple levels of representation, or a hierarchy of features, with higher-level, more abstract features defined in terms of lower-level features. It connects artificial, software-based calculators that approximate the function of brain neurons. Deep learning architectures are based on multi-layered artificial neural networks with

---

11 Further factors to consider when choosing and applying a learning algorithm include the heterogeneity of the data (if the feature vectors include features of many different kinds (discrete, discrete ordered, counts, continuous values), some algorithms are easier to apply than others), the redundancy in the data (if the input features contain redundant information (e. g., highly correlated features), some learning algorithms will perform poorly because of numerical instabilities), or the presence of interactions and non-linearities (if each of the features makes an independent contribution to the output, then algorithms based on linear functions and distance functions generally perform well).

each layer transforming data nonlinearly, allowing hierarchical feature extraction from raw input to abstract representations. Each level learns to transform its input data into a slightly more abstract and composite representation. Deep learning is built on a set of sophisticated algorithms that learn to extract and predict complicated patterns from massive volumes of data. In contrast to the "traditional" machine learning paradigms, unstructured data is not required to be pre-processed and organized into a structured format, since deep learning algorithms adjust and fit themselves for accuracy.

Additionally, further learning paradigms have been developed, e. g. Self Learning, Feature Learning, Anomaly Detection, or Association Rule Learning.

2. Data Inference

According to data inference, parametric and nonparametric ML algorithms are categorized.

Parametric ML algorithms make (strong) assumptions about the form of the mapping function which can be either linear or nonlinear, trying to adjust the response of the model by means of this mapping function (representing the relation between input and output variables). This makes these algorithms simple (easy to understand and to interpret results), faster learning, and, they require less data and can work well even if the fit to the data is not perfect. The benefits come at the expense of reduced performance due to being constrained to the linear form (by choosing the linear functional form a priori, they are relatively unlikely to match the "real" underlying mapping function) and only capturing limited model complexity (hence, they are more suited to simpler problems).

Nonparametric ML algorithms do not make (strong) assumptions about the form of the mapping function and, hence, are free to learn any functional form from the training data. They seek to best fit the training data in constructing the mapping function simultaneously maintaining some ability to generalize to unseen data. They are flexible (capable of fitting manifold functional forms) and have a higher model complexity that results in more powerful models. The benefits come at the expense of requiring larger sets of training data, they are slower to train (since more parameters must be trained), they carry the risk to overfit the training data and, the results are harder to interpret and explain.

ML algorithms try to best estimate the mapping function. Since every mapping causes an error, the algorithm tries to minimize this error resulting in a Bias-Variance Trade-Off (Bias-Variance Dilemma). The resulting prediction error can be broken down into three parts: (i) The irreducible error refers to the error introduced from the chosen framing of the problem and may be caused by factors like unknown features. It cannot be reduced regardless of what algorithm is used. (ii) The bias error refers to the simplifying (erroneous) assumptions made by the model to make the mapping function easier to learn. Low (high) bias suggests less (strong) assumptions about the form of the mapping function. (iii) The variance error refers to the sensitivity of a model to fluctuations in the training data. ML algorithms that have a high (low) variance are strongly (weakly) influenced by the specifics of the training data, exercising a strong (weak) im-

pact on the number and types of features employed to characterize the mapping function.

The goal is to achieve both, low bias, and low variance, simultaneously. But there is a trade-off between bias and variance (the "bias-variance dilemma"), since increasing (decreasing) bias will decrease (increase) variance automatically (although the trade-off is not always automatic and/or linear).[12] Parametric (nonparametric) ML algorithms often have a high (low) bias but a low (high) variance. Parametric algorithms have a high bias making them fast to learn and easier to understand but less flexible. In turn they have lower predictive performance on complex problems that fail to meet the simplifying assumptions, and they are exposed to fail to capture important regularities (i. e., to underfit the data). In contrast, nonparametric ML algorithms have a low bias making them flexible so that they can fit the data well. But if the algorithm is too flexible, it will fit each training data set differently, making it exposed to overfit to noisy and unrepresentative training data.

3. Tasks

According to tasks, regression, classification, clustering, and dimensionality reduction ML algorithms may be categorized. Regression algorithms capture outputs that may have any numerical value within a range. Regression is used to predict continuous outputs (responses). Classification algorithms capture categorical outputs restricted to a limited set of values (binary, multiclass or multilabel classification). Classification is used to predict discrete outputs (responses). Clustering algorithms discover inherent groupings in the data where no assumptions are made about the likely relationships within the data. Dimensionality Reduction algorithms aim at reducing the dimension of the feature set by obtaining a set of principal input variables, transforming the features via mathematical operations.

4. Variable Type

ML algorithms may be further categorized according to the type of variables employed. Continuous variables are obtained by measuring and can take an uncountable set of values. Continuous random variables are described statistically by a continuous distribution.[13] Categorical variables can take exactly two possible values (binary variable, dichotomous variable, Bernoulli variable) or a limited number of possible values (polytomous variables), assigning each unit of observation to a particular group or nominal category. Categorical random variables are normally described statistically by a categorical distribution.[14]

5. Further Categorization Criteria

Furthermore, ML algorithms may be categorized according to their average predictive accuracy, the training time need-

---

12 The bias-variance dilemma is a core problem especially for supervised ML algorithms.

13 Examples for continuous distributions are the Normal distribution, Student's t-distribution, F-distribution, Chi-squared distribution, Log-normal distribution, Pareto distribution, Poisson distribution, Exponential distribution, or the Gamma distribution.

14 Examples for categorical distributions are the Bernoulli distribution, Binomial distribution, Negative binomial distribution, Multinomial distribution, Beta distribution, Beta-binomial distribution, Geometric distribution, Hypergeometric distribution, Dirichlet distribution, or the Wishart distribution.

ed, the prediction speed achieved, the amount of parameter tuning needed, their performance with only a limited number of observations available, the handling of irrelevant features, their capability of automatically learning feature interaction, the scaling of features needed and, the interpretability and explainability of the results generated.

V. Types of Machine Learning Algorithms according to Functions and Similarities

According to functions and similarities, ML algorithms may be categorized into the following types:

Figure 2: Algorithm Categorizations According to Functions and Similarities
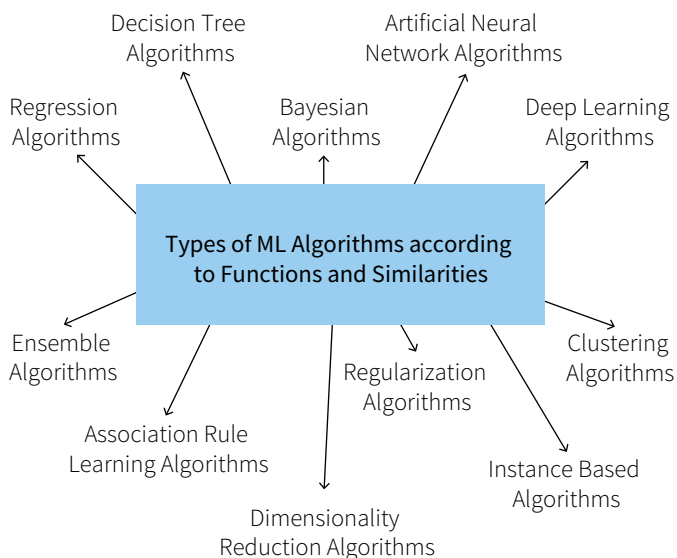
1. Regression Algorithms

Regression methodology is employed in various types: (i) Linear Regression (assuming a linear relationship between the features (explaining variables) and the continuous output (explained) variables), (ii) Regularized linear regression (simultaneously minimizing the sum of squared error of the model and reducing its complexity),[15] (iii) Logistic regression (intended for binary classification problems), and (iv) Linear Discriminant Analysis (LDA; intended for multi-class classification problems).[16]

2. Decision Tree Algorithms

Decision tree (DT) algorithms use a decision tree as a predictive model to predict the value of an item (represented in the leaves) from a set of observations about the item (represented in the branches). It is a supervised learning approach. DTs where the target variable can take a discrete set of values are

called classification trees, tree models where the target variable can take continuous values are called regression trees.[17]

A DT[18] is built by splitting the source set, constituting the root node of the tree, into subsets, which constitute the successor children. The top is known as the root and the bottom nodes are known as leaf nodes. DTs start with a basic question from where to ask a series of questions to reach a final answer. Starting at the root node, a question is asked about a feature and, depending upon the answer, the data set is split into branches. Then another question is asked about the next feature and the data set is again split. The splitting is based on a set of splitting rules based on classification features. This process is repeated on each derived subset in a recursive manner (recursive partitioning). The recursion is completed when the subset at a node includes similar values of the target feature, or when splitting no longer adds value to the predictions.

DT algorithms generally work top-down by choosing a variable at each step that best splits the set of items. To measure the quality of the split, various metrics that generally measure the homogeneity of the target feature within the subsets, may be employed, e. g. the Estimate of Positive Correctness, Entropy, Information Gain, Gini Impurity, the Variance Reduction or the Measure of Goodness.

DTs face various advantages, but also suffer from some drawbacks. Major advantages of DTs are that they are easy to understand and the results are simple to interpret by using a white-box (open-box) model, they can handle both, continuous and categorical data, they require only little data preparation (normalization, that is, making values across features comparable), feature importance is easy to determine and they automatically rank features according to their importance (the hierarchy of features reflects their importance), the model can be validated using statistical tests, they work well for large data sets, they can handle mussing values, and they are insensitive to underlying relationships between features. These advantages come at the expense of some limitations since DTs are not robust to changes in the training data (they are exposed to variance errors), they cannot guarantee globally optimized results (although locally optimal decisions/splits are made at each node), they are prone to overfitting the data, the average depth of the tree is not guaranteed to be minimal, and, information gain is biased in favor of features with more levels (for data including categorical variables with different numbers of levels).

3. Bayesian Algorithms

Bayesian algorithms are based upon Bayesian theorem, assuming that each feature is independent from the remaining

---

15 Common types of regularization procedures for linear regression are e. g. (i) Lasso Regression (L1 regularization) modifying OLS to also minimize the absolute sum of the coefficients, and (ii) Ridge Regression (L2 regularization) modifying OLS to also minimize the squared absolute sum of the coefficients.

16 Common extensions of the Linear Discriminant Analysis are (i) Quadratic Discriminant Analysis, (ii) Flexible Discriminant Analysis (using nonlinear combinations of inputs), and (iii) Regularized Discriminant Analysis (moderating the impact of different features).

17 The term Classification and Regression Tree (CART) Analysis is an umbrella term for both types of DTs.

18 Ensemble methods construct more than one decision tree: Boosted trees build an ensemble by training each new instance to emphasize the training instances previously mis-modeled. Bootstrap aggregated (or bagged) decision trees build multiple decision trees by repeatedly resampling training data with replacement, and voting the trees for a consensus prediction. Rotation Forests train every decision tree by first applying Principal Component Analysis (PCA) on a random subset of the features.

ones. It provides a way that the probability of hypotheses may be computed based on prior knowledge.

Common Bayesian algorithms are (ii) the Naïve (Idiot) Bayes algorithm (performing classification for binary as well as multiclass classification problems by calculating the probability of different classes or outcomes based on previously known data; it first estimates the probability of the feature in a class and then multiplies the probabilities to determine the influence of features in a class); (ii) the Gaussian Naïve Bayes algorithm (is an extension of the Naïve Bayes algorithm to real-valued attributes by assuming a Gaussian (Normal) distribution); (iii) Bayesian Belief Networks (BBNs) (are similarly based upon Bayesian theorem, except considering the probabilities and dependencies among features).

4. Artificial Neural Network Algorithms and Deep Learning Algorithms

Neural networks are not only another method, they are a different way of approaching problems, emulating the functionality of the human brain. Neural networks learn to perform tasks by considering examples, generally without being programmed with any task-specific rules. Each example contains a known input and a result, forming probability-weighted associations, which are stored within the data structure of the net itself.

A neural network is a collection of connected (the connections are called edges) neurons (also known as nodes) that transmit signals (information) from one neuron to another. Typically, the neurons are organized into (multiple) layers. Different layers may perform different kinds of transformations on their inputs. Neurons of one layer connect only to neurons of the immediately preceding and immediately following layers. The input layer receives external data, the output layer produces the ultimate result. In between them are the hidden layers.

Deep Learning Algorithms generally come as multi-layered (consisting of multiple hidden layers) neural networks, called Deep Neural Networks (DNNs) or Artificial Neural Networks (ANNs).[19] They are trained on large amounts of data to identify and classify phenomena, recognize patterns and relationships, evaluate possibilities, and make predictions and decisions, with the additional hidden layers helping to refine and optimize the outcomes for greater accuracy. They allow machines to perform sophisticated tasks by developing (unlike other neural networks) the ability to understand which features are more relevant for predicting the output and extracting different details about the data at different levels.

As the data is passed to the input layer, the neurons get activated. The signals received are processed and the output is computed by an activation function (of the sum of its inputs). Activation functions shape the outputs of the neurons. They decide whether a neuron should be activated or not (according to

importance of the neuron's input to the network).[20] Afterwards, the neurons of the input layer pass the data to the next layer. But before passing it to the next layer, they tune themselves (randomly) in relation to the data they just received. This layer, just like the previous layer, tunes itself and then forwards the data to the next layer. This continues until the data is passed to the output layer predicting the value. This predicted value is compared to the actual target value, and the difference between the two determines the feedback that is provided to the neurons to enable them improving their tuning (this back-feeding into the network to fine-tune is known as back propagation).[21]

The loss function (cost function) evaluates the ability of the model to predict the expected output. It is to be minimized. Common loss functions are:

- For classification tasks: (i) the Binary Cross-Entropy Loss/ Log Loss function, measuring the performance where the predicted output is a probability value between 0 and 1, and (ii) the Hinge Loss function, penalizing both, wrong predictions and true predictions not being confident.
- For regression tasks: (i) the Mean Square Error (MSE, Quadratic Loss, L2 Loss) function, measuring the average of squared differences between the actual and the value predicted by the model; it penalizes the model for making large errors by squaring, making the loss function less robust to outliers; (ii) the Mean Absolute Error (MAE, L1 Error) function, measuring the average of absolute differences between the actual and the value predicted by the model (without considering their directions); it is more robust to outliers as compared to the MSE loss function; (iii) the Huber Loss function, a combination of the MSE and the MAE loss function, defined as MAE for larger errors and MSE for smaller errors; (iv) the Log-Cosh Loss function, measuring the logarithm of the hyperbolic cosine of the prediction error; and (v) the Quantile Loss function, measuring the average errors of quantiles.

---

20 Common (non-linear) activation functions are the Sigmoid/Logistic function (an S-shaped curve that takes any real value as input and output in the range of 0 to 1), the Tanh (Hyperbolic Tangent) function (takes any real value as input and output in the range of -1 to 1), the Arctan function (takes any real value as input and output in the range of $-(\pi/2)$ to $(\pi/2')$, additionally introducing symmetry), the ReLU (Rectified Linear Unit, or simply: Ramp) function (does not activate all neurons simultaneously; instead, they are only activated if the output of the linear transformation is greater than 0; it has the slope 1 for positive values and a constant function with zero value otherwise), the ELU (Exponential Linear Units) function (is a variant of the ReLU function that modifies the slope of the negative part of the function by using a log curve to define the negative values), the Softmax function (is a combination of multiple sigmoids), the Swish function (is an interpolation between the ReLU function and a scaled version of a linear function), the GELU (Gaussian Error Linear Unit) function (combines properties from dropout, ReLU, etc.), and the SELU (Scaled Exponential Linear Unit) function (captures internal normalization, that is, each layer preserves the mean and variance from the previous layers; therefore, both, positive and negative values shift to the mean).

21 The training from a given example is usually conducted by determining the difference (error) between the processed output (prediction) of the network and a target output. The network then successively adjusts its weighted associations (the weight increases or decreases the strength of the signal) according to a learning rule (loss function) using this error value to produce an output that increasingly corresponds to the target output. The size of the corrective steps to adjust for errors is defined by the learning rate. A high learning rate shortens the training time, but at the expense of a lower ultimate accuracy, while a lower learning rate takes longer, but at the expense of the potential for higher accuracy.

---

19 The word "deep" in deep learning stands for multiple hidden layers.

There is in increasing number of neural networks available. Common (artificial) neural network algorithms are:

- Learning Vector Quantization (LVQ) is a prototype-based supervised (binary and multiclass) classification and regression algorithm. An LVQ system is represented by prototypes (a fixed pool of codebook vectors) which are defined in the feature space that have the same input and output attributes as the training data. While each codebook vector represents a neuron, the collection of all codebook vectors forms the network. For each data point, the codebook vector closest to the input is determined according to a distance metric (BMU – Best Matching Unit).[22] The algorithm starts with a pool of random codebook vectors and the most similar codebook vector is selected. The process is completed for a previously determined number of epochs (iterations on the training dataset).
- Convolutional Neural Networks (CNNs) process structured grid data. They are built on three types of layers: (i) Convolutional layers applying a set of filters (kernels) to the input grid, where each filter extracts significant characteristics from the input data building a feature map that captures essential aspects of the grid, (ii) pooling layers, reducing the dimensionality of the feature maps while retaining the most essential information, and (iii) fully connected layers, flattening the output and feeding it into one or more fully connected (dense) layers, culminating in the output layer that makes the final classification or prediction.
- Recurrent Neural Networks (RNNs) are designed to recognize patterns in data sequences, such as time series (stock price prediction), natural language processing or speech recognition. They maintain a hidden state that memorizes information about previous inputs, and an output generated by the hidden state at each time step. The network is trained using backpropagation through time to minimize the prediction error. RNNs work by using feedback loops to connect the output of each time step back to the input of the next step. This enables the network to utilize prior time step information to inform its predictions for future steps.
- Long Short-Term Memory Networks (LSTMs) are a special kind of RNNs capable of handling sequential input and learning long-term dependencies. They are referred to as "long-short term" because they can recall knowledge from a long time ago while simultaneously disregarding (forgetting) unnecessary information. They include (i) a cell state that runs through the entire sequence and carries information across many learning steps and (ii) three gates that control the flow of information (the input gate determines which information from the current input should be updated, the forget gate decides what information should be discarded, and the output gate controls the information that should be outputted).
- Generative Adversarial Networks (GANs) generate new (realistic) data that resembles the original data by training two neural networks in a competitive setting. In the training process, the generator network (creating new data that is comparable to the original, that is, fake data from random noise) and the discriminator network (evaluating the au-

thenticity of the data, distinguishing between real and fake data) are trained simultaneously. The generator tries to fool the discriminator by producing better fake data, while the discriminator tries to get better at detecting counterfeit data. The two networks are trained simultaneously, with the generator network improving at making plausible fakes and the discriminator network improving at recognizing them. This adversarial process leads to the generator producing increasingly realistic data.[23]
- Transformer Networks (Transformers) process input data using self-attention, allowing for parallelization and improved handling of long-range dependencies. They analyze incoming data and employ attention processes to capture long-range relationships. The Self-Attention Mechanism computes the importance of each input item relative to others (e. g. enabling the model to weigh the significance of different words in a sentence differently). They consist of an encoder and a decoder. Positional Encoding adds information about the position of data items in the sequence since self-attention does not inherently capture sequence order. Following the Encoder-Decoder Architecture, the encoder processes the input sequence and the decoder generates the output sequence. Each consists of multiple layers of self-attention and feed-forward networks.
- Autoencoders (AEs) encode data into a lower-dimensional representation and then decode it back to the original data to compress or denoise it. They are intended to reconstruct the input data, which implies that they learn to encode the information into a compact representation and then decode it back into the original input. The encoder maps the input data to a lower-dimensional latent space representation. This latent space represents the compressed version of the input data. Subsequently, the decoder reconstructs the input data from the latent representation.[24] Variational Autoencoders (VAEs) use variational inference to generate new data points similar to the training data. They encode data into a lower-dimensional space and then decode it back to the original format, learning to spot patterns in the data and producing new data that is similar (comparable) to the original. The encoder maps input data to a probability distribution in the latent space. The latent space sampling samples from the latent space distribution to introduce variability in the generated data. Finally, the decoder generates output data from the sampled latent representation.
- Deep Q-Networks (DQNs) combine deep learning with Q-learning, a reinforcement learning algorithm to handle environments with high-dimensional state spaces. They learn a Q-function that predicts the expected reward for performing a certain action in a particular condition. The Q-function is taught by trial and error, with the algorithm attempting various actions and learning from the outcomes. While Q-Learning[25] uses a Q-table to represent the value of taking an action in a given state, Deep Q-Networks replace

---

22 E. g. the Euclidean distance or the Manhattan distance.

23 GANs may be used for data augmentation, which involves combining produced data with real data to build a bigger dataset for training ML models.
24 Autoencoders are very effective for data compression, noise removal, and anomaly detection. Furthermore, they may be used for feature learning.
25 Q-Learning operates by assessing the value of doing a certain action in a particular state and updating that estimate as the agent interacts with the environment. The agent then utilizes these estimations to determine which action is most likely to result in the largest reward.

the Q-table with a neural network that approximates the Q-values for different actions given a state.

- Self Organizing Maps (SOMs) learn and represent complicated data in a low-dimensional environment. They operate by transforming high-dimensional input data into a two-dimensional grid, with each neuron representing a different area of the input space. The neurons are linked and create a topological structure, allowing them to learn and adjust to the input data. SOMs use the statistical features of the input data to discover patterns and correlations among the variables, and then self-organize into a meaningful structure.
- Capsule Networks (CNs) organize neurons into "capsules" that encode certain aspects of the input data, extracting progressively complicated properties by employing numerous layers of capsules.
- Radial Basis Function Networks (RBFNs) can approximate functions and perform classification tasks. They operate by transforming the input data into a higher-dimensional space employing a collection of radial basis functions. The output created is a linear combination of the basis functions, and each radial basis function represents a center point in the input space. RBFNs are especially effective for situations with complicated input-output interactions.
- Multilayer Perceptrons (MLPs) operate by stacking several layers of neurons, with each layer nonlinearly changing the input data. Each neuron receives input from the neurons in the layer below and sends a signal to the neurons in the layer above. The output of each neuron is determined using an activation function, that gives the network nonlinearity.

5. Clustering Algorithms

Clustering algorithms group a set of data points into clusters based on their similarity. Some clustering algorithms like the k-means, the k-medians, the k-medoids or the expectation-maximization algorithm demand determining the number of clusters to detect (regularly labelled with $\mathbf{k}$), while others like the DBSCAN and the OPTICS algorithm as well as Hierarchical Clustering do not require the specification of this parameter. The correct choice of $\mathbf{k}$ depends on the shape and scale of the distribution of the data points and the desired clustering resolution. The optimal choice of $\mathbf{k}$ strikes a balance between maximum compression of the data using a single cluster, and maximum accuracy by assigning each data point to its own cluster.[26]

6. Instance Based Algorithms

Instance-based (memory-based) algorithms learn – instead of performing explicit generalization – by comparing new problem instances with memorized instances seen in training, and build hypotheses directly from the training instances themselves. They learn and memorize (a subset of) their training data set and subsequently generalize to new instances

computing distances or similarities between the new instance and the training instances by employing some distance metric.[27]

The primary advantage of Instance-based algorithms is that they can easily adopt to new data. Major drawbacks are the large amount of memory required to store the data, and the time complexity depending upon the size of the training data (since in every new instance, all previously stored data is examined).

7. Regularization Algorithms

Regularization algorithms prevent overfitting by adding constraints to the model. It improves the model generalization by penalizing large weights (large coefficients), entailing reduced model complexity (focusing on the primary features), reduced impact of noise in the data, improved stability (building robust models) and interpretability, and enhanced performance on test data. Regularization is particularly crucial for high-dimensional datasets where the risk of overfitting is higher.

Common regularization techniques (each penalizing model complexity differently) are:

- L1 Regularization (Lasso): L1 regularization (Lasso – Least Absolute Shrinkage and Selection Operator) adds a penalty equal to the absolute value of the magnitude of coefficients to the loss function. It encourages sparsity by driving some coefficients to zero, effectively performing feature selection by reducing the number of features in the model and focusing on the most relevant features. Lasso is particularly useful when dealing with high-dimensional data with many irrelevant features, since it reduces the complexity of the model, enhances interpretability, and improves model performance on new data.
- L2 Regularization (Ridge): L2 regularization (Ridge) adds a penalty equal to the square of the magnitude of coefficients to the loss function. It tends to shrink coefficients evenly, but rarely forcing them to zero, ensuring that all features contribute to the model to some extent which leads to more stable and generalizable models.
- Elastic Net Regularization: Elastic Net regularization combines both L1 and L2 regularization penalties. The penalties are balanced employing a mixing parameter that allows to interpolate between L1 and L2 regularization. It is particularly useful when there are multiple correlated features where Lasso might over-penalize and Ridge might under-penalize, providing a balanced approach.
- Weight Regularization: Weight regularization adds a penalty to the loss function to prevent the model from learning excessively large weights. By penalizing large weights, it learns only the most significant features, resulting in more robust models and an improved generalization of the model.
- Activation Function Regularization: Activation function regularization adds a penalty to the activations of the neu-

---

26 If an appropriate value of $\mathbf{k}$ is not apparent from prior knowledge about the properties of the data set, it may be determined employing various techniques, such as the Elbow method, the Silhouette method, the Gap statistic, the Davis-Bouldin Index, the Calinski-Harabasz Index, or the Rand Index.

27 Common distance metrics are the Euclidean distance (measuring a straight line between the memorized and the new instance), the Manhattan distance (measuring the absolute difference between the memorized and the new instance), the Minkowski distance (is the generalized form of the Euclidean and the Manhattan distance metrics), and the Mahalanobis distance (measuring the distance between the mean of the distribution of the memorized and the new instance). Further distance metrics are e. g. the Hamming distance, the Tanimoto distance, the Jaccard distance and, the Cosine distance.

rons in the network. Commonly employed techniques to regularize activations are dropout (randomly deactivating neurons during training, preventing them from becoming too specialized and improving the network's ability to generalize) and batch normalization (regularizing the activations by standardizing them, thus stabilizing the training process).

- Pruning: Pruning is used in decision tree-based models. The process of pruning branches simplifies the decision rules of the tree to prevent it from overfitting the training data.
- Data augmentation: Data augmentation techniques use prior knowledge about the data distribution to prevent model overfitting.

8. Dimensionality Reduction Algorithms

When dealing with high dimensional data, it is often useful to reduce dimensionality by transforming the data from the high-dimensional space into a lower-dimensional subspace that captures all meaningful properties of the original data (close to its intrinsic dimension). Consequently, dimensionality reduction refers to techniques for reducing the number of features in the data prior to modeling. Fewer input dimensions entail a simpler structure (that is, less degrees of freedom) in the ML model, avoiding overfitting the training dataset causing the model not to perform well on new data. Dimensionality reduction algorithms are commonly categorized into two main classes: (i) Linear algebra-based algorithms employing matrix factorization methods include Principal Components Analysis, Singular Value Decomposition, and Non-Negative Matrix Factorization. (ii) Manifold learning-based (nonlinear) algorithms include Isomap Embedding, Locally Linear Embedding, Multidimensional Scaling, Spectral Embedding, and t-distributed Stochastic Neighbor Embedding.

9. Association Rule Learning Algorithms

Association rule learning algorithms discover (relevant) relations between variables in large databases. They attempt to identify strong rules employing some measures of interestingness.

Common association rule learning algorithms are:[28]

- Apriori algorithm: Apriori employs prior knowledge of frequent item-set properties. It proceeds by identifying the frequent individual items in the database and extends them to larger item-sets. Frequent subsets are extended one item at a time (candidate generation) and groups of candidates are tested against the data. The algorithm terminates when no further successful extensions are found.
- Eclat algorithm: The Eclat (Equivalence Class Transformation) algorithm is a backtracking algorithm that traverses the frequent item-set lattice graph in a depth-first search (DFS) fashion. Whereas the breadth-first search (BFS) traversal used in the Apriori algorithm will end up checking every subset of an item-set before checking it, DFS traversal checks larger item-sets and can save on checking the support of some of its subsets by virtue of the downward-closer property.
- FP-growth algorithm: The FP-growth (Frequent Pattern) algorithm first counts the occurrences of items in the dataset and

stores these counts, and, second, builds the FP-tree structure by inserting transactions into the tree. Items in each transaction that do not meet the minimum support requirement are discarded. Growth begins from the bottom (i. e. the item with the smallest support by finding all sorted transactions that end in that item). A new conditional tree is created which is the original FP-tree projected. The supports of all nodes in the projected tree are re-counted with each node getting the sum of its children counts. Nodes (and, hence, subtrees) that do not meet the minimum support are pruned. Recursive growth ends when no individual items meet the minimum support threshold. After this step, processing continues with the next least-supported item of the original FP-tree. Once the recursive process has completed, all frequent item sets are found, and association rule creation begins.

10. Ensemble Algorithms

Ensemble algorithms combine the predictions from multiple models to obtain better predictive performance than could be obtained from any of the constituent learning algorithms on its own. Ensembles combine multiple hypotheses to form a superior hypothesis. They are made up of a set of classifiers aggregating their predictions to identify the best result. Empirically, ensembles tend to yield better results when there is a significant diversity among the individual models. Therefore, they generally seek to promote diversity among the combined models.

Common types of ensemble algorithms are:

- The Bayes Optimal Classifier is an ensemble of all hypotheses in the hypothesis space. The hypothesis represented by the Bayes optimal classifier is the optimal hypothesis. Each hypothesis is given a vote proportional to the likelihood that the training dataset would be sampled from a system assuming that hypothesis to be true. To facilitate training data of finite size, the vote of each hypothesis is multiplied by its prior probability. The Bayes classifier minimizes the probability of misclassification.
- Bootstrap Aggregating (Bagging) trains an ensemble on bootstrapped data sets, achieving inference by voting of predictions of ensemble members (aggregation). Since the features are randomly picked from the original training data set with replacement, a bootstrap set may contain a given feature zero, one, or multiple times. Ensemble members can also have limits on the features (e. g., nodes of a decision tree), to encourage exploring of diverse features. After creating the bootstrapped data sets, decision trees are generated by determining for how many samples the feature's presence or absence yields a positive or negative result. This information is used to compute a confusion matrix that lists the true positives, false positives, true negatives, and false negatives of the feature when used as a classifier. These features are then ranked according to various classification metrics based on their confusion matrices[29] and used to partition the samples into a set that contains the top feature (generally classified as positive) and a set that contains the remaining features (generally classified as negative). This process is repeated recursi-

---

28 Further common association rule learning algorithms are e. g. ASSOC, OPUS search or Lore.

29 Common metrics are the Estimate of Positive Correctness, Entropy, Information Gain, or the Measure of Goodness.

vely for successive levels of the tree until the desired depth is obtained. The trees are then employed as predictors to classify new data. Bagging is designed to improve stability and accuracy, it reduces variance, and it helps to avoid overfitting.

- Boosting trains models successively by emphasizing training data misclassified by previously learned models. Initially, all data has equal weight and is used to learn a base model. The examples misclassified by the base model are assigned a weight greater than correctly classified examples. This boosted data is used to train further base models. Inference is done by voting or averaging. Most boosting algorithms consist of iteratively learning weak classifiers with respect to a distribution, adding them to a final strong(er) classifier.[30]

- Bayesian Model Averaging (BMA) makes predictions by averaging the predictions of models weighted by their posterior probabilities given the data. It generally gives better results than a single model, especially where very different models have nearly identical performance in the training set, but may otherwise perform quite differently.

- Bayesian Model Combination (BMC) is an algorithmic correction to BMA. Instead of sampling each model in the ensemble individually, it samples from the space of possible ensembles (employing model weights drawn randomly from a Dirichlet distribution). This modification overcomes the tendency of BMA to converge toward giving all the weight to a single model.

- A Bucket of Models is an ensemble technique where the best model for every problem is selected by a model selection algorithm. When tested with only one problem, a bucket of models can produce no better results than the best model in the set, but when evaluated across many problems, it will typically produce much better results (on average) than any model in the set.

- Stacking (stacked generalization) trains a model combining the predictions of several other learning algorithms. A combiner algorithm (final estimator) is trained to make a final prediction using all the predictions of the other algorithms (base estimators) as additional inputs or using cross-validated predictions from the base estimators. Stacking typically yields performance better than any single trained models.

VI. Conclusion

Artificial Intelligence will significantly impact virtually every industry, and the business valuation profession will be no exception. It may support valuation analysts (and, thus, may change the valuation profession) by easing repetitive tasks concerning data collection and exploitation prior to any decision making or human judgement, and guiding them when making decisions by providing a better understanding of the key drivers of a valuation opinion. Consequently, AI is a promising source of value creation for valuation analysts, since it reduces cost of routine-tasks (data collection, data analysis, data processing, validation checks, modelling, model review), shifts human expertise to the more complex and, correspondingly, higher value-added tasks, allows for a better management of time, and, allows for a better use and leverage of information.

Part one of this article provides the basics of AI (especially Machine Learning) models and algorithms, model performance optimization, learning paradigms, data inference, algorithm tasks as well as types of variables captured, and describes the basics of different types of common algorithms categorized according to functions and similarities. Built on these basics, the upcoming part two provides an overview on the primary tasks AI (Machine Learning) algorithms have been applied in business valuation research currently emphasizing predictive (rather than generative) algorithms, including their capability to predict company values directly (stand-alone or automated valuation approach), their capability to identify explanatory variables in predicting company values, their performance in selecting peers forming a peer group, and their performance to develop information extraction systems of financial data for business valuation purposes. ◆

---

30 Traditional boosting algorithms are the AdaBoost algorithm (Decision Stumps) for (binary) classification problems.
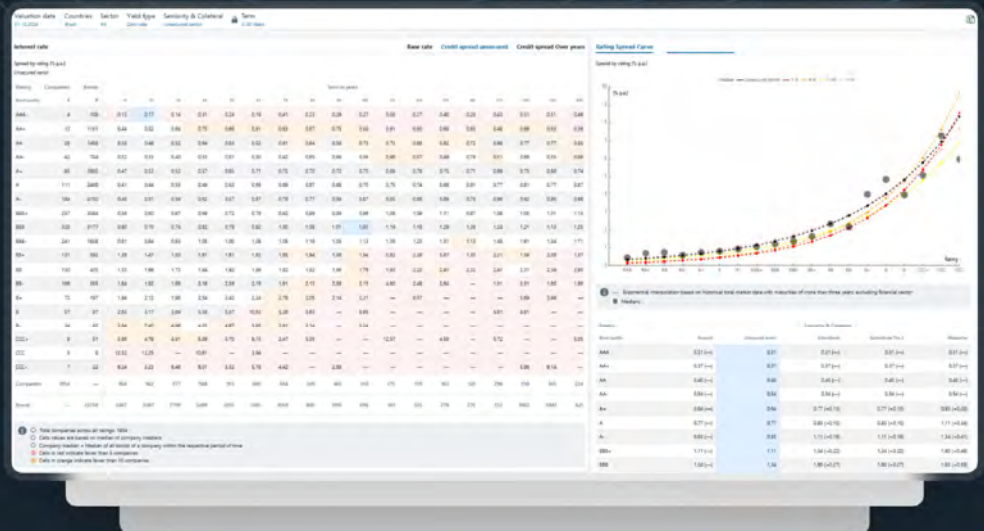
Dr. Christian Reichert, CVA, MBA
Würzburg, Germany.
Contact: ebvm@eacva.de

Valuation of Intangible Assets – An Integrated Relief-from-Royalty Method and Monte Carlo Simulation Approach

This article explores the synergies between Relief-from-Royalty Method (RFR) and Monte Carlo Simulation (MCS), highlighting the advantages of this combined approach in the valuation of intangible assets and illustrating how it can enhance the strategic decision-making process in the face of uncertainty. For illustration purposes, a full calculation example for a technology-related intangible asset in Germany is presented after a detailed explanation of the methodological approach.

Introduction

The precise evaluation of intangible assets, such as patents, trademarks, copyrights, and goodwill, remains a critical yet challenging task in both financial reporting and strategic decision-making. One of the most widely accepted approaches for valuing intangible assets is the Relief-from-Royalty Method (RFR). This method estimates the value of an intangible asset by determining the royalties a company would have to pay if it did not own the asset and had to license it from a third party. While intuitive, the RFR method involves several assumptions, including the determination of royalty rates (RR), the expected future revenues associated with the asset, and the appropriate discount rate (DR). Due to the lack of e.g. robust market and financial data for the intangible asset, significant (epistemic) uncertainties are associated with the valuation process. In that context, epistemic uncertainties are understood as uncertainties that arise due to imperfect information or limited data about the system under review and which can be reduced by more data or by model refinements.[1]

In order to address this uncertainty and improve the valuation model, Monte Carlo Simulation (MCS) provides a powerful tool. MCS is a statistical technique that uses random sampling and probability distributions to model and evaluate the impact of uncertainty on the asset's value. By simulating a range of possible scenarios and incorporating variability in key inputs (e.g. RR, Compound Annual Growth Rate (CAGR), DR), MCS can generate a probability distribution for the intangible asset's fair value (FV). Consequently, offering a more robust and comprehensive understanding of the asset's potential worth.

Methodology

The proposed methodology combines RFR and MCS as well as an expert-based survey with the aim of assessing key inputs of the RFR. The common and deterministic approach of the RFR method comprises in general the following steps:[2] 1) estimation of the revenue stream, 2) determination of the RR, 3) definition of the DR and eventually 4) calculation of the FV. Initially, step 1) tries to specify the future revenue directly attributable to the intangible asset under review. Regarding step 2) market comparable RR are used very often to derive a hypothetical RR which is multiplied with the annual revenue to determine the royalty payments. In the past, the "25% Rule" has been applied widely as an alternative starting point to calculate the RR as 25% of the expected profits of the IP-related product.[3] Since the royalty payments are planned to occur over the whole useful asset lifetime, e.g. patent term, discounting with an appropriate DR is required in step 3) to gain the current value of the royalty payments. As an estimate of the DR specific to the asset the industry average Weighted Average Cost of Capital (WACC)

is normally used in IP valuation projects, and if necessary, the WACC will be adjusted to asset-specific risks by means of an additional risk premium or an adapted debt/equity ratio.[4] Finally, the FV of the intangible asset is calculated in step 4) by adding up the discounted annual royalty payments over the entire useful asset lifetime. Whereas the FV represents the economic benefit derived from owning the asset instead of paying royalties for the use of the intangible asset to another party. In the context of the process described above, there are three main (exogenous) factors that have theoretically an influence on the FV but that are subject to (epistemic) uncertainty at the time of the RFR valuation. These are namely the asset-specific future revenue in terms of CAGR, the market comparable RR and the appropriate DR.

The proposed combined (probabilistic) RFR and MCS approach tries to address the three uncertain model factors mentioned above quantitatively by means of three main methodological steps:

1. Expert-based survey of uncertain factors

The CAGR, RR and DR in particular are characterized by a high degree of uncertainty, as they often relate to (registered) intellectual property (IP) for which, by its very nature, no reliable historical market or financial data as benchmarks are available. An initial expert-based survey, e.g. by means of the Delphi method, aims at assessing the uncertain RFR factors throughout the remaining useful lifetime (RUL): 1) the specific CAGR of the future revenues, 2) the RR as well as 3) the appropriate DR related to the intangible asset. However, to derive the required parameters of the hypothetical distribution based on the expert assessment, a direct two-step approach is applied, i.e. each expert provides in a first step information on every required distribution parameter of each uncertain model factor. In a second step, the experts' estimates will be statistically analyzed by identifying the maximum of the surveyed maximum values, the minimum of the surveyed minimum values and the mode of the surveyed modal values. By relying on expert judgment and experience combined with MCS, the approach may help to overcome (commercial and financial) data limitations.

2. Simulation of the CAGR, RR and DR

Regarding the three assumed uncertainty factors **CAGR**, **RR** and **DR** introduced into the model by an MCS, a triangular distribution was selected. The triangular distribution is often referred to as "lack of knowledge" distribution and recommended as most suitable in case of applications in which only limited information is available, but at least upper and lower bounds as well as the most likely value of the distribution can be estimated.[5] Given the limited availability of quantitative data but being able to conduct an expert-based survey on the worst, best as well as most likely case of the future revenue, possible royalties and most appropriate DR of a highly innovative technology, the triangular distribution was applied as the most fitting probability distribution.

---

1   For a more comprehensive introduction to epistemic uncertainty, please see, Der Kiureghian/Ditlevsen, Aleatory or epistemic? Does it matter?, Structural Safety, vol. 31, no. 2 (2009): 105-112.

2   For a detailed description of the methodology, please see, Reilly, Relief from Royalty Method of Intellectual Property Valuations, les Nouvelles – Journal of the Licensing Executives Society, vol. LVII, no. 1 (2022): 15-30.

3   For a further discussion of the "25% Rule", please see, Goldscheider, The Classic 25% Rule and the Art of Intellectual Property Licensing, Duke Law & Technology Review, vol. 10, no. 1 (2011): 1-22.

4   Grant Thornton International Ltd., Intangible assets in a business combination – Identifying and valuing intangibles under IFRS 3, (2013): 26-29.

5   Wang/Pinsky, Geometry of deviation measures for triangular distributions, Frontiers in Applied Mathematics and Statistics, vol. 9 (2023): 1-14.

Inversion of the cumulative distribution function (CDF) represents the standard approach to generate random variates from the triangular distribution with minimum a, mode c, and maximum b.[6] In detail, if U is a uniform variate drawn from the interval $(0, 1)$, the required (triangularly distributed) random variates are produced by the inverse CDF equations (1), (2) and (3).[6] Whereas $F(c) = (c-a)/(b-a)$ follows a triangular distribution with parameters a, b and c. The $CAGR_{MCS}$, RR_{MCS} and DR_{MCS} are integrated into the MCS model with a total of 10,000 runs.

$$CAGR_{MCS} = \begin{cases} a_{CAGR} + \sqrt{U\left(b_{CAGR} - a_{CAGR}\right)\left(c_{CAGR} - a_{CAGR}\right)} \\ b_{CAGR} - \sqrt{(1-U)\left(b_{CAGR} - a_{CAGR}\right)\left(b_{CAGR} - c_{CAGR}\right)} \end{cases} \tag{1}$$
$$\text{if } 0 < U < F(c)$$
$$\text{if } F(c) \leq U < 1$$

$$RR_{MCS} = \begin{cases} a_{RR} + \sqrt{U\left(b_{RR} - a_{RR}\right)\left(c_{RR} - a_{RR}\right)} \\ b_{RR} - \sqrt{(1-U)\left(b_{RR} - a_{RR}\right)\left(b_{RR} - c_{RR}\right)} \end{cases} \tag{2}$$
$$\text{if } 0 < U < F(c)$$
$$\text{if } F(c) \leq U < 1$$

$$DR_{MCS} = \begin{cases} a_{DR} + \sqrt{U\left(b_{DR} - a_{DR}\right)\left(c_{DR} - a_{DR}\right)} \\ b_{DR} - \sqrt{(1-U)\left(b_{DR} - a_{DR}\right)\left(b_{DR} - c_{DR}\right)} \end{cases} \tag{3}$$
$$\text{if } 0 < U < F(c)$$
$$\text{if } F(c) \leq U < 1$$

3. Determination of the FV

Initially, the achievable first period revenue REV_1 for the intangible asset is estimated based on, e.g. market analysis. The starting revenue serves as the basis for the forecast of annual revenue until the end of the RUL. The level of the CAGR within the RUL is then modeled using the MCS (see (1)). The corresponding distribution parameters of the CAGR simulation are determined by the expert survey mentioned above. The specific revenue (REV) in each year t of the RUL is calculated in the following way:

$$REV_t = REV_{t-1} \cdot \left(1 + CAGR_{MCS}\right) \tag{4}$$

Next, the RR is modeled as a triangular-distributed random variable (see (2)), with parameters derived from the expert survey. This results in the specific RR in year t. By applying the MCS-based RR as well as considering the domestic corporate income tax (CIT), the post-tax Royalty Savings (RS) in a given year t can be derived:

$$RS_t = REV_t \cdot RR_{MCS} \cdot \left(1 - CIT\right) \tag{5}$$

As already explained, also the DR is incorporated into the model as survey-based MCS parameter with the purpose of calculating the present value of the RS in a given year t (see (3)). Ultimately, the total of the annual present values of the RS ($PVRS_{total}$) of the intangible asset throughout the entire RUL is determined as follows:

$$PVRS_{total} = \sum_{i=1}^{n} RS_t \cdot \frac{1}{\left(1 + DR_{MCS}\right)^t} \tag{6}$$

The final calculation step includes the determination of n FV of the intangible asset (by means of the MCS) related to the RUL including a tax amortization benefit factor (TABF) by equation (8). Specifically, the tax amortization benefit is the present value of tax savings resulting from the deduction of an intangible asset when it is amortized and is applied in income-based valuation approaches such as discounted cash flow (DCF) or RFR. The respective TABF is calculated by equation (7).

$$TABF = \frac{1}{\left[1 - \frac{CIT}{RUL} \cdot \left(\frac{1}{DR_{MCS}} - \frac{1}{\left(DR_{MCS} \cdot \left(1 + DR_{MCS}\right)^{RUL}\right)}\right)\right]} \tag{7}$$

$$FV = PVRS_{total} \cdot TABF \tag{8}$$

III. Example: Technology-related intangible asset in Germany

The following section will explain in detail the case study including key assumptions. The example refers to a technology-related intangible asset of a company located in Germany. The intangible asset, i.e. registered IP, represents the Unique Selling Proposition (USP) and thus the technological core of the company's business model. Regarding the example case, the patent was recently granted, and therefore commercialization did not take place yet. Furthermore, the following analyses, and evaluations are conducted at time t_0 without substantial quantitative information or data at hand. German patent law prescribes a patent term of 20 years.[7] Consequently, the RUL is set at 20 years, which means that IP protection for the underlying technology according to German patent law still exist for another 20 years. Currently, in Germany the CIT including a solidarity surcharge account for 15.825% added by a trade tax ranges from 8.75% to 20.30%, depending on the location of the business.[8] Therefore, an average CIT of 30.35% is assumed.

The initial revenue REV_1 at time t_1 dedicated directly to the technology-related intangible asset is set at 500 kEUR based on a preliminary market analysis.

6 Stein/Keblis, A new method to simulate the triangular distribution, Mathematical and Computer Modelling, vol. 49, no. 5-6 (2009): 1143-1147.

7 Patent Act (Patentgesetz – PatG), Link ».
8 PwC, Worldwide Tax Summaries – Germany (Last reviewed 30 June 2025), Link ».

Concerning the survey-based estimation of the three uncertain model factors, a panel of 10 experts[9] was asked to provide an estimate of the three triangular distribution parameters a, b and c for each of the CAGR, RR and DR. This results in 30 responses per uncertainty factor, 10 responses per parameter and 90 responses in the overall survey. The aggregation is realized in a second step by identifying for each uncertain model factor parameter a based on the surveyed minimum values, parameter b based on the surveyed maximum values and parameter c based on the surveyed modal values at the time t_0. Based on the judgment of the internal panel of 10 experts, the necessary distribution parameters of the triangular distribution were derived according to the direct two-step approach at time t_0 (see Table 1).

Table 1: Expert-based assessment of triangular distribution parameters of uncertainty factors CAGR, RR and DR at time t_0

Parameter	CAGR	RR	DR
Minimum (a)	2.0%	8.0%	8.0%
Maximum (b)	14.0%	25.0%	25.0%
Mode (c)	6.5%	15.0%	17.0%

Looking at the results of the simulation model, various observations can be made regarding the statistical properties of the FV. For instance, the mean value of the FV is 554.12 kEUR and the median is 519.67 kEUR. The simulated minimum value is 214.88 kEUR and the maximum value 1,990.57 kEUR.

When analysing the distribution shown in Table 2 and Figure 1 (see page 20), it can be determined that an FV value between greater than 750 and lower or equal to the maximum FV of approx. 1,991 kEUR occurs with a probability of 7.11%. An FV value below or equal 650 kEUR has a probability of 75.70%. Whereas a value lower or equal than 450 kEUR can occur with a probability of 32.15%.

Table 2: Relevant FV categories and related probabilities

#	FV [kEUR]			
	>	<=	p	p cumulated
1	0	250	0.43%	0.43%
2	250	350	8.98%	9.41%
3	350	450	22.74%	32.15%
4	450	550	25.25%	57.40%
5	550	650	18.30%	75.70%
6	650	750	11.17%	86.87%
7	750	850	6.02%	92.89%
8	850	950	3.14%	96.03%
9	950	1050	1.72%	97.75%
10	1050	1150	1.18%	98.93%
11	1150	1250	0.47%	99.40%
12	1250	1350	0.29%	99.69%
13	1350	1450	0.17%	99.86%
14	1450	1550	0.06%	99.92%
15	1550	1650	0.03%	99.95%
16	1650	1991	0.05%	100.00%

In addition to the above-mentioned statistical analysis of the simulation model results, the data generated can also be used to perform e.g. a Risk-of-Loss (RoL) analysis to provide adequate support for strategic decision-making. For instance, the decision-maker can determine the RoL by means of the simulation model for a defined investment amount that is assumed to be necessary for the full commercialization of the technology-related intangible asset. In other words, the probability that the FV will fall short of a certain given investment amount can be quantified. As an example, Table 3 shows a possible RoL distribution depending on possible investment amounts for commercialization. According to the analysis carried out, for example, the RoL for an investment of 600 kEUR is 67.4%. In contrast, the risk for an investment of 300 kEUR is only 2.8%.

Table 3: Results of the RoL analysis for commercialization investments

Investment [kEUR]	Risk-of-Loss (RoL)
600	67.4%
500	45.4%
400	19.6%
300	2.8%
200	0,0%

IV. Conclusions

The combined approach presented above utilizes both the RFR and MCS based on an expert survey and enables a more dynamic evaluation of intangible assets compared to deterministic valuations. This approach not only benefits from the traditional RFR methodology but also incorporates the flexibility of MCS to capture the inherent epistemic uncertainties and resulting risks in the valuation process. The integration of these aforementioned techniques (expert survey, three-point estimation, MCS) can lead to more accurate and reliable valuations compared to simple and deterministic rates as well as single-point estimations and may empower businesses, investors, and accountants to make

---

9   The expert panel can be surveyed in a systematic and structured way, for example, by the application of the Delphi technique.

Figure 1: FV Histogram and S-Curve
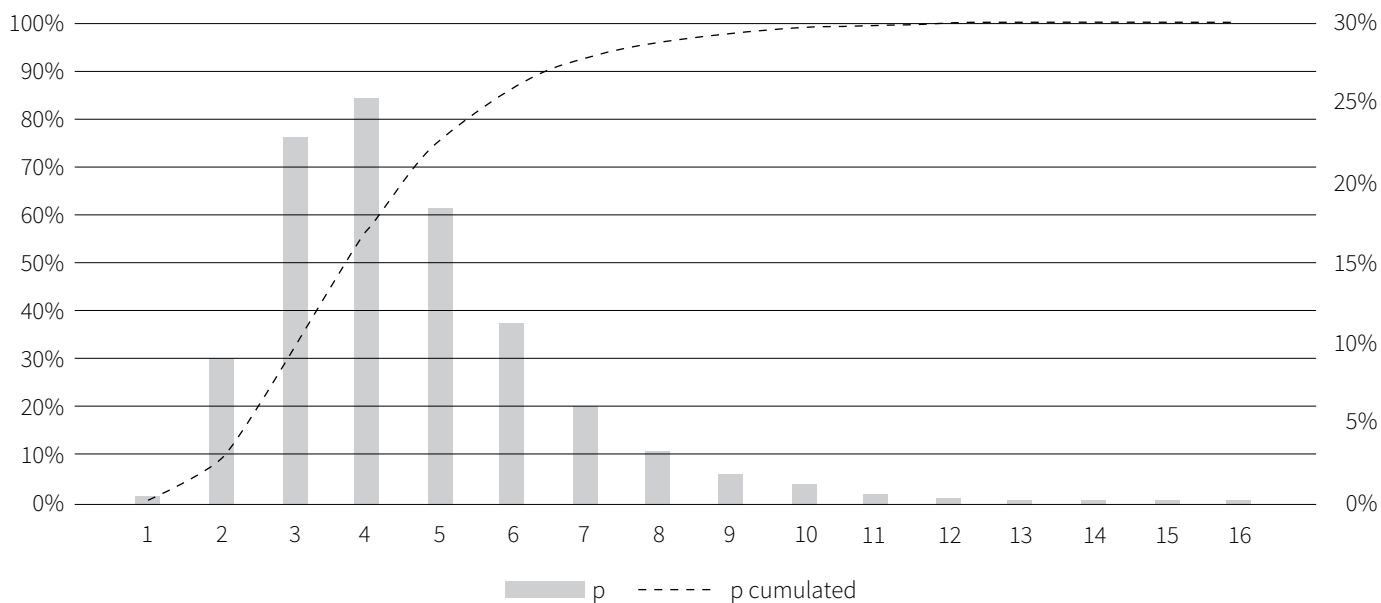
more informed decisions regarding the FV of intangible assets. In that context, the model developed in this study has shown that a combination of RFR and MCS can provide a pragmatic analytical approach in absence of robust data for the valuation of intangible assets. As part of this approach, significant (epistemic) uncertainties can be integrated into the FV valuation model in a structured manner. In particular, the involvement of expert judgments for the purpose of quantifying uncertain model factors may compensate for the lack of reliable data. Regarding the quality of the decision support, the combined approach delivers reliable quantitative results about the specific FV that should be taken into account for strategic decisions related to the commercialization of the intangible asset. ◆

Certified Valuation Analyst (CVA)

– the most widely recognised business valuation credential worldwide –

CVA Courses 2025/2026:

» **20 – 24 October 2025 in Luxembourg** – *in English*

» **24 – 29 November 2025 in Munich** – *in German (fully booked)*

» **21 – 28 January 2026 live online** – *in German*

» **23 – 28 March 2025 in Frankfurt** – *in German*

» **13 – 17 April 2026 in Luxembourg** – *in English*

» **4 – 9 May 2026 in Vienna** – *in German*

» **24 – 29 August 2026 in Berlin** – *in German*

» **30 November – 4 December 2026 in Munich** – *in German*

Learn more and register:

» courses in English language: www.EACVA.com/certified-valuation-analyst-cva

» courses in German language: www.EACVA.de/certified-valuation-analyst-cva

For further information, please contact the EACVA team
e-mail: info@eacva.de / phone: +49 6108 97 444 20

Industry Betas and Multiples

Dr. Martin H. Schmidt
Senior Manager Deal Advisory KPMG AG WPG Germany
Contact: ebvm@eacva.de

Dr. Andreas Tschöpel, CVA, CEFA, CIIA
Partner Deal Advisory KPMG AG WPG Germany, Member of the Technical Committee for Business Valuation and Economics (FAUB) of the IDW e.V., Board Member of the EACVA e.V.

General

To derive the provided betas and multiples, only companies from the Eurozone have been considered. The included companies have been grouped on an industry level and on a sub-industry level based on the Global Industry Classification Standard (GICS). In each issue of the journal, aggregates for all eleven main industries and one individually selected sub-industry will be shown. Due to the special characteristics of companies operating in the financial industry (high leverage, leverage as part of the operating business, high dependency on the interest rate level, etc.), we only provide levered betas and equity-based multiples for that industry.

All presented values are based on raw data and raw calculations. They have carefully been checked and evaluated but have not been audited nor have individual values been verified. Certain results may be misleading in your setup or specific context. All results should be critically evaluated and interpreted. The data and usage are at your own risk.

Data source

All data has been obtained from the KPMG Valuation Data Source. The data source provides access to cost of capital parameters from more than 150 countries and sectors as well as peer-group-specific data from over 16,500 companies worldwide. The data covers the period from 2012 to the present. The data is updated monthly and is accessible from anywhere around the clock.
See www.kpmg.de/en/valuation-data-source for details.

Eurozone Cost of Capital Parameters as at 31 July 2025

The typified, uniform risk-free rate based on AAA-rated government bonds currently lies at 3% for the Eurozone. It is derived from yield curves based on Svensson parameters and results published by the European Central Bank. The overall long-term market return for the Eurozone is estimated at around 8.5%, leading to a market risk premium of 5.5%. Estimations of the market return rely on historical returns, as well as on forward-looking return estimates and risk premiums based on Eurozone companies with current market share prices and earnings forecasts from financial analysts.

Betas

Levered, debt and unlevered betas are calculated over an observation period of a single five-year period (monthly returns) and for five one-year periods (weekly returns).

Raw levered betas are obtained from a standard OLS regression, with stock returns being the dependent variable and stock market index returns (S&P Eurozone BMI Index) being the independent variable. Stock and index returns are total returns, thus including dividends, stock splits, rights issues, etc. (if available). Levered betas below zero and above three are treated as outliers and are excluded.

Unlevered betas have been estimated based on Harris-Pringle, assuming uncertain tax shields and including debt beta:

$$\beta_u = \beta_L \frac{E}{E + D} + \beta_D \frac{D}{E + D},$$

where β_U = unlevered beta, β_D = debt beta, D = net debt, E = market value of equity. Debt betas rely on a company's individual rating on a given date. Monthly rating-specific levels of debt betas are extracted from a broad market analysis. Net debt consists of total debt (incl. lease liabilities) + net pensions + total preferred equity - total cash - short-term investments. In accordance with the observation period, parameter averages of debt beta, net debt and market equity over the individual periods are applied when unlevering levered betas. Unlevered betas below zero and above two are treated as outliers and are excluded.

Table 1: Median Levered Industry Betas for five single 1y-periods and one 5y-period

31 July 2025	Median Levered Betas								
	1-Year, weekly returns							5-Year, monthly returns	
Industries	Comps incl. (Average*)	8/2020 to 7/2021	8/2021 to 7/2022	8/2022 to 7/2023	8/2023 to 7/2024	8/2024 to 7/2025	Average*	Comps incl.	8/2021 to 7/2025
Industrials	268	0.92	0.82	0.83	0.84	0.89	0.86	246	1.01
Consumer Discretionary	173	0.98	1.01	0.94	0.92	0.94	0.96	156	1.07
Health Care	133	0.71	0.68	0.74	0.79	0.78	0.74	124	0.72
Financials	143	1.15	0.89	0.89	0.73	0.91	0.91	137	1.05
Utilities	50	0.75	0.56	0.62	0.75	0.35	0.61	47	0.68
Materials	87	0.88	0.83	0.93	0.82	0.96	0.88	86	0.99
Real Estate	80	0.59	0.57	0.84	0.79	0.41	0.64	76	0.80
Communication Services	86	0.70	0.55	0.68	0.58	0.62	0.62	84	0.80
Information Technology	155	0.77	0.94	0.91	0.82	0.80	0.85	141	1.03
Consumer Staples	77	0.49	0.69	0.48	0.43	0.40	0.50	75	0.54
Energy	35	1.22	0.46	0.68	0.41	0.86	0.73	35	0.79

Table 2: Median Industry Equity-Ratios for five single 1y-periods and one 5y-period

31 July 2025	Median Equity-Ratios								
	1-Year							5-Year	
Industries	Comps incl. (Average*)	8/2020 to 7/2021	8/2021 to 7/2022	8/2022 to 7/2023	8/2023 to 7/2024	8/2024 to 7/2025	Average*	Comps incl.	8/2021 to 7/2025
Industrials	274	83.5%	81.6%	80.6%	78.3%	80.9%	81.0%	241	79.9%
Consumer Discretionary	174	85.4%	78.5%	75.2%	73.5%	68.3%	76.2%	147	75.8%
Health Care	138	99.5%	97.1%	91.4%	93.6%	94.9%	95.3%	123	97.6%
Utilities	50	64.4%	67.0%	61.5%	59.7%	59.4%	62.4%	47	63.3%
Materials	90	79.7%	78.5%	77.7%	77.6%	74.3%	77.6%	87	75.7%
Real Estate	82	58.5%	51.5%	46.6%	48.7%	49.7%	51.0%	72	49.8%
Communication Services	90	83.3%	81.1%	75.1%	74.1%	77.3%	78.2%	80	77.5%
Information Technology	155	98.4%	95.2%	93.3%	93.2%	92.1%	94.4%	138	95.7%
Consumer Staples	80	78.4%	74.3%	71.3%	68.9%	70.5%	72.7%	74	72.9%
Energy	37	68.3%	72.0%	83.9%	86.1%	79.9%	78.0%	34	70.8%

Table 3: Median Unlevered Industry Betas for five single 1y-periods and one 5y-period

31 July 2025	Median Unlevered Betas								
	1-Year, weekly returns							5-Year, monthly returns	
Industries	Comps incl. (Average*)	8/2020 to 7/2021	8/2021 to 7/2022	8/2022 to 7/2023	8/2023 to 7/2024	8/2024 to 7/2025	Average*	Comps incl.	8/2021 to 7/2025
Industrials	254	0.76	0.67	0.67	0.65	0.72	0.69	228	0.80
Consumer Discretionary	155	0.74	0.78	0.71	0.69	0.70	0.72	137	0.86
Health Care	116	0.53	0.52	0.59	0.54	0.56	0.55	108	0.58
Utilities	49	0.60	0.41	0.46	0.53	0.30	0.46	45	0.47
Materials	85	0.78	0.69	0.72	0.65	0.78	0.72	82	0.76
Real Estate	75	0.44	0.46	0.49	0.55	0.32	0.45	66	0.59
Communication Services	80	0.58	0.52	0.55	0.45	0.52	0.52	77	0.64
Information Technology	144	0.72	0.84	0.75	0.69	0.69	0.74	129	0.92
Consumer Staples	75	0.48	0.54	0.44	0.37	0.39	0.44	71	0.49
Energy	33	1.00	0.39	0.58	0.37	0.71	0.61	34	0.71

Source: KPMG Valuation Data Source, see www.kpmg.de/en/valuation-data-source
Average = Arithmetic Mean

Data

Table 4: Median Levered Subindustry (Health Care) Betas for five single 1y-periods and one 5y-period

31 July 2025	Median Levered Betas								
	1-Year, weekly returns							5-Year, monthly returns	
Subindustry: Health Care	Comps incl. (Average*)	8/2020 to 7/2021	8/2021 to 7/2022	8/2022 to 7/2023	8/2023 to 7/2024	8/2024 to 7/2025	Average*	Comps incl.	8/2021 to 7/2025
Health Care Equipment & Supplies	24	0.78	0.74	0.82	0.95	0.88	0.83	26	0.69
Health Care Providers & Services	22	0.41	0.59	0.60	0.58	0.49	0.53	23	0.55
Health Care Technology	7	0.78	0.60	0.85	0.33	0.78	0.67	7	0.74
Biotechnology	44	0.96	0.87	0.70	0.98	1.01	0.90	38	0.84
Pharmaceuticals	44	0.96	0.87	0.70	0.98	1.01	0.66	23	0.60
Life Sciences Tools & Services	8	0.49	0.76	0.98	1.34	0.72	0.86	7	0.88

Table 5: Median Subindustry (Health Care) Equity-Ratios for five single 1y-periods and one 5y-period

31 July 2025	Median Equity-Ratios								
	1-Year							5-Year	
Subindustry: Health Care	Comps incl. (Average*)	8/2020 to 7/2021	8/2021 to 7/2022	8/2022 to 7/2023	8/2023 to 7/2024	8/2024 to 7/2025	Average*	Comps incl.	8/2021 to 7/2025
Health Care Equipment & Supplies	25	97.8%	89.2%	86.7%	84.9%	88.1%	0.89	24	91.2%
Health Care Providers & Services	24	80.3%	69.9%	55.4%	64.6%	67.0%	0.67	23	69.1%
Health Care Technology	8	99.0%	100.7%	93.5%	88.8%	97.8%	0.96	8	97.4%
Biotechnology	43	116.7%	115.8%	107.6%	108.0%	108.1%	1.11	38	113.3%
Pharmaceuticals	30	96.2%	88.6%	93.1%	93.5%	96.7%	0.94	23	93.2%
Life Sciences Tools & Services	9	93.6%	93.8%	92.2%	89.1%	88.2%	0.91	7	89.7%

Table 6: Median Unlevered Subindustry (Health Care) Betas for five single 1y-periods and one 5y-period

31 July 2025	Median Unlevered Betas								
	1-Year, weekly returns							5-Year, monthly returns	
Subindustry: Health Care	Comps incl. (Average*)	8/2020 to 7/2021	8/2021 to 7/2022	8/2022 to 7/2023	8/2023 to 7/2024	8/2024 to 7/2025	Average*	Comps incl.	8/2021 to 7/2025
Health Care Equipment & Supplies	23	0.58	0.75	0.87	0.71	0.60	0.70	23	0.61
Health Care Providers & Services	22	0.43	0.44	0.45	0.46	0.46	0.45	22	0.50
Health Care Technology	7	0.59	0.62	0.81	0.29	0.55	0.57	6	0.69
Biotechnology	31	0.68	0.37	0.57	0.59	0.62	0.57	28	0.48
Pharmaceuticals	25	0.62	0.41	0.46	0.56	0.57	0.52	22	0.55
Life Sciences Tools & Services	8	0.48	0.68	0.89	1.21	0.64	0.78	7	0.82

Source: KPMG Valuation Data Source, see www.kpmg.de/en/valuation-data-source
*Average = Arithmetic Mean

Multiples

Multiples are computed based on actuals (based on the annual report) and forecasts (based on consensus estimates by analyst) for the trailing year and the forward +1 year. Trading multiples for Sales, EBITDA and EBIT are each derived by dividing a companies' enterprise value (market value of equity plus net debt) by its sales, EBITDA or EBIT. Earnings multiples are derived by dividing a companies' market value of equity by earnings (net income). The market-to-book ratio is derived by dividing a companies' market value of equity by its book value of equity. Multiples below zero and above 500 are treated as outliers and are excluded. ◆

Table 7: Median Industry Multiples

31 July 2025 Industries	Sales			EBITDA			EBIT			Earnings			Market to Book-Ratio		
	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.
Industrials	1.0	1.0	210	7.6	6.6	203	12.5	10.9	199	15.4	13.3	190	1.9	1.7	187
Consumer Discretionary	0.8	0.8	130	6.7	6.1	126	12.3	10.5	122	13.5	10.7	117	1.7	1.5	114
Health Care	2.5	2.4	100	9.5	9.2	75	14.5	12.9	74	18.0	15.1	69	2.0	2.0	83
Financials	n/m	n/m	n/a	n/m	n/m	n/a	n/m	n/m	n/a	10.8	10.2	108	1.3	1.2	99
Utilities	2.9	2.9	38	8.6	8.2	38	15.0	14.3	38	14.0	14.6	37	1.6	1.5	35
Materials	1.0	0.9	75	6.8	6.0	74	11.7	10.0	68	12.9	11.3	65	1.1	1.1	68
Real Estate	11.7	11.4	45	18.9	17.1	47	19.1	16.7	48	13.8	12.6	35	0.8	0.8	43
Communication Services	1.6	1.5	61	5.7	5.8	61	12.2	10.5	54	13.1	11.5	51	1.6	1.5	53
Information Technology	1.3	1.2	110	9.1	7.6	105	14.4	12.1	98	19.0	15.5	87	2.4	2.2	90
Consumer Staples	0.7	0.6	59	7.2	6.9	60	12.6	11.4	57	14.3	13.9	55	1.4	1.4	55
Energy	1.1	1.1	26	5.3	5.0	26	7.7	7.7	27	11.9	9.6	27	1.1	1.1	26

Table 8: Median Subindustry (Health Care) Multiples

31 July 2025 Subindustry: Health Care	Sales			EBITDA			EBIT			Earnings			Market to Book		
	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.	Trai-ling	Fwd. +1	Comps incl.
Health Care Equipment & Supplies	3.2	2.5	18	12.6	11.2	16	15.7	14.4	16	19.2	16.4	16	2.3	2.2	15
Health Care Providers & Services	1.0	1.1	18	7.3	7.3	17	11.9	12.1	17	13.3	11.5	13	1.3	1.6	16
Health Care Technology	2.3	1.9	4	9.3	7.9	3	15.8	13.9	4	17.8	12.1	5	1.9	1.9	5
Biotechnology	4.0	4.0	28	13.3	11.5	10	14.7	12.8	10	21.4	15.5	9	1.6	1.8	16
Pharmaceuticals	2.7	2.6	23	10.0	9.0	20	13.7	11.6	19	15.8	13.0	18	2.5	2.4	22
Life Sciences Tools & Services	3.8	3.5	9	14.1	12.7	9	18.5	16.4	8	22.0	19.5	8	2.9	2.6	9

Source: KPMG Valuation Data Source, see www.kpmg.de/en/valuation-data-source
*Average = Arithmetic Mean

Data

Discounts for Lack of Marketability

Professor Dr. Stefan O. Grbenic, StB, CVA

Professor of Management Control, Accounting and Finance at Webster University St. Louis/Vienna and Graz University of Technology and Visiting Professor at University of Maribor, Istanbul Medeniyet University and University of Twente.

Contact: ebvm@eacva.de

Over the years, a variety of Option Pricing Models (hereinafter OPM) have been introduced to estimate Discounts for Lack of Marketability (hereinafter DLOM), capturing the key value drivers stock price volatility, period of illiquidity, and dividend yield.[1] The DLOM are computed employing three OPM generally proved to generate DLOM estimates that comport with DLOM empirically observed on the European market[2] according to varying assumptions about the period of illiquidity, the size of the underlying DLOM benchmarks, the volatility of the underlying stock return and, the dividend yield (employing closed-form solution formulae):[3]

- Lookback Put OPM:[4]

$$\text{DLOM}_i = \frac{1}{P_i} P_i \left[\theta_i \right]$$

with $\theta_i = \left(2 + \frac{\sigma_i^2 T}{2} \right) N\left(\frac{\sqrt{\sigma_i^2 T}}{2} \right) + \sqrt{\frac{\sigma_i^2 T}{2\pi}} e^{-\frac{\sigma_i^2 T}{8}} - 1$

- Adjusted Lookback Put OPM:[5]

$$\text{DLOM}_i = \frac{P_i \left[\theta_i \right]}{1 + P_i \left[\theta_i \right]}$$

- Perpetual Exchange Put OPM:[6]

$$\text{DLOM}_i = \frac{1}{P_i} \left(\frac{P_i}{-\psi_i - \frac{1}{2}} \right) \left(\frac{-\psi_i - \frac{1}{2}}{\frac{1}{2} - \psi_i} \right)^{\left(\frac{1}{2} - \psi_i \right)}$$

with $\psi_i = \sqrt{\frac{1}{4} + \frac{2q_i}{\sigma_i^2}}$

where **i** is the index on the stocks related to DLOM estimates, P_i is the current price of the underlying stock as on end of computation period date, σ_i is the volatility of the underlying stock return, **T** is the period of illiquidity (holding period) indicating the period the stock is expected to remain non-marketable, q_i is the dividend yield of the underlying stock and, **N()** is the cumulative normal distribution function.

The computations are based on stock and company data directly collected from the stock exchanges as well as from yahoo!finance.

When using the data, please consider the following:

- DLOM are computed employing (stock and company) data for the year 2024.
- DLOM reported in the tables for all three OPM are computed employing the arithmetic mean of all values available.
- The tables for all three OPM are separated for various periods of illiquidity (holding periods) 3 months, 6 months, 9 months, 1 year, 1.5 years and 2 years with the choice on the holding period depending on the specific valuation. The final table for the Perpetual Exchange Put OPM holds irrespective of choosing a specific holding period.
- Countries with less than 20 observations (10 observations for the Perpetual Exchange Put OPM) remain unreported, but are included in the regional breakdown.
- The various regions (see bottom of the tables) are compounded as follows:
  - **Central and Western Europe:** Andorra, Austria, Belgium, France, Germany, Liechtenstein, Luxembourg, Monaco, The Netherlands, Switzerland;
  - **Southern Europe:** Croatia, Cyprus, Gibraltar, Greece, Italy, Malta, Portugal, San Marino, Slovenia, Spain, Turkey;
  - **Scandinavia:** Denmark, Finland, Iceland, Norway, Sweden
  - **Britain:** Ireland, United Kingdom;
  - **Eastern Europe:** Albania, Armenia, Azerbaijan, Belarus, Bosnia and Herzegovina, Bulgaria, Czech Republic, Estonia, Georgia, Hungary, Kazakhstan, Kosovo, Latvia, Lithuania, Moldova, Montenegro, North Makedonia, Poland, Romania, Russia, Serbia, Slovakia, Ukraine.
- The volatility σ_i of the underlying stock return is computed by the standard deviation of daily logarithmic stock returns (adjusted close prices) over the year 2024. To avoid distortions by thin trading, stocks with too many observations

---

1 For a theoretical analysis see e. g. Hitchner/Aldering/Angell/Morris, Discount for Lack of Marketability, 2011: 305-351.
2 See Grbenic/Baumüller, Zum Fungibilitätsabschlag am europäischen Markt, WPg, vol. 75, no. 22 (2022): 1291-1301.
3 See Grbenic, The Performance of Option Pricing Models Estimating the Marketability Discount in a Pre-IPO Real-World Data Setting: Evidence from Europe, Journal of Business Valuation and Economic Loss Analysis, vol. 17, no. 1 (2022): 1-37.
4 See Longstaff, How Much Can Marketability Affect Security Values?, The Journal of Finance, vol. 50, no. 5 (2005): 1767-1774.
5 See Abbott, Discount for Lack of Liquidity: Understanding and Interpreting Option Models, Business Val-uation Review, vol. 28, no. 3 (2009): 114-148.
6 See Ghaidarov, The Cost of Illiquidity for Private Equity Investments, Working Paper, 2010: 1-28.

missing were either omitted or missing or invalid stock returns, respectively, were replaced employing the Uniform (Average) Returns Procedure

$$r_{i,t} = \sqrt[d+1]{\frac{p_{i,t+1+j}}{p_{i,t-d+j}}}$$

where i is the index on the stocks related to DLOM, $r_{i,t}$ is the return of stock i at day t, $p_{i,t}$ is the price of stock i at day t, d is the length (number of days) of the non-trading interval and, j is the number of remaining days without trading at day t in the non-trading interval.

The dividend yield q_i of the underlying stock is computed in a sustainable shape[7]

$$q_i = \ln\left[\left(1+\frac{EPS_i}{PPS_i}\right)\cdot\left(1-\frac{g_i}{ROE_i}\right)\right]$$

where EPS_i are the earnings per share of stock i, PPS_i is the price of stock i as on end of computation period date, ROE_i is the return on equity of stock i and, g_i is the compound annual growth rate of operating sales over the preceding 5 years.

The data is evaluated carefully; however, the author denies liability for the accuracy of all computations.

Notes for application:

n indicates the number of DLOM (sample size) computed. $\overline{x}_a$ indicates the arithmetic mean, $\overline{x}_h$ indicates the harmonic mean

$$\bar{x}_h = \frac{n}{\sum_{i=1}^{n}\frac{n}{x_i}}$$

and $\overline{x}_t$ indicates the truncated mean (10% level = 10 % of the observations sorted in ascending order being eliminated upside and downside)

$$\bar{x}_t = \frac{\sum_{2}^{n-1} x_i}{n-2}$$

The first quartile $Q1$ indicates the boundary of the lowest 25%, the third quartile $Q3$ indicates the boundary of the highest 25% of the computed DLOM. Using this information, the effectively employed DLOM may be related to the group of the 25% lowest (highest) discounts computed. $Q2$ indicates the median of the DLOM computed. The confidence interval reports the range (lower confidence limit to upper confidence limit) of the DLOM applying a 95% confidence level. Assuming the DLOM to be normally distributed, this indicates all DLOM lying within these limits.

To evaluate the assumption of normally distributed DLOM computed, the p-value for the Jarque-Bera Test for Normality

$$JB = n\left[\frac{(\text{skewness})^2}{6} + \frac{(\text{kurtosis}-3)^2}{24}\right]$$

is reported in brackets. P-values below (above) the defined level of significance (0.01, 0.05 or 0.10) indicate that the null hypothesis of the DLOMs being normally distributed is rejected (accepted). Consequently, a p-value above (below) the defined level of significance indicates the DLOMs (not) to be normally distributed.

The skewness sk indicates the symmetry of the distribution of the computed DLOM. A negative skewness indicates the distribution to be skewed to the left, whereas a positive skewness indicates the distribution to be skewed to the right (a skewness of zero indicates the distribution to be symmetric). The kurtosis kurt indicates the weight in the tails of the distribution of the computed DLOM (for the normal distribution, the kurtosis is 3). The standard deviation sd indicates the dispersion of the computed DLOM. Finally, the coefficient of variation cv indicates the dispersion of the computed DLOM adjusting for the scale of units in the DLOM, expressed by the standard deviation as a percentage of the mean. It allows for a comparison of the dispersion of the DLOM across countries/regions. A lower (higher) coefficient of variation indicates a lower (higher) dispersion of the computed DLOM and, similarly, a higher (lower) reliability. ◆

---

7 See Ghaidarov, Analysis and Critique of the Average Strike Put Option Marketability Discount Model, White Paper, 2009: 1-15; Ghaidarov, The Cost of Illiquidity for Private Equity Investments, Working Paper, 2010: 1-28.

Lookback Put OPM, Adjusted Lookback Put OPM and Perpetual Exchange Put OPM, 2024, Holding Period = 3 months

Country / Region	n	$\bar{x}_a$	$\bar{x}_h$	$\bar{x}_t$	Q_1	Q_2	Q_3	95% (JB)	sk	kurt	sd	cv
Austria	107	16.34%	10.61%	15.07%	9.20%	11.13%	20.47%	[14,00% ; 18,68%] (0,0000)	1.81	3.31	0.12	0.75
Belgium	244	19.96%	13.05%	17.39%	9.38%	12.95%	23.54%	[17,46% ; 22,46%] (0,0000)	5.27	46.23	0.20	0.99
Bosnia and Herzegovina	22	16.75%	11.55%	14.87%	7.82%	13.85%	18.94%	[10,95% ; 22,54%] (0,0000)	2.60	8.11	0.13	0.78
Bulgaria	26	11.42%	9.47%	11.21%	7.46%	11.35%	13.46%	[9,51% ; 13,33%] (0,4004)	0.64	0.21	0.05	0.41
Croatia	27	11.21%	9.22%	10.50%	7.63%	9.68%	13.08%	[9,03% ; 13,40%] (0,0000)	3.23	14.16	0.06	0.58
Cyprus	83	21.21%	12.08%	17.18%	10.06%	13.72%	19.61%	[15,52% ; 26,90%] (0,0000)	4.16	19.87	0.26	1.23
Czech Republic	40	17.84%	9.87%	14.15%	6.92%	11.89%	16.83%	[11,07% ; 24,61%] (0,0000)	3.05	9.77	0.21	1.19
Denmark	329	25.31%	15.49%	20.26%	11.37%	16.52%	27.12%	[19,26% ; 31,36%] (0,0000)	15.55	265.78	0.56	2.20
Estonia	44	12.35%	7.01%	9.10%	5.13%	7.31%	12.53%	[6,80% ; 17,90%] (0,0000)	4.92	27.00	0.18	1.48
Finland	401	18.31%	13.75%	16.86%	10.90%	14.35%	20.63%	[17,13% ; 19,50%] (0,0000)	2.61	10.62	0.12	0.66
France	1,021	21.54%	14.29%	19.40%	10.54%	16.04%	27.15%	[20,39% ; 22,68%] (0,0000)	5.16	53.44	0.19	0.86
Germany	1,043	23.47%	15.46%	21.44%	12.02%	16.78%	29.83%	[22,38% ; 24,56%] (0,0000)	2.18	7.60	0.18	0.76
Greece	171	18.49%	12.05%	14.45%	9.41%	12.80%	17.51%	[12,48% ; 24,50%] (0,0000)	11.81	148.29	0.40	2.15
Hungary	68	17.48%	13.04%	16.16%	9.05%	15.25%	21.70%	[14,67% ; 20,28%] (0,0000)	2.64	10.74	0.12	0.66
Iceland	58	17.85%	9.09%	14.81%	7.49%	9.13%	13.89%	[11,47% ; 24,23%] (0,0000)	4.45	25.02	0.24	1.36
Ireland	139	21.51%	13.94%	19.83%	9.84%	13.37%	32.18%	[18,69% ; 24,33%] (0,0000)	1.59	2.21	0.17	0.78
Italy	593	15.32%	12.35%	14.05%	10.12%	12.78%	16.82%	[14,56% ; 16,09%] (0,0000)	3.14	14.03	0.09	0.62
Kazakhstan	26	8.93%	2.03%	8.88%	6.09%	8.58%	10.83%	[7,05% ; 10,81%] (0,7679)	0.34	0.13	0.05	0.52
Lithuania	45	9.73%	7.67%	8.82%	6.87%	8.61%	10.62%	[7,86% ; 11,60%] (0,0000)	3.27	13.62	0.06	0.64
Luxembourg	121	20.94%	16.11%	19.41%	12.45%	15.62%	25.36%	[18,56% ; 23,32%] (0,0000)	1.75	2.49	0.13	0.63
Malta	21	16.58%	13.31%	16.58%	9.50%	14.23%	18.80%	[10,84% ; 22,31%] (0,0228)	1.59	2.23	0.09	0.54
Netherlands	295	22.09%	14.51%	19.93%	10.77%	16.12%	28.19%	[19,92% ; 24,26%] (0,0000)	4.60	39.38	0.19	0.86
North Macedonia	85	13.64%	1.13%	10.69%	4.85%	8.16%	15.40%	[9,12% ; 18,17%] (0,0000)	5.72	41.34	0.21	1.54
Norway	411	22.91%	15.41%	20.34%	11.71%	16.83%	26.96%	[20,91% ; 24,91%] (0,0000)	4.67	36.23	0.21	0.90
Poland	1,204	21.48%	17.09%	20.27%	13.45%	18.42%	26.30%	[20,81% ; 22,15%] (0,0000)	2.53	13.12	0.12	0.55
Portugal	58	13.97%	10.69%	13.22%	7.55%	11.95%	15.90%	[11,79% ; 16,15%] (0,0000)	1.81	3.99	0.08	0.59
Romania	174	15.76%	12.57%	14.78%	9.49%	13.54%	18.78%	[14,40% ; 17,12%] (0,0000)	2.47	9.75	0.09	0.58
Russia	323	27.10%	16.61%	18.49%	13.54%	17.37%	22.82%	[17,62% ; 36,59%] (0,0000)	13.80	210.78	0.87	3.20
Slovenia	24	10.89%	9.15%	10.39%	7.41%	8.72%	11.62%	[8,58% ; 13,20%] (0,0002)	1.69	2.30	0.05	0.50
Spain	287	16.05%	11.83%	14.70%	9.26%	12.24%	17.89%	[14,77% ; 17,33%] (0,0000)	2.13	5.22	0.11	0.68
Sweden	1,671	30.67%	20.76%	28.18%	14.98%	24.26%	39.85%	[29,59% ; 31,76%] (0,0000)	2.97	19.79	0.23	0.74
Switzerland	542	19.00%	11.04%	16.90%	8.71%	11.99%	24.37%	[17,48% ; 20,53%] (0,0000)	4.01	34.23	0.18	0.95
Turkey	654	23.42%	21.42%	22.53%	18.31%	21.67%	26.49%	[22,77% ; 24,07%] (0,0000)	2.85	13.80	0.08	0.36
United Kingdom	2,531	19.29%	11.69%	17.73%	9.48%	14.62%	24.75%	[18,72% ; 19,86%] (0,0000)	2.03	6.72	0.15	0.76
Central and Western Europe	3,378	21.46%	13.77%	19.35%	10.40%	15.22%	26.96%	[20,84% ; 22,07%] (0,0000)	3.95	33.84	0.18	0.85
Southern Europe	1,910	18.56%	14.01%	17.11%	10.96%	15.93%	22.35%	[17,83% ; 19,28%] (0,0000)	16.92	491.50	0.16	0.87
Scandinavia	2,870	26.96%	17.50%	24.14%	12.90%	19.53%	34.10%	[25,94% ; 27,98%] (0,0000)	15.70	497.39	0.28	1.03
Britain	2,670	19.41%	11.79%	17.83%	9.50%	14.55%	24.90%	[18,85% ; 19,96%] (0,0000)	2.00	6.35	0.15	0.76
Eastern Europe	2,086	20.41%	9.40%	18.01%	11.72%	16.23%	23.52%	[18,86% ; 21,97%] (0,0000)	29.75	1,089.37	0.36	1.77
Total	12,914	21.66%	12.99%	19.42%	11.05%	16.17%	26.31%	[21,25% ; 22,07%] (0,0000)	24.18	1,185.97	0.24	1.09

Lookback Put OPM, Adjusted Lookback Put OPM and Perpetual Exchange Put OPM, 2024, Holding Period = 6 months

Country / Region	n	$\bar{x}_a$	$\bar{x}_h$	$\bar{x}_t$	Q_1	Q_2	Q_3	95% (JB)	sk	kurt	sd	cv
Austria	107	19.94%	14.42%	18.64%	12.81%	15.68%	24.17%	[17,56% ; 22,31%] (0,0000)	1.82	3.58	0.12	0.62
Belgium	244	25.22%	17.56%	21.64%	13.09%	17.50%	28.42%	[21,62% ; 28,82%] (0,0000)	8.45	98.98	0.29	1.13
Bosnia and Herzegovina	22	23.73%	16.22%	20.69%	11.05%	19.43%	26.10%	[15,08% ; 32,39%] (0,0000)	2.92	10.36	0.20	0.82
Bulgaria	26	16.10%	13.32%	15.76%	10.45%	15.59%	19.43%	[13,36% ; 18,85%] (0,2400)	0.76	0.58	0.07	0.42
Croatia	27	15.09%	12.87%	14.56%	10.92%	13.73%	18.27%	[12,82% ; 17,36%] (0,0000)	1.95	6.16	0.07	0.44
Cyprus	83	29.42%	16.83%	22.87%	14.26%	19.02%	28.04%	[20,51% ; 38,34%] (0,0000)	5.01	27.39	0.41	1.39
Czech Republic	40	25.24%	13.69%	19.11%	9.61%	15.68%	23.60%	[14,79% ; 35,69%] (0,0000)	3.41	12.24	0.33	1.29
Denmark	329	35.19%	21.20%	26.62%	16.05%	23.26%	34.95%	[23,90% ; 46,47%] (0,0000)	16.74	294.27	1.04	2.96
Estonia	44	17.34%	9.80%	12.45%	7.20%	10.40%	17.67%	[8,73% ; 25,95%] (0,0000)	5.47	32.70	0.28	1.63
Finland	401	23.65%	18.37%	21.87%	15.29%	19.66%	27.22%	[22,17% ; 25,13%] (0,0000)	4.19	29.96	0.15	0.64
France	1,021	27.92%	19.39%	24.90%	14.78%	21.62%	32.86%	[26,28% ; 29,56%] (0,0000)	8.50	120.95	0.27	0.96
Germany	1,043	29.14%	21.01%	27.06%	16.77%	23.27%	34.74%	[27,92% ; 30,37%] (0,0000)	3.79	34.37	0.20	0.69
Greece	171	25.43%	16.54%	19.12%	13.30%	17.59%	23.77%	[14,63% ; 36,24%] (0,0000)	12.66	163.54	0.72	2.81
Hungary	68	24.05%	18.24%	22.54%	12.98%	20.73%	30.91%	[20,54% ; 27,55%] (0,0000)	1.71	3.40	0.14	0.60
Iceland	58	24.22%	11.96%	19.00%	10.53%	12.97%	20.07%	[14,25% ; 34,19%] (0,0000)	5.45	34.71	0.38	1.57
Ireland	139	27.66%	19.09%	25.10%	13.89%	18.36%	35.35%	[24,03% ; 31,29%] (0,0000)	2.41	7.59	0.22	0.78
Italy	593	20.54%	17.11%	19.27%	14.09%	17.69%	23.61%	[19,66% ; 21,42%] (0,0000)	2.82	12.28	0.11	0.53
Kazakhstan	26	12.60%	2.87%	12.50%	8.53%	12.07%	15.57%	[9,92% ; 15,28%] (0,6473)	0.42	0.33	0.07	0.53
Lithuania	45	13.62%	10.74%	12.31%	9.71%	12.16%	15.01%	[10,92% ; 16,32%] (0,0000)	3.56	16.26	0.09	0.66
Luxembourg	121	26.72%	21.89%	24.99%	17.38%	21.51%	30.75%	[24,07% ; 29,38%] (0,0000)	2.14	5.67	0.15	0.55
Malta	21	21.01%	18.22%	21.01%	13.63%	19.85%	27.35%	[15,78% ; 26,24%] (0,4916)	0.84	0.11	0.08	0.39
Netherlands	295	28.24%	19.64%	25.47%	15.33%	21.77%	33.87%	[25,21% ; 31,28%] (0,0000)	7.81	94.06	0.26	0.94
North Macedonia	85	19.21%	1.60%	14.54%	6.78%	11.68%	20.92%	[12,03% ; 26,38%] (0,0000)	6.65	52.59	0.33	1.73
Norway	411	30.91%	20.92%	26.66%	16.01%	22.77%	35.17%	[27,82% ; 33,99%] (0,0000)	6.24	58.07	0.32	1.03
Poland	1,204	30.41%	23.91%	28.42%	18.79%	25.55%	36.32%	[29,39% ; 31,43%] (0,0000)	3.04	18.22	0.18	0.59
Portugal	58	18.20%	14.65%	17.55%	10.63%	16.25%	21.48%	[15,79% ; 20,60%] (0,0000)	1.30	1.61	0.09	0.50
Romania	174	22.27%	17.64%	20.77%	13.44%	18.79%	26.55%	[20,26% ; 24,28%] (0,0000)	2.83	12.75	0.13	0.60
Russia	323	41.52%	23.22%	26.00%	18.83%	24.39%	32.01%	[23,40% ; 59,63%] (0,0000)	14.26	222.44	1.65	3.99
Slovenia	24	15.36%	12.88%	14.59%	10.36%	12.43%	16.41%	[12,04% ; 18,68%] (0,0000)	1.81	3.01	0.08	0.51
Spain	287	20.46%	16.17%	19.22%	13.00%	16.79%	23.53%	[19,10% ; 21,82%] (0,0000)	1.91	4.17	0.12	0.57
Sweden	1,671	41.25%	28.18%	36.85%	20.67%	32.50%	50.18%	[39,61% ; 42,88%] (0,0000)	4.42	39.19	0.34	0.83
Switzerland	542	23.22%	14.92%	20.61%	12.07%	16.38%	28.75%	[21,25% ; 25,19%] (0,0000)	8.04	112.61	0.23	1.00
Turkey	654	32.69%	29.78%	31.56%	25.10%	29.91%	37.39%	[31,81% ; 33,58%] (0,0000)	2.10	7.40	0.12	0.35
United Kingdom	2,531	24.87%	15.90%	22.85%	13.15%	19.80%	31.20%	[24,15% ; 25,59%] (0,0000)	2.93	17.66	0.19	0.75
Central and Western Europe	3,378	27.06%	18.67%	24.48%	14.51%	20.67%	32.40%	[26,26% ; 27,86%] (0,0000)	7.55	106.53	0.24	0.88
Southern Europe	1,910	25.24%	19.37%	23.38%	15.36%	21.90%	30.29%	[24,08% ; 26,41%] (0,0000)	24.63	845.60	0.26	1.03
Scandinavia	2,870	36.27%	23.69%	31.49%	18.00%	26.17%	42.41%	[34,56% ; 37,97%] (0,0000)	22.75	845.03	0.47	1.28
Britain	2,670	25.01%	16.04%	22.96%	13.21%	19.77%	31.35%	[24,30% ; 25,72%] (0,0000)	2.90	16.81	0.19	0.75
Eastern Europe	2,086	29.35%	13.20%	25.20%	16.38%	22.78%	32.90%	[26,44% ; 32,26%] (0,0000)	32.61	1,246.41	0.68	2.31
Total	12,914	28.78%	17.81%	25.53%	15.39%	22.14%	34.01%	[28,10% ; 29,47%] (0,0000)	36.47	2,134.95	0.40	1.37

Lookback Put OPM, Adjusted Lookback Put OPM and Perpetual Exchange Put OPM, 2024, Holding Period = 9 months

Country / Region	n	x̄_a	x̄_h	x̄_t	Q₁	Q₂	Q₃	95% (JB)	sk	kurt	sd	cv
Austria	107	22.71%	17.14%	21.29%	15.51%	19.12%	25.82%	[20,17% ; 25,24%] (0,0000)	2.13	6.00	0.13	0.58
Belgium	244	29.40%	20.72%	24.76%	15.70%	20.36%	30.98%	[24,63% ; 34,18%] (0,0000)	9.96	126.33	0.38	1.29
Bosnia and Herzegovina	22	29.18%	19.76%	25.09%	13.53%	23.67%	30.76%	[18,04% ; 40,33%] (0,0000)	3.14	11.87	0.25	0.86
Bulgaria	26	19.70%	16.25%	19.22%	12.70%	18.68%	24.18%	[16,28% ; 23,12%] (0,1311)	0.86	0.89	0.08	0.43
Croatia	27	18.06%	15.62%	17.67%	13.45%	16.56%	21.10%	[15,62% ; 20,50%] (0,0001)	1.27	2.31	0.07	0.40
Cyprus	83	36.11%	20.38%	27.12%	17.67%	23.11%	34.38%	[24,06% ; 48,17%] (0,0000)	5.36	30.47	0.55	1.53
Czech Republic	40	31.19%	16.52%	22.81%	11.70%	17.27%	29.03%	[17,40% ; 44,98%] (0,0000)	3.62	13.65	0.43	1.38
Denmark	329	43.52%	25.33%	31.32%	19.35%	27.54%	41.15%	[26,95% ; 60,08%] (0,0000)	17.07	302.33	1.53	3.51
Estonia	44	21.35%	11.91%	15.02%	8.79%	12.69%	19.64%	[9,93% ; 32,77%] (0,0000)	5.76	35.60	0.38	1.76
Finland	401	27.78%	21.58%	25.55%	18.31%	22.97%	31.33%	[25,95% ; 29,62%] (0,0000)	5.27	44.15	0.19	0.67
France	1,021	32.97%	23.03%	28.92%	17.73%	25.09%	38.39%	[30,80% ; 35,13%] (0,0000)	10.25	160.74	0.35	1.07
Germany	1,043	33.55%	24.97%	31.15%	20.31%	27.76%	39.67%	[32,12% ; 34,99%] (0,0000)	5.66	69.80	0.24	0.70
Greece	171	31.31%	19.81%	22.53%	16.25%	21.00%	28.24%	[15,67% ; 46,94%] (0,0000)	12.85	167.02	1.04	3.31
Hungary	68	29.12%	22.13%	27.41%	15.85%	24.75%	37.99%	[24,91% ; 33,33%] (0,0000)	1.60	2.63	0.17	0.60
Iceland	58	29.42%	13.91%	22.15%	12.82%	16.03%	24.06%	[15,99% ; 42,86%] (0,0000)	5.88	39.03	0.51	1.74
Ireland	139	32.48%	22.82%	28.84%	16.78%	22.13%	37.60%	[27,97% ; 36,98%] (0,0000)	3.03	11.72	0.27	0.83
Italy	593	24.55%	20.61%	23.14%	17.04%	21.44%	28.31%	[23,53% ; 25,58%] (0,0000)	3.05	15.65	0.13	0.52
Kazakhstan	26	15.41%	3.51%	15.25%	10.38%	14.73%	19.28%	[12,10% ; 18,71%] (0,5213)	0.49	0.51	0.08	0.53
Lithuania	45	16.61%	13.05%	14.99%	11.74%	14.42%	17.99%	[13,22% ; 20,00%] (0,0000)	3.76	18.15	0.11	0.68
Luxembourg	121	31.19%	26.02%	29.03%	20.85%	25.78%	33.77%	[28,10% ; 34,28%] (0,0000)	2.78	10.70	0.17	0.55
Malta	21	24.40%	21.77%	24.40%	16.89%	24.08%	30.92%	[19,18% ; 29,63%] (0,6028)	0.41	-1.16	0.08	0.34
Netherlands	295	33.10%	23.29%	29.47%	17.74%	25.56%	38.13%	[29,11% ; 37,08%] (0,0000)	9.49	125.19	0.35	1.05
North Macedonia	85	23.70%	1.96%	17.47%	8.22%	14.45%	24.86%	[14,02% ; 33,37%] (0,0000)	7.13	58.40	0.45	1.89
Norway	411	37.29%	24.84%	31.33%	19.14%	27.00%	41.29%	[33,14% ; 41,45%] (0,0000)	6.98	69.54	0.43	1.15
Poland	1,204	37.37%	29.03%	34.62%	22.73%	30.93%	43.87%	[36,04% ; 38,70%] (0,0000)	3.38	21.89	0.24	0.63
Portugal	58	21.44%	17.52%	20.70%	13.05%	19.47%	25.31%	[18,70% ; 24,18%] (0,0000)	1.31	1.88	0.10	0.49
Romania	174	27.29%	21.48%	25.32%	16.46%	22.41%	32.39%	[24,73% ; 29,86%] (0,0000)	3.09	14.98	0.17	0.63
Russia	323	54.14%	28.19%	31.76%	22.90%	29.52%	39.30%	[27,38% ; 80,89%] (0,0000)	14.42	226.65	2.44	4.51
Slovenia	24	18.80%	15.71%	17.78%	12.57%	15.38%	20.37%	[14,66% ; 22,93%] (0,0000)	1.92	3.61	0.10	0.52
Spain	287	23.86%	19.30%	22.42%	15.50%	20.07%	27.75%	[22,34% ; 25,37%] (0,0000)	2.09	5.70	0.13	0.55
Sweden	1,671	49.72%	33.46%	43.36%	24.85%	37.98%	57.09%	[47,53% ; 51,91%] (0,0000)	5.15	50.37	0.46	0.92
Switzerland	542	26.54%	17.68%	23.36%	14.64%	19.53%	30.82%	[24,05% ; 29,04%] (0,0000)	10.74	173.26	0.30	1.11
Turkey	654	39.87%	35.98%	38.43%	29.74%	35.86%	45.84%	[38,74% ; 41,01%] (0,0000)	2.10	7.73	0.15	0.37
United Kingdom	2,531	29.21%	18.91%	26.56%	15.76%	23.41%	35.76%	[28,32% ; 30,10%] (0,0000)	3.73	27.56	0.23	0.78
Central and Western Europe	3,378	31.45%	22.15%	28.19%	17.46%	24.57%	36.80%	[30,43% ; 32,48%] (0,0000)	9.92	161.45	0.30	0.97
Southern Europe	1,910	30.47%	23.30%	28.05%	18.49%	26.14%	35.98%	[28,85% ; 32,09%] (0,0000)	27.92	1,008.60	0.36	1.19
Scandinavia	2,870	43.75%	28.08%	36.91%	21.52%	31.21%	48.53%	[41,35% ; 46,16%] (0,0000)	25.87	1,014.94	0.66	1.50
Britain	2,670	29.38%	19.08%	26.67%	15.83%	23.35%	35.89%	[28,50% ; 30,25%] (0,0000)	3.68	26.19	0.23	0.78
Eastern Europe	2,086	36.52%	16.08%	30.67%	19.79%	27.89%	39.94%	[32,26% ; 40,79%] (0,0000)	33.74	1,310.03	0.99	2.72
Total	12,914	34.43%	21.30%	29.99%	18.45%	26.37%	39.47%	[33,47% ; 35,40%] (0,0000)	41.58	2,572.48	0.56	1.62

Lookback Put OPM, Adjusted Lookback Put OPM and Perpetual Exchange Put OPM, 2024, Holding Period = 1 year

Country / Region	n	$\bar{x}_a$	$\bar{x}_h$	$\bar{x}_t$	Q_1	Q_2	Q_3	95% (JB)	sk	kurt	sd	cv
Austria	107	25.05%	19.31%	23.51%	17.83%	21.92%	28.10%	[22,30% ; 27,80%] (0,0000)	2.55	9.44	0.14	0.57
Belgium	244	33.05%	23.22%	27.27%	17.89%	23.16%	33.34%	[27,08% ; 39,03%] (0,0000)	10.77	141.50	0.47	1.43
Bosnia and Herzegovina	22	33.86%	22.71%	28.76%	15.62%	26.82%	34.72%	[20,39% ; 47,33%] (0,0000)	3.30	12.93	0.30	0.90
Bulgaria	26	22.72%	18.70%	22.13%	14.77%	21.17%	27.72%	[18,71% ; 26,74%] (0,0691)	0.95	1.16	0.10	0.44
Croatia	27	20.56%	17.89%	20.28%	15.40%	18.94%	23.74%	[17,91% ; 23,22%] (0,0333)	0.98	0.83	0.08	0.38
Cyprus	83	42.06%	23.30%	30.68%	19.91%	27.03%	37.50%	[26,90% ; 57,22%] (0,0000)	5.54	32.08	0.69	1.65
Czech Republic	40	36.42%	18.84%	25.88%	13.53%	18.87%	32.84%	[19,44% ; 53,41%] (0,0000)	3.75	14.55	0.53	1.46
Denmark	329	51.08%	28.64%	35.17%	21.92%	30.95%	45.85%	[29,22% ; 72,94%] (0,0000)	17.22	306.05	2.02	3.95
Estonia	44	24.87%	13.66%	17.18%	10.17%	14.61%	21.07%	[10,73% ; 39,01%] (0,0000)	5.93	37.34	0.47	1.87
Finland	401	31.31%	24.09%	28.58%	20.35%	26.11%	35.54%	[29,11% ; 33,50%] (0,0000)	5.93	53.11	0.22	0.71
France	1,021	37.33%	25.92%	32.21%	20.11%	28.27%	42.31%	[34,62% ; 40,04%] (0,0000)	11.27	185.49	0.44	1.18
Germany	1,043	37.33%	28.12%	34.47%	23.16%	31.26%	44.19%	[35,65% ; 39,00%] (0,0000)	7.13	100.21	0.28	0.74
Greece	171	36.65%	22.46%	25.32%	18.21%	24.10%	31.46%	[16,17% ; 57,13%] (0,0000)	12.93	168.33	1.36	3.70
Hungary	68	33.43%	25.36%	31.35%	18.34%	27.75%	42.81%	[28,54% ; 38,33%] (0,0000)	1.67	2.98	0.20	0.61
Iceland	58	34.06%	15.40%	24.78%	14.70%	18.45%	25.75%	[17,22% ; 50,90%] (0,0000)	6.12	41.49	0.64	1.88
Ireland	139	36.63%	25.83%	31.97%	19.12%	24.74%	41.19%	[31,23% ; 42,02%] (0,0000)	3.42	14.37	0.32	0.88
Italy	593	27.95%	23.48%	26.36%	19.40%	24.42%	32.06%	[26,77% ; 29,13%] (0,0000)	3.35	19.47	0.15	0.52
Kazakhstan	26	17.77%	4.05%	17.56%	11.93%	16.96%	22.17%	[13,92% ; 21,62%] (0,4065)	0.55	0.67	0.10	0.54
Lithuania	45	19.14%	14.97%	17.24%	13.40%	16.79%	20.49%	[15,13% ; 23,14%] (0,0000)	3.92	19.68	0.13	0.70
Luxembourg	121	35.00%	29.32%	32.38%	23.29%	29.73%	38.61%	[31,41% ; 38,58%] (0,0000)	3.28	14.59	0.20	0.57
Malta	21	27.27%	24.62%	27.27%	19.69%	27.60%	32.24%	[21,76% ; 32,78%] (0,5863)	0.39	-1.24	0.09	0.32
Netherlands	295	37.30%	26.18%	32.72%	20.02%	28.85%	42.61%	[32,34% ; 42,26%] (0,0000)	10.43	143.19	0.43	1.16
North Macedonia	85	27.66%	2.26%	19.93%	9.42%	16.58%	27.99%	[15,54% ; 39,79%] (0,0000)	7.42	62.01	0.56	2.03
Norway	411	42.87%	27.96%	35.23%	21.63%	30.33%	45.63%	[37,66% ; 48,09%] (0,0000)	7.43	76.93	0.54	1.26
Poland	1,204	43.33%	33.26%	39.84%	25.93%	35.30%	50.54%	[41,70% ; 44,96%] (0,0000)	3.63	24.74	0.29	0.66
Portugal	58	24.18%	19.84%	23.23%	14.38%	21.30%	28.70%	[21,07% ; 27,28%] (0,0000)	1.45	2.74	0.12	0.49
Romania	174	31.57%	24.68%	29.14%	18.76%	25.47%	37.25%	[28,49% ; 34,66%] (0,0000)	3.30	16.73	0.21	0.65
Russia	323	65.86%	32.29%	36.61%	25.98%	33.63%	44.04%	[30,46% ; 101,26%] (0,0000)	14.50	228.77	3.23	4.91
Slovenia	24	21.69%	18.08%	20.45%	14.41%	17.91%	23.79%	[16,83% ; 26,55%] (0,0000)	2.01	4.11	0.12	0.53
Spain	287	26.73%	21.82%	25.04%	17.62%	22.60%	30.22%	[25,04% ; 28,42%] (0,0000)	2.36	7.94	0.15	0.54
Sweden	1,671	57.16%	37.65%	48.90%	28.32%	41.86%	62.68%	[54,42% ; 59,90%] (0,0000)	5.59	57.83	0.57	1.00
Switzerland	542	29.40%	19.86%	25.61%	16.26%	21.88%	33.27%	[26,35% ; 32,45%] (0,0000)	12.37	212.88	0.36	1.23
Turkey	654	45.99%	41.05%	44.22%	33.57%	40.78%	53.28%	[44,60% ; 47,38%] (0,0000)	2.15	8.17	0.18	0.39
United Kingdom	2,531	32.93%	21.32%	29.59%	17.78%	26.10%	39.52%	[31,87% ; 33,99%] (0,0000)	4.30	35.07	0.27	0.82
Central and Western Europe	3,378	35.24%	24.92%	31.19%	19.80%	27.67%	40.24%	[33,98% ; 36,50%] (0,0000)	11.36	197.59	0.37	1.06
Southern Europe	1,910	34.95%	26.51%	31.91%	21.04%	29.59%	40.56%	[32,86% ; 37,03%] (0,0000)	29.70	1,100.18	0.46	1.33
Scandinavia	2,870	50.34%	31.54%	41.47%	24.47%	35.03%	53.25%	[47,23% ; 53,45%] (0,0000)	27.64	1,115.36	0.85	1.69
Britain	2,670	33.12%	21.52%	29.69%	17.86%	26.09%	39.67%	[32,08% ; 34,16%] (0,0000)	4.24	33.26	0.27	0.83
Eastern Europe	2,086	42.82%	18.48%	35.25%	22.59%	31.97%	45.62%	[37,20% ; 48,44%] (0,0000)	34.34	1,344.17	1.31	3.06
Total	12,914	39.34%	24.11%	33.66%	20.87%	29.68%	43.88%	[38,09% ; 40,59%] (0,0000)	44.31	2,816.43	0.72	1.84

Lookback Put OPM, Adjusted Lookback Put OPM and Perpetual Exchange Put OPM, 2024, Holding Period = 1.5 years

Country / Region	n	xa	xh	xt	Q1	Q2	Q3	95% (JB)	sk	kurt	sd	cv
Austria	107	29.00%	22.72%	27.13%	20.83%	25.64%	32.33%	[25,72% ; 32,28%] (0,0000)	3.30	15.80	0.17	0.59
Belgium	244	39.44%	27.08%	31.42%	21.13%	27.17%	38.02%	[31,05% ; 47,84%] (0,0000)	11.59	157.42	0.67	1.69
Bosnia and Herzegovina	22	41.91%	27.58%	34.89%	19.13%	31.57%	42.84%	[24,04% ; 59,79%] (0,0000)	3.52	14.34	0.40	0.96
Bulgaria	26	27.82%	22.76%	27.00%	18.20%	25.14%	33.08%	[22,73% ; 32,91%] (0,0194)	1.09	1.58	0.13	0.45
Croatia	27	24.76%	21.62%	24.52%	18.61%	22.78%	28.97%	[21,63% ; 27,90%] (0,0877)	0.88	0.41	0.09	0.37
Cyprus	83	52.73%	28.06%	36.62%	24.49%	32.97%	43.22%	[31,41% ; 74,05%] (0,0000)	5.72	33.68	0.98	1.85
Czech Republic	40	45.72%	22.60%	30.96%	16.04%	22.89%	38.29%	[22,55% ; 68,90%] (0,0000)	3.91	15.64	0.72	1.58
Denmark	329	64.92%	33.89%	41.52%	26.07%	36.23%	51.46%	[32,46% ; 97,38%] (0,0000)	17.36	309.59	2.99	4.61
Estonia	44	31.11%	16.53%	20.82%	12.49%	17.81%	25.90%	[11,68% ; 50,55%] (0,0000)	6.13	39.31	0.64	2.05
Finland	401	37.32%	27.92%	33.55%	24.19%	30.73%	41.11%	[34,40% ; 40,25%] (0,0000)	6.64	63.31	0.30	0.80
France	1,021	44.93%	30.44%	37.57%	23.67%	32.94%	47.96%	[41,14% ; 48,73%] (0,0000)	12.41	214.87	0.62	1.37
Germany	1,043	43.80%	33.06%	39.78%	27.60%	36.16%	49.67%	[41,60% ; 45,99%] (0,0000)	8.96	141.17	0.36	0.83
Greece	171	46.42%	26.67%	29.96%	21.77%	28.40%	37.81%	[16,22% ; 76,62%] (0,0000)	12.98	169.33	2.00	4.31
Hungary	68	40.77%	30.64%	37.85%	22.80%	32.62%	51.97%	[34,53% ; 47,01%] (0,0000)	1.86	3.93	0.26	0.63
Iceland	58	42.40%	17.65%	29.17%	17.78%	22.48%	31.01%	[18,82% ; 65,99%] (0,0000)	6.38	44.20	0.90	2.12
Ireland	139	43.82%	30.60%	37.21%	22.90%	29.59%	46.19%	[36,66% ; 50,98%] (0,0000)	3.84	17.33	0.43	0.97
Italy	593	33.69%	28.11%	31.67%	23.15%	29.53%	38.41%	[32,20% ; 35,18%] (0,0000)	3.84	25.20	0.18	0.55
Kazakhstan	26	21.74%	4.96%	21.42%	14.45%	20.67%	27.32%	[16,94% ; 26,55%] (0,2328)	0.67	0.95	0.12	0.55
Lithuania	45	23.40%	18.14%	21.02%	16.10%	20.33%	25.51%	[18,27% ; 28,53%] (0,0000)	4.17	22.04	0.17	0.73
Luxembourg	121	41.48%	34.47%	37.83%	28.25%	34.23%	44.67%	[36,84% ; 46,11%] (0,0000)	3.83	18.87	0.26	0.62
Malta	21	32.09%	29.16%	32.09%	24.50%	31.58%	35.37%	[25,56% ; 38,62%] (0,5373)	0.78	-0.22	0.10	0.32
Netherlands	295	44.62%	30.68%	37.97%	23.40%	33.23%	48.40%	[37,69% ; 51,56%] (0,0000)	11.41	162.79	0.61	1.36
North Macedonia	85	34.71%	2.76%	23.90%	11.38%	20.35%	32.97%	[17,75% ; 51,67%] (0,0000)	7.77	66.32	0.79	2.27
Norway	411	52.71%	32.83%	41.77%	25.61%	35.50%	53.60%	[45,39% ; 60,03%] (0,0000)	7.97	86.32	0.75	1.43
Poland	1,204	53.59%	40.19%	48.65%	31.33%	42.79%	61.05%	[51,40% ; 55,79%] (0,0000)	3.98	28.97	0.39	0.72
Portugal	58	28.79%	23.53%	27.44%	17.16%	25.41%	33.95%	[24,94% ; 32,64%] (0,0000)	1.73	4.19	0.15	0.51
Romania	174	38.85%	29.95%	35.49%	22.82%	31.59%	45.76%	[34,79% ; 42,90%] (0,0000)	3.60	19.30	0.27	0.70
Russia	323	87.81%	39.00%	44.77%	31.02%	40.79%	54.00%	[35,09% ; 140,53%] (0,0000)	14.58	230.85	4.82	5.48
Slovenia	24	26.57%	22.01%	24.92%	17.53%	22.09%	29.70%	[20,41% ; 32,73%] (0,0000)	2.15	4.91	0.15	0.55
Spain	287	31.58%	25.80%	29.34%	21.18%	26.95%	36.70%	[29,50% ; 33,66%] (0,0000)	2.84	11.67	0.18	0.57
Sweden	1,671	70.35%	44.17%	58.36%	33.50%	48.41%	71.92%	[66,51% ; 74,19%] (0,0000)	6.11	67.61	0.80	1.14
Switzerland	542	34.36%	23.27%	29.17%	19.15%	26.27%	36.57%	[30,17% ; 38,56%] (0,0000)	14.13	257.93	0.50	1.45
Turkey	654	56.44%	49.24%	53.97%	39.58%	49.30%	67.07%	[54,56% ; 58,33%] (0,0000)	2.20	8.49	0.25	0.44
United Kingdom	2,531	39.31%	25.10%	34.54%	20.79%	30.37%	45.35%	[37,92% ; 40,71%] (0,0000)	5.03	45.55	0.36	0.91
Central and Western Europe	3,378	41.79%	29.24%	36.07%	23.38%	31.82%	45.88%	[40,06% ; 43,53%] (0,0000)	12.95	240.15	0.51	1.23
Southern Europe	1,910	42.65%	31.66%	38.33%	25.08%	35.30%	48.30%	[39,64% ; 45,66%] (0,0000)	31.60	1,200.34	0.67	1.57
Scandinavia	2,870	62.02%	36.94%	49.19%	28.93%	40.93%	61.67%	[57,50% ; 66,54%] (0,0000)	29.60	1,229.72	1.23	1.99
Britain	2,670	39.55%	25.34%	34.65%	20.97%	30.33%	45.41%	[38,17% ; 40,92%] (0,0000)	4.94	43.01	0.36	0.92
Eastern Europe	2,086	53.93%	22.44%	42.92%	27.29%	38.24%	54.35%	[45,61% ; 62,26%] (0,0000)	34.96	1,379.77	1.94	3.60
Total	12,914	47.91%	28.58%	39.70%	24.80%	34.88%	50.92%	[46,09% ; 49,74%] (0,0000)	47.17	3,079.54	1.06	2.21

Lookback Put OPM, Adjusted Lookback Put OPM and Perpetual Exchange Put OPM, 2024, Holding Period = 2 years

Country / Region	n	x̄ₐ	x̄ₕ	x̄ₜ	Q₁	Q₂	Q₃	95% (JB)	sk	kurt	sd	cv
Austria	107	32.37%	25.39%	30.05%	23.31%	27.87%	36.71%	[28,52% ; 36,22%] (0,0000)	3.80	20.11	0.20	0.62
Belgium	244	45.13%	30.05%	34.82%	23.49%	30.47%	42.44%	[34,29% ; 55,96%] (0,0000)	12.00	165.73	0.86	1.90
Bosnia and Herzegovina	22	48.94%	31.60%	40.04%	22.10%	35.51%	50.23%	[26,82% ; 71,06%] (0,0000)	3.66	15.20	0.50	1.02
Bulgaria	26	32.13%	26.13%	31.09%	21.06%	28.71%	38.82%	[26,07% ; 38,20%] (0,0060)	1.21	1.90	0.15	0.47
Croatia	27	28.31%	24.69%	27.98%	21.06%	26.17%	33.66%	[24,67% ; 31,95%] (0,0426)	0.96	0.73	0.11	0.38
Cyprus	83	62.46%	31.93%	41.47%	27.79%	37.93%	49.94%	[34,99% ; 89,92%] (0,0000)	5.81	34.47	1.26	2.01
Czech Republic	40	54.13%	25.64%	35.22%	18.14%	26.23%	42.99%	[24,89% ; 83,37%] (0,0000)	4.01	16.27	0.91	1.69
Denmark	329	77.74%	38.02%	46.89%	29.25%	40.43%	57.29%	[34,67% ; 120,82%] (0,0000)	17.43	311.31	3.97	5.11
Estonia	44	36.73%	18.90%	23.89%	14.47%	19.00%	30.01%	[12,07% ; 61,38%] (0,0000)	6.24	40.36	0.81	2.21
Finland	401	42.52%	30.84%	37.71%	26.60%	34.84%	46.59%	[38,87% ; 46,16%] (0,0000)	7.02	69.13	0.37	0.87
France	1,021	51.65%	33.96%	42.08%	26.42%	36.76%	52.27%	[46,77% ; 56,52%] (0,0000)	13.05	232.14	0.79	1.54
Germany	1,043	49.41%	36.91%	44.10%	30.49%	40.08%	55.37%	[46,67% ; 52,16%] (0,0000)	9.96	165.58	0.45	0.91
Greece	171	55.45%	30.02%	33.88%	24.88%	31.73%	40.85%	[15,52% ; 95,38%] (0,0000)	13.00	169.72	2.65	4.77
Hungary	68	47.08%	34.96%	43.36%	26.14%	36.97%	58.77%	[39,54% ; 54,62%] (0,0000)	2.01	4.64	0.31	0.66
Iceland	58	50.03%	19.32%	32.88%	20.17%	24.75%	34.76%	[19,73% ; 80,33%] (0,0000)	6.52	45.68	1.15	2.30
Ireland	139	50.13%	34.36%	41.65%	25.84%	33.33%	51.34%	[41,23% ; 59,04%] (0,0000)	4.06	18.90	0.53	1.06
Italy	593	38.58%	31.84%	36.07%	26.22%	33.53%	43.56%	[36,79% ; 40,38%] (0,0000)	4.16	28.82	0.22	0.58
Kazakhstan	26	25.10%	5.72%	24.66%	16.45%	23.65%	31.76%	[19,45% ; 30,75%] (0,1281)	0.77	1.19	0.14	0.56
Lithuania	45	27.03%	20.76%	24.20%	18.30%	23.09%	29.87%	[20,86% ; 33,20%] (0,0000)	4.36	23.79	0.21	0.76
Luxembourg	121	47.06%	38.50%	42.37%	31.89%	38.04%	50.46%	[41,35% ; 52,77%] (0,0000)	4.08	20.77	0.32	0.67
Malta	21	36.18%	32.74%	36.18%	28.14%	33.65%	39.19%	[28,35% ; 44,00%] (0,2738)	1.11	0.51	0.12	0.34
Netherlands	295	51.09%	34.18%	42.40%	26.03%	35.58%	53.56%	[42,17% ; 60,00%] (0,0000)	11.92	173.44	0.78	1.52
North Macedonia	85	41.08%	3.19%	27.13%	13.00%	22.46%	37.23%	[19,31% ; 62,85%] (0,0000)	7.97	68.81	1.01	2.46
Norway	411	61.51%	36.62%	47.36%	28.28%	39.98%	58.82%	[52,10% ; 70,92%] (0,0000)	8.29	92.23	0.97	1.58
Poland	1,204	62.54%	45.84%	56.18%	35.57%	48.42%	69.57%	[59,79% ; 65,29%] (0,0000)	4.21	32.05	0.49	0.78
Portugal	58	32.71%	26.45%	31.01%	19.97%	28.60%	39.36%	[28,12% ; 37,30%] (0,0000)	1.92	5.08	0.17	0.53
Romania	174	45.11%	34.30%	40.85%	26.34%	35.72%	50.51%	[40,13% ; 50,09%] (0,0000)	3.81	21.13	0.33	0.74
Russia	323	108.58%	44.48%	51.56%	34.78%	46.53%	62.94%	[38,53% ; 178,62%] (0,0000)	14.62	231.85	6.40	5.89
Slovenia	24	30.71%	25.28%	28.67%	19.92%	25.36%	34.84%	[23,36% ; 38,06%] (0,0000)	2.26	5.52	0.17	0.57
Spain	287	35.71%	28.95%	32.90%	23.64%	30.28%	41.06%	[33,24% ; 38,19%] (0,0000)	3.14	14.02	0.21	0.60
Sweden	1,671	82.25%	49.22%	66.58%	37.63%	53.34%	78.34%	[77,32% ; 87,18%] (0,0000)	6.43	73.95	1.03	1.25
Switzerland	542	38.71%	25.91%	32.04%	21.39%	29.44%	41.41%	[33,35% ; 44,08%] (0,0000)	15.03	282.15	0.64	1.64
Turkey	654	65.48%	55.83%	62.29%	43.93%	55.10%	78.81%	[63,10% ; 67,87%] (0,0000)	2.21	8.53	0.31	0.47
United Kingdom	2,531	44.87%	28.05%	38.68%	23.23%	33.95%	50.50%	[43,14% ; 46,60%] (0,0000)	5.48	52.74	0.44	0.99
Central and Western Europe	3,378	47.54%	32.61%	40.07%	26.16%	35.18%	50.74%	[45,32% ; 49,76%] (0,0000)	13.81	264.34	0.66	1.38
Southern Europe	1,910	49.35%	35.79%	43.73%	28.38%	39.48%	54.82%	[45,41% ; 53,29%] (0,0000)	32.62	1,254.73	0.88	1.78
Scandinavia	2,870	72.56%	41.10%	55.82%	32.18%	45.56%	68.79%	[66,63% ; 78,49%] (0,0000)	30.67	1,293.47	1.62	2.23
Britain	2,670	45.14%	28.33%	38.80%	23.52%	33.93%	50.51%	[43,44% ; 46,84%] (0,0000)	5.37	49.67	0.45	0.99
Eastern Europe	2,086	63.88%	25.72%	49.41%	31.01%	43.41%	61.70%	[52,84% ; 74,92%] (0,0000)	35.28	1,398.06	2.57	4.03
Total	12,914	55.51%	32.11%	44.81%	27.81%	39.02%	56.96%	[53,11% ; 57,91%] (0,0000)	48.67	3,219.60	1.39	2.50

Data

Perpetual Exchange Put OPM, 2024

Country / Region	n	x̄ₐ	x̄ₕ	x̄ₜ	Q₁	Q₂	Q₃	95% (JB)	sk	kurt	sd	cv
Austria	25	32.44%	28.22%	31.87%	24.51%	29.46%	37.57%	[27,23% ; 37,65%] (0,0660)	1.06	0.83	0.13	0.39
Belgium	56	33.63%	27.49%	32.71%	21.73%	30.05%	40.47%	[29,34% ; 37,93%] (0,0013)	1.13	0.78	0.16	0.48
Denmark	41	40.10%	32.03%	39.24%	28.46%	37.21%	49.21%	[34,54% ; 45,67%] (0,0543)	0.84	0.75	0.18	0.44
Finland	71	30.73%	19.48%	30.57%	20.42%	30.23%	37.85%	[27,68% ; 33,77%] (0,5277)	0.33	0.00	0.13	0.42
France	181	36.55%	29.86%	35.38%	24.12%	34.02%	45.17%	[34,16% ; 38,94%] (0,0000)	1.07	1.45	0.16	0.45
Germany	219	46.82%	39.95%	46.25%	32.78%	42.75%	58.31%	[44,43% ; 49,20%] (0,0011)	0.57	-0.44	0.18	0.38
Greece	23	31.57%	27.90%	30.53%	24.08%	30.18%	38.83%	[26,56% ; 36,58%] (0,0011)	1.23	2.87	0.12	0.37
Ireland	23	41.39%	37.39%	41.38%	33.40%	39.12%	52.41%	[35,95% ; 46,82%] (0,7079)	0.32	-0.57	0.13	0.30
Italy	45	34.63%	26.19%	33.75%	24.64%	32.02%	43.18%	[29,56% ; 39,71%] (0,0013)	1.04	1.66	0.17	0.49
Luxembourg	21	40.04%	35.48%	39.70%	32.14%	36.31%	48.27%	[33,93% ; 46,14%] (0,6574)	0.39	-0.60	0.13	0.33
Netherlands	59	37.92%	31.96%	37.11%	26.39%	32.67%	44.65%	[33,62% ; 42,23%] (0,0069)	0.99	0.32	0.17	0.44
Norway	55	32.10%	25.45%	31.68%	20.48%	28.78%	41.10%	[28,23% ; 35,97%] (0,1993)	0.59	-0.18	0.14	0.45
Spain	47	32.45%	27.43%	31.57%	21.46%	29.01%	42.36%	[28,38% ; 36,51%] (0,0293)	0.93	0.41	0.14	0.43
Sweden	221	46.21%	38.69%	45.79%	33.77%	44.45%	56.46%	[43,92% ; 48,51%] (0,0503)	0.40	-0.12	0.17	0.38
Switzerland	138	35.49%	28.17%	34.89%	23.69%	32.53%	44.09%	[32,81% ; 38,17%] (0,0096)	0.63	-0.08	0.16	0.45
Turkey	12	51.99%	44.58%	51.99%	36.31%	44.46%	71.52%	[38,98% ; 65,00%] (0,5014)	0.51	-1.31	0.20	0.39
United Kingdom	445	33.47%	24.93%	32.69%	21.94%	31.00%	43.03%	[32,01% ; 34,94%] (0,0000)	0.77	0.46	0.16	0.47
Central and Western Europe	700	39.37%	32.04%	38.44%	26.91%	35.56%	48.45%	[38,08% ; 40,66%] (0,0000)	0.79	0.19	0.17	0.44
Southern Europe	144	34.73%	28.00%	33.57%	22.63%	32.04%	42.36%	[32,07% ; 37,40%] (0,0000)	1.11	1.28	0.16	0.47
Scandinavia	396	40.43%	28.54%	39.84%	28.04%	39.10%	50.17%	[38,69% ; 42,17%] (0,0001)	0.52	0.09	0.18	0.44
Britain	468	33.86%	25.34%	33.13%	22.28%	32.11%	43.33%	[32,44% ; 35,29%] (0,0000)	0.72	0.37	0.16	0.46
Eastern Europe	12	25.87%	20.46%	25.87%	17.29%	19.95%	25.96%	[14,92% ; 36,83%] (0,0000)	2.68	7.89	0.17	0.67
Total	1,720	37.63%	28.71%	36.77%	24.82%	34.79%	47.35%	[36,82% ; 38,45%] (0,0000)	0.75	0.29	0.17	0.46

News from EACVA

EACVA's 18th International Business Valuation Conference 2025

Share knowledge ✓ Build networks ✓ Ensure quality ✓

EACVA is pleased to invite business valuation professionals across Europe and beyond to its 18th International Business Valuation Conference, held on **13 and 14 November 2025** in **Munich**. As the premier networking event of the year, our conference offers unparalleled opportunities for knowledge exchange, professional development, and relationship building within the valuation community.

Attendees can look forward to two inspiring days featuring sessions led by renowned experts such as Prof. Dr. Hans-Werner Sinn, Thorsten Jekel, Antonella Puca, Prof. dr. Marc Goedhart, Seth Bernström and many others. The program includes 2 keynotes and 24 parallel sessions that address the latest trends, challenges, and innovations in business valuation.

A highlight will be the exclusive Networking Dinner at the traditional Augustiner Bräu Festsaal, fostering informal contacts and collaborations with a high-caliber international audience. Spaces are limited – <u>secure your place today!</u>

Mastering the Valuation of Asset-Light, IP-Driven Start-Ups

Live Web Seminar | Thursday, 20 November 2025 | 13:30–16:15 CET

In a rapidly evolving business world where intellectual property, data, and digital scalability drive enterprise value – yet traditional valuation models often fail to capture these dynamics. Are you equipped to value the highly asset-light start-ups? EACVA's exclusive live online seminar "**Valuation of Highly Asset-Light Start-Up Companies**" provides business valuation professionals, analysts and corporate finance experts with the necessary tools and insights to confidently tackle this challenge.

Led by renowned valuation expert **Prof. Dr. Matthias Meitner, CFA**, this seminar will provide you with a hands-on approach to:
- Distilling valuation-relevant information from data overload
- Identifying competitive advantages in digital business models
- Developing robust forecasting approaches for intangible-heavy firms
- Turning complex analyses into actionable valuations

This seminar is designed for valuation professionals, analysts, auditors, tax advisors, controllers, consultants, investors, and corporate finance experts. Reserve your spot today and gain expert insights into valuing the business models shaping tomorrow's markets: <u>Learn more and register now!</u>

News

News from IVSC

IVSC Europe Committee Launches Valuation Survey

The IVSC Europe Committee has launched a wide-ranging survey to gather views from across the valuation ecosystem. The survey aims to capture perspectives from valuers, VPOs, regulators, end users, and academics on the challenges and opportunities shaping the profession in Europe. Insights will help guide the Committee's priorities and inform IVSC's global technical agenda.

Take the survey

Perspectives Paper: AI and Technology in Valuation

IVSC has published a new **Perspectives Paper** examining the role of artificial intelligence and technology in valuation. The paper explores how AI is reshaping practice — from data gathering and modelling to reporting and quality control — while emphasising the continuing importance of professional judgement and accountability. Stakeholder feedback is invited to inform future standard-setting.

Download the Perspectives Paper

Perspectives Paper: Managing Risk in Valuation

A further **Perspectives Paper** focuses on managing risk in valuation. It highlights how valuers can identify, assess, and communicate risks that influence their work, drawing on case studies and practical examples. The paper emphasises the need for transparent disclosure and robust processes, supporting greater trust and consistency in valuation outcomes.

Download the Perspectives Paper
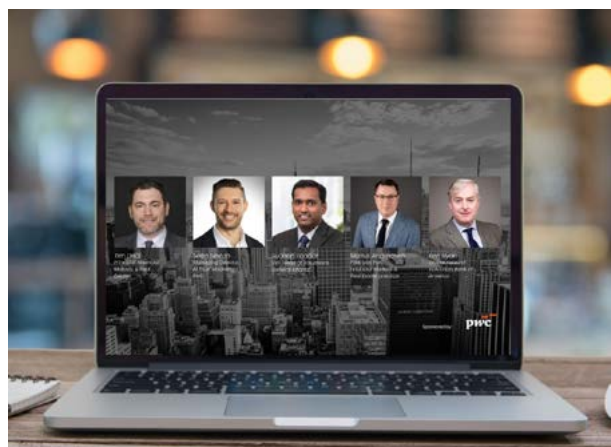
Technology and Valuation – Watch the Latest Dialogue Webinar

The latest session in the **IVSC Valuation Dialogue Series**, hosted with PwC, explored how technology is transforming valuation across asset classes. Panellists shared perspectives on AI, data quality, and tools shaping valuation practice today and in the future. The series continues to provide an open forum for sharing global insights on emerging themes in valuation.

Watch the recording

Shaping the Future of Valuation – IVS Agenda Consultation

Following a global consultation, IVSC has published a summary of feedback received on its technical priorities. The **Agenda Consultation 2024-25** attracted responses from stakeholders worldwide, helping the Standards Boards to identify areas requiring further guidance, research, or new standards. Insights will inform the technical work plan in the lead-up to the next edition of IVS.

Read the summary

Standards Boards Meet in Prague

In June 2025, IVSC's Standards Boards convened in **Prague** for a week of technical discussions aligned with Czech Valuation Day. Board members from more than 40 countries advanced their work on feedback to the Agenda Consultation and began preparations for the next cycle of standard-setting. Meetings also included outreach with local stakeholders, including the Czech Ministry of Finance and Czech National Bank.

Register Now: IVSC–Kroll Valuation Webinar Series 2025

Now in its sixth year, the **IVSC–Kroll Valuation Webinar Series** is a leading platform for global dialogue on valuation. Taking place from **4–13 November 2025**, the series will convene leading experts and more than 3,000 participants worldwide. This year's five sessions will cover the global economic outlook, VPO-led guidance on intangibles, the evolving role of technology and AI, valuation in legal disputes, and divergence in public versus private real estate markets. Each session is free to attend, but places are limited.

Find out more and register here

EBVM

The European Business Valuation Magazine