Robo Advisory Customer Groups: Who Requires Advice?*

Justus Blaschke and Johannes Kriebel

Abstract: Prior literature has often investigated how robo advisors can broaden their customer base. This study is based on the observation that some customers value the risk elicitation of robo advisors (guidance customers), whereas others value other aspects such as the simplicity and convenience of these services. Based on empirical robo advisory data, we build machine learning models to identify guidance customers. The models make predictions based on the financial knowledge of customers to a large extent. The age of a customer, the amount invested, income, and available assets are further important determinants.

Keywords: Robo Advisory, Robo Advice, Financial Decision Support, Financial Advice, Customer Segmentation, Machine Learning, Financial Literacy

Robo Advisory Kundengruppen: Wer benötigt Beratung?

Zusammenfassung: Die bestehende Literatur im Bereich Robo Advisory hat oft die Frage adressiert, wie Robo Advisor ihre Kundenbasis erweitern können. Unser Forschungspapier basiert auf der Beobachtung, dass einige Kunden die Einordnung der Risikobereitschaft (Beratungskunden) suchen, während andere Kunden andere Aspekte wie den einfachen und komfortablen Zugang zu Geldanlagen suchen. Wir bauen Modelle des maschinellen Lernens, um diese Beratungskunden basierend auf Robo Advisory-Daten zu identifizieren. Die Vorhersage stützt sich gerade auf die finanzielle Vorbildung. Das Kundenalter, der investierte Betrag, das Einkommen und das verfügbare Vermögen sind weitere wichtige Determinanten.

Stichwörter: Robo Advisory, Robo Advice, Finanzielle Entscheidungshilfen, Finanzberatung, Kundensegmentierung, Maschinelles Lernen, Finanzielle Bildung

1 Introduction

As defined by *Sironi* (2016, p. 25) robo advisors are automated investment solutions that use digital technologies to guide customers in their investment decisions. Robo advisors mainly use rebalancing techniques based on passive investments and diversification as their asset management strategy. The advisory process usually starts by assessing a person's aims, risk tolerance, and risk appetite either online or via a smartphone app.

* The authors wish to thank Anselm Hüwe, Andreas Pfingsten (the editor), and an anonymous reviewer for their comments, suggestions, and valuable feedback.

Based on this information, the robo advisor suggests an investment portfolio, in which the customer can then directly invest. The investment portfolio is periodically rebalanced to remain congruent with customer needs.

Sironi (2016, p. 24) calls robo advisory a game-changer in personal finances. There are, indeed, good arguments that robo advisors benefit customers in many important ways. *Ruf et al.* (2015) point out that robo advisors are particularly transparent in the advisory and investment process. In addition, the standardized process leads to a constant level of quality (*Faloon/Scherer* 2017). *Beketov et al.* (2018) emphasize that robo advisors utilize efficient processes and that the decreased costs are forwarded to consumers. Furthermore, robo advisory services are ubiquitous and available at any time over multiple channels. Robo advisors provide a service similar to wealth management services but without common entry barriers such as high minimum investments. In addition, *D'Acunto et al.* (2019) point out that robo advisors help reduce consumers' exposure to common behavioral biases such as trend-chasing. In summary, robo advisory is often seen as a means to provide better and more affordable advisory services, particularly to customers with low income and financial education.

Despite these potential benefits to individual investors, it is interesting that the dissemination of robo advisory services is proceeding more slowly than many observers expected (*Bruckes et al.* 2019). There is, hence, literature that investigates the factors that hinder customers' use of robo advisory services. For example, *Jung et al.* (2018) point out that customers tend to stick to services they are familiar with. *Bruckes et al.* (2019) determine trust in technology and institutions as well as the perceived risk of technology to be important factors. *Woodyard/Grable* (2018) argue that robo advisors are primarily used by young customers. *Hohenberger et al.* (2019) highlight the importance of experience.

One important factor for higher acceptance and better-quality robo advisory services is arguably the design of robo advisory questionnaires, which is discussed by *Faloon/Scherer* (2017), *Jung et al.* (2018), and *Tertilt/Scholz* (2018) In practical operations, the number of respondents tends to decrease with an increase in the number of questions asked by robo advisors. *Jung et al.* (2018) discuss ways of designing questionnaires that customers will complete. *Tertilt/Scholz* (2018) assert that there is a trade-off between extensive risk assessments through long questionnaires and an increase in business through short questionnaires. This trade-off is arguably a reason that robo advisory questionnaires tend to be shorter compared to those used in offline advisory sessions. In our study, we use data from a German robo advisory service that offers customers the choice between an extensive and a short questionnaire. As argued by *Tertilt/Scholz* (2018), many customers prefer short questionnaires. However, there is one customer group that chooses the long questionnaire option, which provides a more extensive risk assessment. Based on this observation, we argue that this group of customers particularly seeks the risk elicitation aspect of robo advisory services, and we label this group *guidance customers*. Although *Tertilt/Scholz* (2018) argue in favor of a general tendency to uniformly use short questionnaires, robo advisors could identify guidance customers and offer them customized services. This could increase value for this group by better adapting services to their needs.

Our study thus investigates how to distinguish guidance customers from other customers. We build a random forest model to identify guidance customers and benchmark it with a logistic regression model as a comparison. *Siroky et al.* (2009) highlight the strength of ensemble methods in general and the random forest in classification problems.

The random forest model is found to outperform the logistic regression in our analysis. The model identifies guidance customers with an area under the curve (AUC) of 0.66, which is a strong performance especially compared to a benchmark using a logistic regression of 0.63.

We further assess the characteristics of guidance customers using partial prediction plots and permutation importance. This allows determining the variables that mainly drive predictions. Based on these results, customers with more financial education require less advice. This is also the case for customers with higher initial assets and income. In contrast, customers with lower financial education, income, and assets are more likely to require advice. The results for customer age are interesting, as older customers require less assistance for risk elicitation. Overall, a robo advisory service could offer lower-income customers and customers with less financial education more extensive advice and provide simpler services with fewer risk elicitation options to other customers.

The remainder of the paper is structured as follows: Section 2 presents our methods. Section 3 describes the data. Section 4 offers the results. Section 5 concludes.

2 Methodology

2.1 Machine Learning Methods

2.1.1 Classification and Regression Trees

In this section, the random forest, a common machine learning method, is introduced. It is later used to predict the individual demand for advice. The random forest is based on the concept of recursive partitioning. It was made popular in the 1980s by *Breiman et al.* (1984) and remains a prevalent machine learning tool for many prediction tasks. The most common algorithm from this group is the classification and regression trees (CART) method suggested in the original publication of *Breiman et al.* (1984).

Recursive partitioning is a non-parametric approach for learning structures in a dataset that uses them to make predictions for new data points. During the learning step, a training set is used to build a binary tree structure partitioning the dataset into subsets. Every leaf node accounts for one subset of the training data and the inner nodes of the tree account for one decision rule. In later prediction steps, new observations are categorized using these decision rules. Every new observation walks down the tree according to those decision rules. The predicted value is then calculated from the observations in the final leaf node. For regression problems, the average of the response values of the observations is taken as the prediction. For categorization problems, the majority class in the leaf nodes is taken as the prediction.

A simple example of such a decision tree is given in *Figure 1*. The tree is built on the robo advisory dataset and uses *knowledge* and the clients' *net income* as predictors. There are three splits produced during the learning step. The first is a split of *knowledg*e. All observations with *knowledge* lower than or equal to three are sent to the left of the tree and arrive at a terminal node to accept the advice. All other observations are sent to the right of the tree. These are further split by *net income*. All observations with an outstanding amount higher than or equal to 60,000 euro go to the right of the subtree and arrive at a final node this way. In the last node, the remaining observations are split up into net incomes higher than or equal to 45,000 euro and net incomes lower than 45,000 euro.

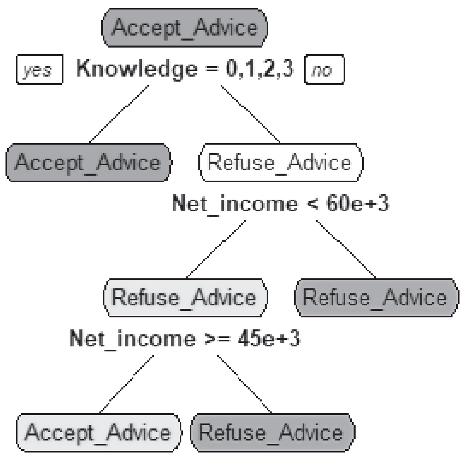

Figure 1: Classification tree for whether advice is accepted or refused

2.1.2 Random Forest

When using the CART method, several drawbacks are often mentioned in the literature. These include a tendency to overfit data, unstable predictions through small leaf nodes, and a poor approximation of functional relations with the discrete prediction values in the leaf nodes (see *Siroky et al.* 2009 or *Berk* 2020, pp. 198–199). To cope with these shortcomings and to improve the predictive power, ensemble methods aggregating a group of decision trees instead of using a single tree have been suggested. More generally, ensemble methods combine a set of base learners. The prediction of these base learners is then later aggregated. This is usually done by taking a majority vote in categorization problems or an average of the predictions for regression problems. In the case of the random forest, the ensemble is a set of decision trees.

Breiman (1996) introduced a method called bagging (bootstrap aggregation). The idea behind bagging is to produce a set of decision trees by repeatedly sampling from a dataset and building a decision tree for each bootstrap sample. The advantages of bagging mainly lie in the reduction of the variance of predictions by averaging the outcomes of the ensemble of trees and reducing bias by having a larger variety of possible predictions using the averages of the predictions of the single trees. Later, *Breiman* (2001) combined the idea of bagging with methods from other authors, such as the random split selection method by *Dietterich* (2000), and named the new method the random forest. *Breiman* (2001) defined the random forest as an ensemble of decision trees in which K decision trees are built depending on identical and independently distributed random vectors $\Theta_0, \Theta_1, \ldots, \Theta_K$. Each tree is built using a random sample from the dataset and in every node of the tree, a selection of m explaining variables from all available explaining variables M is drawn at random. The random vector Θ_k for the k-th tree determines the subsample from the training dataset that is used for the tree k and the randomly selected features that are available for the split selection in each node.

The tuning parameters that are usually used in the random forest method are the number of trees K and the number of features M used for the split selection in each node. Concerning the number of trees K, one could build as many trees as possible, as there is no overfitting with an increasing number of trees (*Breiman* 2001). Concerning the number of split features M, often, the number of $M = \sqrt{p}$ in regression or $M = \frac{p}{3}$ in classification is used as the rule-of-thumb (*Siroky et al.* 2009), but it can also be determined using a grid-based search and cross-validation.

The random forest model is set up as follows: In line with prior applications of machine learning methods, the dataset is split into a training set on which the random forest model is set up and trained and a test set on which the trained model performance is evaluated. We use a split-ratio of 70 % of the data for the training set and 30 % for the test set.

We further apply a 10-fold cross-validation method to the training data to select the optimal number of the random forest split features (mtry), resulting in mtry = 5, and split features between 2 and 6 were tested. The mtry is chosen to increase the AUC in the cross-validation. Cross-validation splits the training set into 10 folds and iteratively builds random forest models with 9 folds for training and 1 fold for testing. This procedure is repeated until every fold serves at least one time as a test set. Cross-validation ensures less biased results and increases the model's out-of-sample performance.

2.2 Machine Learning Interpretability

Although the methods presented here can offer substantial advantages in terms of prediction quality, they come at the cost of being more difficult to analyze than the usual output of linear models. This is less of a problem when a CART has few nodes, such as in *Figure 1*. In this case, the direction and importance of variables can be read by looking at the individual splits. This becomes increasingly difficult the more complex the tree becomes and virtually impossible when evaluating a whole forest of trees. In this sense, the prediction of a random forest could be considered a result of a "black box" (*Palczewska et al.* 2013). However, although researchers and practitioners want to benefit from high prediction quality, they need to understand how these predictions are derived.

There are several approaches to increasing the interpretability of predictions. We employ two of these approaches in this study. The first is based on plotting the partial variation in predictions over the values of predictive variables, as suggested by *Frank/Bouckaert* (2009, pp. 65–66). In such a plot, one could see how much the prediction varies with the predictive variable. One could further evaluate the monotony of the relationship between the predictive variable and the outcome.

Several measures of variable importance can also be calculated. In this study, the variable permutation importance of one variable is calculated by permuting its values in the test data. Next, performance metrics for predictions on the dataset using the permuted variable and then using the unpermuted variable are calculated. The importance measure is taken either as the change in overall categorization accuracy for classification problems or as the change in overall squared errors in regression problems (*Gregorutti et al.* 2017). These variable importance measures are computed for all variables and their values can be compared. It is assumed that variables with a large difference between the permuted and the non-permuted data have a higher influence on the model than variables with a lower permuted variable importance.

3 Data

3.1 Dataset

Our empirical analysis is based on a dataset of a leading German robo advisor that offers digital financial investment services through one-time investments and savings plans based on exchange-traded fund portfolios. The data includes the socio-economic client information and responses to the robo advisory questionnaire that every client must complete. The questionnaire marks the start of the advisory process and its outcome is the suggested asset allocation by the robo advisor. The questionnaire contains items regarding the customer's investment goals, risk tolerance, and risk capacity.

The questionnaire used in this study has a major characteristic that allows us to draw conclusions about the demand for advice and the identification of customers who value risk elicitation in the investment process. In the course of the onboarding process, every client must choose between self-determining the level of risk tolerance or answering additional questions to determine his or her risk tolerance as assessed by the robo advisor.

The total sample contained information of 10,447 subjects and was collected during a sample period from the second half of 2019 to the first half of 2020. In a first step, we limited the sample to clients who legitimized themselves and opened an account with the robo advisor. From the remaining 6,432 accounts, 1,224 accounts opened for children and joined accounts were excluded to make sure that the cases are comparable. After excluding 32 data sets with missing values, the sample contained 5,176 individual observations.

Table 1 presents summary statistics for the sample. These allow some interesting observations about the characteristics of investors that use the advisory service. The individuals appear to have relatively high incomes. The average client belongs to the wealthiest 10 % of the German population with a net income of roughly 46,000 euro, according to a recent study on the German income distribution by the Institut der deutschen Wirtschaft (*Niehues/Stockhausen* 2020). In addition, only about 12 % of the clients state that they have no or little knowledge in financial matters, leaving the majority with at least fundamental knowledge, as confirmed by *Bucher-Koenen/Lusardi* (2011), who found the average level of financial literacy in a nationwide survey to be rather moderate. The mean investor age of 44.3 is not particularly young but compared to the mean age of 62 of typical wealth management clientele, as suggested by *McKinsey*'s (2015) study, is below average, supporting the assumptions of *Woodyard/Grable* (2018) and *Fulk et al.* (2018). However, all age groups are well represented and the mean age of customers (44.3 years) is close to the averages of the German (44.5 years) and the European (43.7 years) populations (*Eurostat* 2020; *Statistisches Bundesamt* 2020). Regarding the wealth and financial literacy of the average customer, the sample appears to be shifted from the center of German society. However, this is a typical characteristic of robo advisory customers. *Fan/Chatterjee* (2020) found that the adoption of robo advisory services in the United States is also positively influenced by high subjective knowledge in financial matters and the accumulated wealth of individuals, which *Merkle* (2020) found to be true when surveying German adults on the same topic.

Table 1: Summary statistics

	N	Mean	Median	St. Dev.
Gender	5176			
Female = 0	1516			
Male = 1	3660			
Age	5176	44.29	43	14.38
< 30 years old	993			
31–40 years old	1366			
41–50 years old	1050			
51–60 years old	994			
> 60 years old	773			
Knowledge	5176	2.76	3	1.18
None = 0	120			
Little = 1	511			
Fundamental = 2	1645			
Good = 3	1580			
Very Good = 4	865			
Professional = 5	455			
Net Income	5176	45,765.93	36,000	40,473.6
< 30,000 euro	1886			
30,001–60,000 euro	2485			
60,001–80,000 euro	308			
> 80,000 euro	497			
Total Assets	5176	288,480.3	110,000	569,024.9
< 50,000 euro	1861			
50,001–150,000 euro	1059			
150,001–400,000 euro	1188			
> 400,000 euro	1068			
Invested Amount	5176	12,589.89	5,000	30,347.4
< 1,000 euro	1135			
1,001–5,000 euro	1457			
5,001–10,000 euro	1589			
10,001–30,000 euro	600			
> 80,000 euro	395			

Notes: To guarantee absolute anonymity for clients, the *Age* variable was randomized by a random subtraction or addition of 5 years per customer. Ordinal variables were assumed to be continuous to calculate summary statistics.

3.2 Derivation of Dependent and Independent Variables

Our identification of the guidance customers in this study is quite straightforward. As the customers are asked to choose between self-assessing their risk level or using the risk elicitation questions during the onboarding process, this allows for direct identification.[1] We, therefore, use the customer decision to utilize the offered risk elicitation questions as a measure of demand for advice. *Figure 2* shows that only 12 % of customers choose the risk elicitation questions, whereas 88 % opt out of answering these questions.

As most of the customers reject advice, leaving only a relatively small fraction of customers who seek guidance, we use rebalancing. This is done to draw better conclusions about the guidance customer group. We applied a balancing technique called random oversampling examples (ROSE) to the training set. The ROSE algorithm is theoretically founded and applies smoothed bootstrap resampling to draw artificial samples around the minority class. ROSE is found to outperform classical over- and undersampling methods while enabling unbiased results (*Menardi/Torelli* 2014).

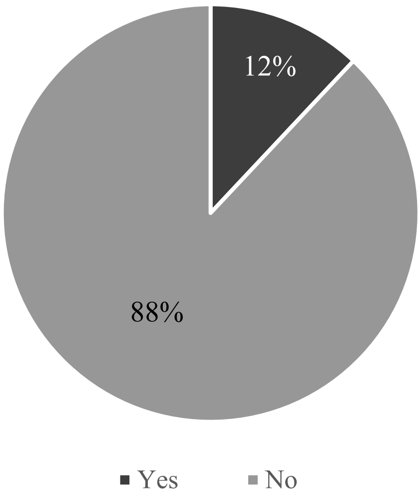

Figure 2: Share of customers that choose advice in the risk elicitation process

The independent variables were selected based on prior research. The first included variable is financial knowledge (*Joo/Grable* 2001; *Hackethal et al.* 2012; *Robb et al.* 2012; *Calcagno/Monticone* 2015). From one perspective, the more financially literate an individual is, the more likely the individual will delegate his or her financial means to a professional. From a different perspective, given that customers realize the need for financial advice, less financially literate customers could seek more extensive advice. Seeking advice also tends to be positively related to an individual's wealth and income (*Hackethal et al.* 2012; *Gibson et al.* 2013). *Hackethal et al.* (2012) suggest that there is a positive relationship between the demand for financial advice and an individual's age and indicate that older investors are more likely to accept advice. Other studies, however, find no or

1 The default option is the choice between different risk levels. In addition, a button allows to open the questionnaire determining the risk level based on additional questions.

even a slightly negative interrelation of age and the need for financial advice. Women are in general more likely to delegate their financial matters than men (*Gentile/Soccorso* 2016). This might be caused by men being more likely to be overconfident in their own abilities (*Barber/Odean* 2001). The invested amount is not investigated to a great extent, but we do expect a non-negligible influence on the decision-making and, therefore, include it as an independent variable.

4 Results

4.1 Model Performance

The receiver operating characteristics (ROC) curve for the random forest model is displayed as the grey line in *Figure 3*. The figure presents the true positive rate plotted against the false positive rate for different decision thresholds of the classifiers. In terms of performance, the random forest model, dealing with the demand for advice in the risk tolerance assessment, yields an AUC of 0.66. Compared to other fields such as credit risk prediction, this is a decent performance.

To benchmark the performance of the random forest model, we set up a logistic regression fitted to the same data as the random forest model. The logistic regression is also fitted on the training set and subsequently used for predictions on the test set. The test set AUC of the logistic regression is 0.63, indicating the validity of the results and relative superiority of the random forest model. The respective ROC curve is displayed as the black line in *Figure 3*.

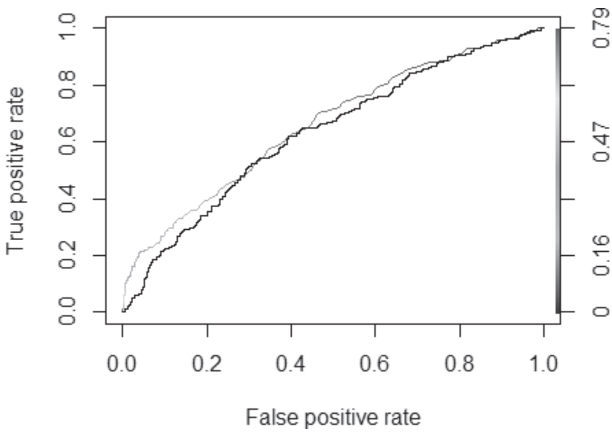

Figure 3: ROC curves of the random forest model (grey) and logistic regression model (black)

4.2 Model Visualizations

Figure 4 presents relationships between predicted outcomes and explanatory variables as a conditional density plot, as explained in Section 2.2. The figure shows partial predictions of the random forest model for whether risk elicitation is required. The predictions are plotted over independent variable values. The panels show that robo advisory clients' need for advice is negatively related to their *knowledge* of financial matters. The same

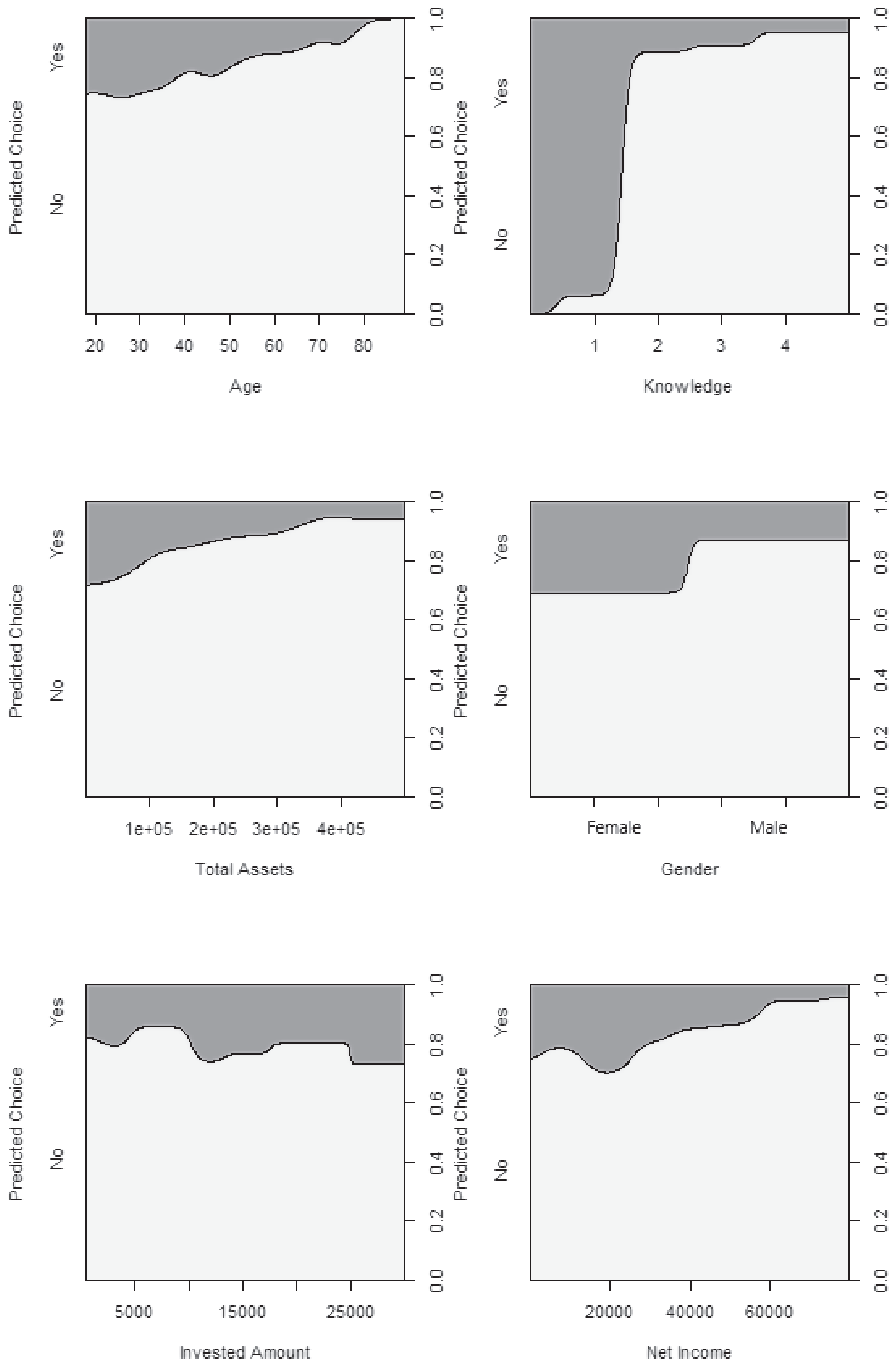

Figure 4: Partial predicted choices for risk elicitation in the risk tolerance assessment process. The graph plots the probabilities of requiring risk elicitation as the dark grey areas in each panel.

conclusion can be inferred when examining the relationships to *net income* and the *total assets* to the demand for advice, even if not to the same strong extent (see the Appendix for the plot for the permutation variable importance, which shows qualitatively similar results).

Older clients are additionally shown to be more confident and avoid seeking advice, which is contrary to the suggestions of *Hackethal et al.* (2012). The opposite is observable when examining the forecasted probability for the *invested amount* variable. The higher the sum a client is willing to spend, the more likely he or she is to demand advice. The *gender* variable, in line with prior research from *Gentile/Soccorso* (2016) and others, suggests that women have a higher tendency to seek assistance while assessing their risk tolerance compared to men. This might be related to the research findings of *Barber/Odean* (2001) that men are more overconfident than women.

The predictions of the random forest model thus suggest that male, financially literate, older, and wealthier clients are more confident in their assessment and, hence, less likely to require risk elicitation. However, these results might not contradict prior research findings in which knowledge and wealth are found to be positively associated with seeking financial advice, as the usage of the robo advisor can be interpreted as seeking advice. Additionally, as stated in Section 3.1, the robo advisory service is mostly used by higher-income clients with above-average education in financial matters. The diverging direction of influence of the *invested amount* with *total assets* and *net income* is particularly interesting. Clients become increasingly more likely to accept advice with higher sums to be invested, whereas it appears to be reversed in relation to their total financial assets. The acceptance of advice seems, therefore, to increase with higher stakes.

5 Conclusion

Our study is based on *Tertilt/Scholz*'s (2018) argument that robo advisors often use short questionnaires to decrease the chance of customers quitting the onboarding process. We study this issue based on a dataset in which customers could choose between a long and a short questionnaire. Although many customers choose the short questionnaire, there is one group of customers that deliberately chooses the long questionnaire and, in this way, a more detailed risk elicitation. This has theoretical implications for better understanding the issue discussed by *Tertilt/Scholz* (2018). Furthermore, from a practical perspective, this suggests that robo advisors could learn about the level of advice customers need and adapt questionnaires accordingly. Based on our results, this learning step can be well-conducted using a random forest model.

From a more general perspective, the nature of our data allows us to identify the features that are characteristic of the guidance customer group. We identify customers with low financial education, assets, and income as those that value more extensive advice. Furthermore, young and female customers prefer more extensive risk elicitation. The demand for risk elicitation further increases with the amount invested. This identifies variables that robo advisor providers should focus on when adapting questionnaires.

The data set used in this study could have the limitation to be specific for the customers of one robo advisor. The customers are further limited to customers on the German market. This might affect whether the results apply to customers in other markets. By comparing the socio-demographic customer characteristics to those stated in prior research, we show that the clients among the studies are comparable. However, one could think of the possibility that the results do not apply to customers in other geographic regions. This could be the focus of further work. Future research could also assess other factors related to the demand for customer advice.

In summary, our results are helpful in better understanding robo advisory customers and improving the design of robo advisory services to meet customer expectations and increase customer value.

References

Barber, B. M.; Odean, T. (2001): Boys will be boys: Gender, overconfidence, and common stock investment. In *Quarterly Journal of Economics* 116 (1), pp. 261–292.

Beketov, M.; Lehmann, K.; Wittke, M. (2018): Robo advisors: quantitative methods inside the robots. In *Journal of Asset Management* 19 (6), pp. 363–370.

Berk, R. A. (2020): Statistical learning from a regression perspective. 3rd ed. 2020. Cham: Springer International Publishing; Imprint: Springer (Springer Texts in Statistics).

Breiman, L. (1996): Bagging predictors. In *Machine Learning* 24 (2), pp. 123–140.

Breiman, L. (2001): Random forests. In *Machine Learning* 45 (1), pp. 5–32.

Breiman, L.; Friedman, J.; Stone, C. J.; Olshen, R. A. (1984): Classification and regression trees. Boca Raton: CRC Press.

Bruckes, M.; Westmattelmann, D.; Oldeweme, A.; Schewe, G. (2019): Determinants and barriers of adopting robo-advisory services. In *Proceedings of the Fortieth International Conference on Information Systems*, pp. 1–10.

Bucher-Koenen, T.; Lusardi, A. (2011): Financial literacy and retirement planning in Germany. In *Journal of Pension Economics and Finance* 10 (4), pp. 565–584.

Calcagno, R.; Monticone, C. (2015): Financial literacy and the demand for financial advice. In *Journal of Banking & Finance* 50, pp. 363–380.

D'Acunto, F.; Prabhala, N.; Rossi, A. G. (2019): The promises and pitfalls of robo-advising. In *Review of Financial Studies* 32 (5), pp. 1983–2020.

Dietterich, T. G. (2000): An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting, and randomization. In *Machine Learning* 40 (2), pp. 139–157.

Eurostat (2020): Bevölkerung: Strukturindikation. http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_pjanind&lang=de, checked on 12/29/2020.

Faloon, M.; Scherer, B. (2017): Individualization of robo-advice. In *Journal of Wealth Management* 20 (1), pp. 30–36.

Fan, L.; Chatterjee, S. (2020): The utilization of robo-advisors by individual investors: An analysis using diffusion of innovation and information search frameworks. In *Journal of Financial Counseling and Planning* 31 (1), pp. 130–145.

Frank, E.; Bouckaert, R. R. (2009): Conditional density estimation with class probability estimators. *Zhi-Hua Zhou, Takahi Washio (Eds.)*: Advances in machine learning. Berlin: Springer, pp. 65–81.

Fulk, M.; Grable, J. E.; Watkins, K.; Kruger, M. (2018): Who uses robo-advisory services, and who does not? In *Financial Services Review* 27, pp. 173–188.

Gentile, M.; Soccorso, P. (2016): Financial advice seeking, financial knowledge and overconfidence. Evidence from the Italian market. In *CONSOB Discussion Papers* (83), pp. 1–45.

Gibson, R.; Michaylik, D.; van de Venter, G. (2013): Financial risk tolerance: An analysis of unexplored factors. In *Financial Services Review* 22, pp. 23–50.

Gregorutti, B.; Michel, B.; Saint-Pierre, P. (2017): Correlation and variable importance in random forests. In *Stat Comput* 27 (3), pp. 659–678.

Hackethal, A.; Haliassos, M.; Jappelli, T. (2012): Financial advisors: A case of babysitters? In *Journal of Banking & Finance* 36 (2), pp. 509–524.

Hohenberger, C.; Lee, C.; Coughlin, J. F. (2019): Acceptance of robo-advisors: Effects of financial experience, affective reactions, and self-enhancement motives. In *Financial Planning Review* 2 (2), pp. 1–14.

Joo, S.; Grable, J. E. (2001): Factors associated with seeking and using professional retirement-planning help. In *Family and Consumer Sciences Research Journal* 30 (1), pp. 37–63.

Jung, D.; Erdfelder, E.; Glaser, F. (2018): Nudged to win: Designing robo-advisory to overcome decision inertia. In *Proceedings of the 26th European Conference on Information Systems (ECIS 2018)*. https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1018&context=ecis2018_rip, checked on 2/18/2020.

McKinsey (2015): The fountain of growth – demographics and wealth management. https://www.mckinsey.com/~/media/McKinsey/Industries/Financial%20Services/PriceMetrix/Our%20Insights/The%20Fountain%20of%20Growth/The-Fountain-of-Growth-Screen-Final.pdf, checked on 1/4/2021.

Menardi, G.; Torelli, N. (2014): Training and assessing classification rules with imbalanced data. In *Data Mining and Knowledge Discovery* 28 (1), pp. 92–122.

Merkle, C. (2020): Robo-advice and the future of delegated investment. In *Journal of Financial Transformation* 51, pp. 20–27.

Niehues, J.; Stockhausen, M. (2020): Einkommensverteilung in Deutschland: Wer zur Oberschicht gehört. https://www.iwkoeln.de/presse/pressemitteilungen/beitrag/judith-niehues-maximilian-stockhausen-wer-zur-oberschicht-gehoert.html, checked on 8/17/2020.

Palczewska, A.; Palczewski, J.; Robinson, R. Marchese; Neagu, D. (2013): Interpreting random forest models using a feature contribution method. In *2013 IEEE 14th International Conference Proceedings*, pp. 112–119.

Robb, C.; Babiarz, P.; Woodyard, A. S. (2012): The demand for financial professionals' advice: The role of financial knowledge, satisfaction, and confidence. In *Financial Services Review* 21 (4), pp. 291–305.

Ruf, C.; Back, A.; Bergmann, R.; Schlegel, M. (2015): Elicitation of requirements for the design of mobile financial advisory services – Instantiation and validation of the requirement data model with a multi-method approach. In *48th Hawaii International Conference on System Sciences*.

Siroky, D. S.; Banks, D.; Bustikova, L.; Lunagomez, S.; Wegman, E. (2009): Navigating random forests and related advances in algorithmic modeling. In *Statistics Surveys* 3, pp. 147–163.

Sironi, P. (2016): FinTech innovation: From robo-advisors to goal based investing and gamification. Chichester, West Sussex: Wiley (Wiley finance series). http://onlinelibrary.wiley.com/book/10.1002/9781119227205.

Statistisches Bundesamt (2020): Durchschnittsalter auf Grundlage des Zensus 2011 nach Geschlecht und Staatsangehörigkeit. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bevoelkerung/Bevoelkerungsstand/Tabellen/durchschnittsalter-zensus-jahre.html, checked on 12/29/2020.

Tertilt, M.; Scholz, P. (2018): To advise, or not to advise – How robo-advisors evaluate the risk preferences of private investors. In *Journal of Wealth Management* 21 (2), pp. 70–84.

Woodyard, A. S.; Grable, J. E. (2018): Insights into the users of robo-advisory firms. In *Journal of Financial Service Professionals* 5, pp. 56–66.

Appendix

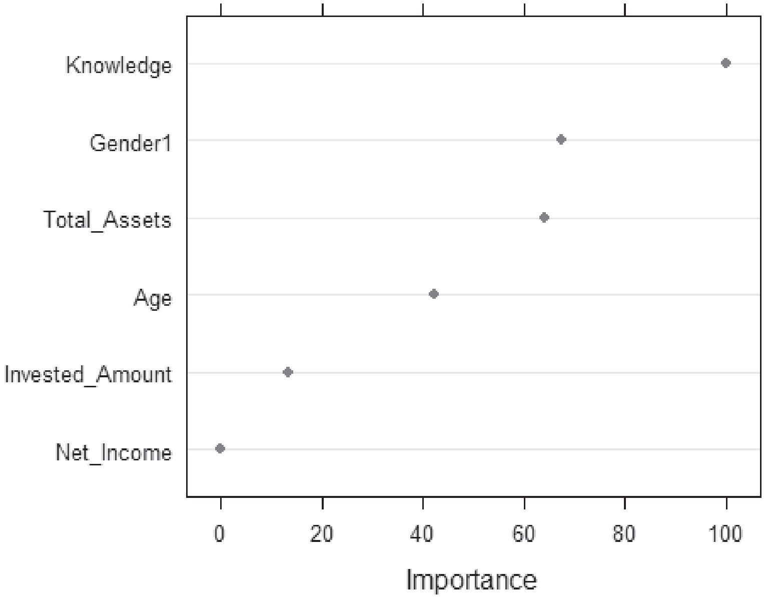

Figure A.1: Permutation variable importance measures for the random forest prediction of accepting advice in the risk tolerance assessment (variable permutation importance scaled to 100).

Justus Blaschke, M.Sc., ist externer Doktorand am Lehrstuhl für Mikroökonomie der Universität Kassel.

Anschrift: Möllmannsweg 23b, 48161 Münster, Deutschland, Tel.: +49 (0) 151/ 4613-1035, E-Mail: justusblaschke@gmail.com

Johannes Kriebel, Dr., ist akademischer Rat am Institut für Kreditwesen der WWU Münster.

Anschrift: WWU Münster, Universitätsstr. 14–16, 48143 Münster, Deutschland, Tel.: +49 (0) 251/8322-692, E-Mail: johannes.kriebel@wiwi.uni-muenster.de