

Zharova, Alona; Mihoci, Andrija; Härdle, Wolfgang Karl

Working Paper

Academic ranking scales in economics: Prediction and imputation

SFB 649 Discussion Paper, No. 2016-020

Provided in Cooperation with:

Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

Suggested Citation: Zharova, Alona; Mihoci, Andrija; Härdle, Wolfgang Karl (2016) : Academic ranking scales in economics: Prediction and imputation, SFB 649 Discussion Paper, No. 2016-020, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:

<https://hdl.handle.net/10419/146189>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

SFB 649 Discussion Paper 2016-020

Academic Ranking Scales in Economics: Prediction and Imputation

Alona Zharova*
Andrija Mihoci*
Wolfgang Karl Härdle*



* Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

Academic Ranking Scales in Economics: Prediction and Imputation*

Alona Zharova,[†] Andrija Mihoci[‡] and Wolfgang Karl Härdle[§]

Abstract

Publications are a vital element of any scientist's career. It is not only the number of media outlets but also the quality of published research that enters decisions on jobs, salary, tenure, etc. Academic ranking scales in economics and other disciplines are, therefore, widely used in classification, judgment and scientific depth of individual research. These ranking systems are competing, allow for different disciplinary gravity and sometimes give orthogonal results. Here a statistical analysis of the interconnection between Handelsblatt (HB), Research Papers in Economics (RePEc, here RP) and Google Scholar (GS) systems is presented. Quantile regression allows us to successfully predict missing ranking data and to obtain a so-called HB Common Score and to carry out a cross-rankings analysis. Based on the merged ranking data from different data providers, we discuss the ranking systems dependence, analyze the age effect and study the relationship between the research expertise areas and the ranking performance.

JEL classification: C14, C53, C81, M10

Keywords: scientometrics, ranking, quantile regression, Handelsblatt, RePEc, Google Scholar

*Financial support from the German Research Foundation (DFG) via Collaborative Research Center 649 "Economic Risk" and International Research Training Group 1792 "High Dimensional Non Stationary Time Series", Humboldt-Universität zu Berlin, is gratefully acknowledged. We are thankful for the research assistance provided by Marius Sterling.

[†]Humboldt-Universität zu Berlin, School of Business and Economics, Ladislaus von Bortkiewicz Chair of Statistics and Collaborative Research Center 649 "Economic Risk", Spandauer Str. 1, 10178 Berlin, Germany, tel: +49 (0)30 2093 5708, E-mail: zharovaa@hu-berlin.de

[‡]Brandenburg University of Technology, Chair of Economic Statistics and Econometrics, Erich-Weinert-Str. 1, 03046 Cottbus, Germany

[§]Humboldt-Universität zu Berlin, C.A.S.E. - Center for Applied Statistics and Economics, Spandauer Str. 1, 10178 Berlin, Germany and School of Business, Singapore Management University, 50 Stamford Road, Singapore 178899

1 Introduction

Academic ranking, as a tool of scientometrics, provides a platform for the evaluation of research results at universities, research centers, institutes, interdisciplinary groups, etc. Ranking systems indeed play an important role in performance comparison and the clarification of individual contribution to the overall institution ranking. For instance, the decisions made during recruitment processes at German universities (economic fields) are typically supported by HB rankings, see Schläpfer and Schneider (2010). The distribution of financial resources at universities is moreover based on performance-related schemes that among others include the comparison of the achieved research results, see Oberschelp and Jaeger (2010).

The evaluation phase of research results is often accompanied by missing observations within certain academic ranking systems. Employed statistical tools from quantile regression enable us here to impute and predict the missing ranking data. Since the focus of academic ranking lies additionally on finding the leading outcomes that significantly contribute to the excellence in research, our work deals with performance comparison among researchers utilizing ranking observations from Handelsblatt (HB), Research Papers in Economics (RP) and Google Scholar (GS) databases.

The goals of this paper include: (i) Developing a forecasting methodology for treatment of incomplete ranking data, (ii) Analyzing the relationship strength between HB, RP and GS rankings and (iii) Understanding the influence of age and investigated research fields on the academic ranks of scientists among analyzed ranking systems.

Our research questions correspondingly include: How to impute a cross ranking for business (BWL) and economic (VWL) sciences as well as to predict an academic rank for researchers, not included in a particular ranking system? How strong is the relation between HB, RP and GS rankings? Does the age, respectively the research area, influence academic ranks? How do the ranking scores change for different age groups and selected research fields in economics?

This paper considers different perspectives for the comparison of research results. Within one ranking system the analysis is straightforward. The challenge arises when the comparison has to be made jointly, through different ranking scales or ranking systems, see Wohlrabe (2011). Consider the HB ranking system that consists of several different rankings itself. Data are therefore often incomplete; for example, a top rated researcher of one ranking may not be listed in the other one(s) and vice versa. Such situations require a careful treatment of missing data. In mitigating these issues, we propose suitable quantile regression and correlation techniques in our work, while the conducted empirical study focuses on the dependence structure between HB, RP and GS ranking systems.

Quantile regression receives currently a relatively strong attention from the research community. A comprehensive description of quantile regression is given in Koenker (2005); the rapidly growing literature shows a variety of approaches and applications. Härdle and Song (2010) construct a uniform confidence band for quantile regression, Fitzenberger and Wilke (2005) provide an empirical application of quantile regression to duration data, Guo et al. (2013) fulfill a functional data analysis of generalized regression quantiles. A statistical analysis of linear as well as nonlinear models with the help of quantile regression through the estimation of conditional quantile curves has been proposed by Spokoiny et al. (2013). Here we employ a quantile (median) regression in order to impute missing academic rankings data.

Based on the conducted analysis, we show that a median regression model successfully interpolates the missing academic rankings. Academic rankings data exhibit different correlation structures between underlying sub-rankings of HB, RP and GS, whereas the academic ranking variation has been documented to be quite sensitive to the age changes. For example, the rank of younger as well as advanced age scientists is marginally changing (increasing) stronger than the rank of middle-age researchers. Interestingly, researchers from mathematical and quantitative methods' fields possess the leading positions across all three discussed ranking systems, whereas scientists specializing in microeconomics, international economics and general economics are presented in the dominant ranking

part within HB, RP and GS respectively.

The proposed approach as well as the findings of this research can be successfully used in practice (i) by selection committee in the recruitment process at universities (economic fields), (ii) as a unique tool in decision making related to the allocation of research funds, (iii) for collaborative purposes and grant proposal applications, etc. Our calculated HB Common Score finally can be used for a simultaneous comparison of candidates' profiles from business (BWL) and economic sciences (VWL).

The paper is structured as follows. The description of ranking systems and our data source is presented in Section 2. Section 3 describes the data and introduces the implementation of the predicting techniques. Section 4 discusses the HB, RP and GS comparison results and provides evidence on the impact of age and the research fields on ranking performance. Finally, Section 5 concludes.

2 Academic Ranking Systems and Scales

For our comparison purposes we choose three ranking systems, namely HB, RP and GS, that are mostly used in economic and business fields among the German speaking countries, such as Germany, Austria and Switzerland. In this paper we repeatedly use the terms *ranking*, *rank* and *score*. The first one represents the academic system or scale that deals with the selected researchers. *Rank* denotes the position of each individual within the ranking scheme, while the *score* denotes the number of points assigned to a particular person. The analysis was furthermore carried out using R and MATLAB. In order to enhance the transparency of this research, all the used programming codes are available on the web-based repository hosting service and collaboration platform GitHub (accessed 22 Apr 2016) and QuantNet (accessed 05 Mai 2016).

2.1 Handelsblatt

The HB rankings provide the lists of the most active publishing researchers in business and economic fields in Germany, Austria and Switzerland, as well as German speaking researchers outside of these countries. The rankings have been developed by the Konjunkturforschungsstelle (KOF) of the ETH Zürich on behalf of HB and German Association for Social Policy (Verein für Sozialpolitik). For this purpose the publication data from several external databases, as well as the data from the Forschungsmonitoring (accessed 14 Oct 2015) are used. The HB ranking has an established reputation among the German speaking economists, since it influences decision making at universities regarding the distribution of funds, the recruitment process and performance evaluations, Schläpfer (2011).

HB creates and publishes a journal ranking out of selected journals which are indexed in the The American Economic Association's electronic bibliography – EconLit, see Combes and Linnemer (2010). Every journal from the HB list receives a weight between 0.05 and 1, where a higher weight indicates a better rank. As reported by Gygli et al. (2014), the individual researcher's rank is built from the number of weighted publications in relevant journals divided by the number of coauthors.

HB considers two fields: business sciences (BWL) and economics sciences (VWL). Within each field we analyze three sub-rankings from 2009 to 2015, namely the Researcher Life's Work (LW), Current Researchers (CR) and Researchers Under 40 (U40). This gives us a total of six BWL and VWL sub-rankings that are usually published once in two years. The CR ranking is based on the rated researchers' publications in predetermined journals over the last five years, whereas the U40 ranking considers all scientists younger than 40 years. The LW ranking, finally, takes all rated researchers publications from the HB journals list into account. It is worth noting that every researcher can be present either in VWL or in BWL rankings. Although inside each of these fields the person can belong to any ranking of LW, CR or U40 (the last only if additionally younger than 40).

It is interesting to trace the rank development over time in an orthogonal coordinate system, as well as to represent a high-dimensional data (up to four dimensions for VWL) of each year in the same scale, namely LW, CR or U40. This task may be performed using parallel coordinates plots, Härdle and Simar (2015), see Figures 1 and 3, for VWL and BWL, respectively. The coordinates representing a single year are drawn in parallel axes and the ranking scores of the selected researchers in various years are connected through the axes with straight lines. Each line represents rankings of a single person through the analyzed years. Red dashed lines denote the three quartiles (25%, 50% and 75%). All variables on the y -axis are scaled to maximum equal to one and minimum equal to zero. The maximal score of each year is shown on the upper x -axis, due to the dynamics of the maximal scores over time.

One can promptly detect a positive trend of the LW and CR rankings for both fields, VWL and BWL. This tendency has been confirmed by the plotted quartiles as well as the maximum ranking score for each year. Outside the quartiles, one may identify outliers evident from the boxplots in Figures 2 (VWL) and 4 (BWL).

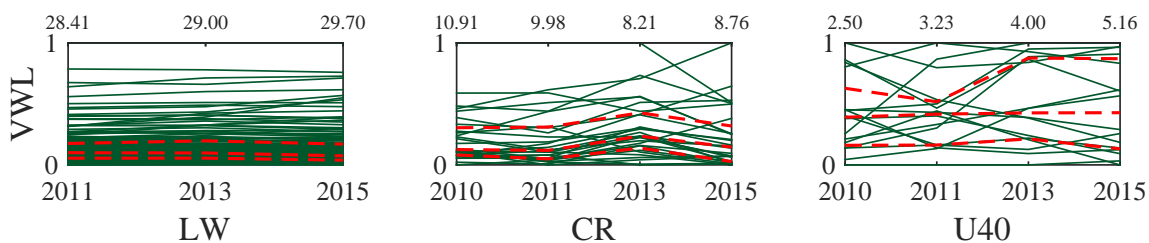


Figure 1: Parallel coordinates plots for standardized scores (y -axis) of researchers within HB VWL sub-rankings from 2010 to 2015. Green lines represent scores of researchers with a specific ranking score for each year. Red dashed lines denote the three quartiles (25%, 50% and 75%). Upper x -axis numbers show the maximum ranking score for the corresponding year. Plotted are scores of researchers that are listed in all analyzed years.



For better understanding of a distributional spread the related data are summarized using boxplots in Figure 2 and Figure 4. By contrast, the LW boxplots for both VWL and BWL appear to have almost similar centers with a slight positive trend. The data

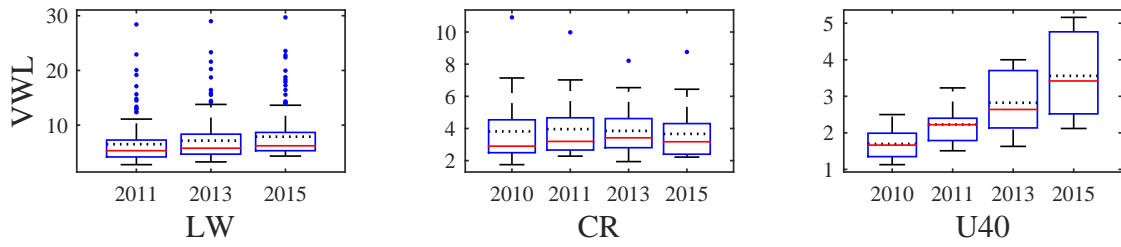



Figure 2: Boxplots for ranking scores of researchers within HB VWL sub-rankings: LW (158 observations), CR (28) and U40 (16) from 2010 to 2015. The red lines denote the median, whereas the dotted lines display the mean.  ARRboxhb

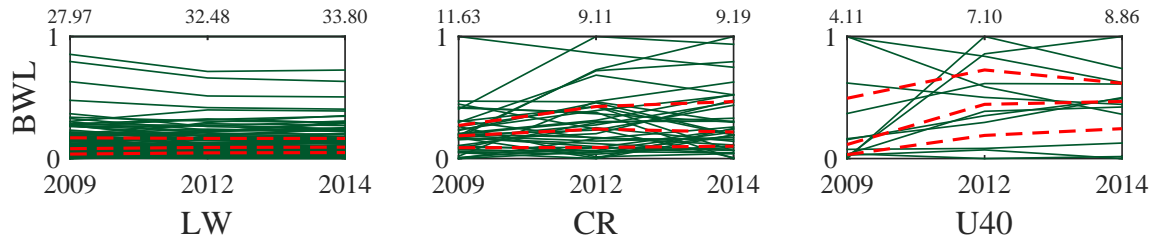



Figure 3: Parallel coordinates plots for standardized scores (y -axis) of researchers within HB BWL sub-rankings from 2009 to 2014. Green lines represent scores of researchers with a specific ranking score for each year. Red dashed lines denote the three quartiles (25%, 50% and 75%). Upper x -axis numbers show the maximum ranking score for the corresponding year. Plotted are scores of researchers that are listed in all analyzed years.  ARRpcphb

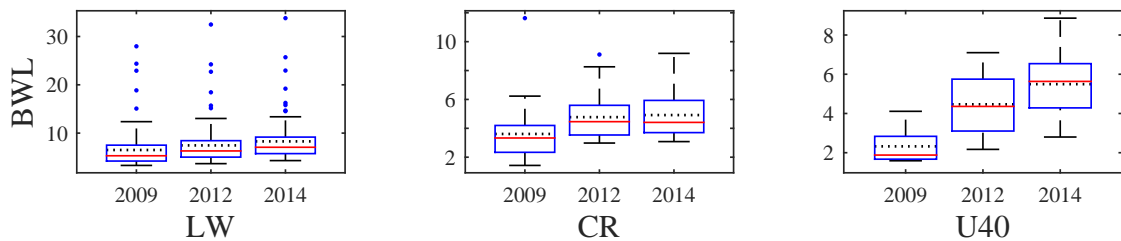



Figure 4: Boxplots for scores of researchers within HB BWL sub-rankings: LW (128 observations), CR (37) and U40 (12) from 2009 to 2014. The red lines denote the median, whereas the dotted lines display the mean.  ARRboxhb

of VWL U40 for 2015 seems to have larger variability than the other ones. The pairwise comparison of the LW, CR and U40 rankings from VWL with the corresponding ones from BWL indicates similar distribution. Almost all boxplots are skewed to the right, as the top whiskers, connected to the larger scores, are much longer than the bottom ones. Strikingly, the U40 scores improve significantly over time that indicates a distinct increase of young researchers' performance for the last several years. Also interesting is the slight increasing change of the BWL CR scores in 2009 – 2014.

For the research in this paper we utilize VWL LW 2015 and BWL LW 2014 rankings, as they are the most recent ones and have the largest number of different considered researchers equal to 500. In order to perform the analyses of the research fields and researcher' age on the score, we have to eliminate the persons with missing observation, i.e. with no information on age or research fields. Thus, the age data is available for 458 persons, while 448 have information about the research area provided by HB.

2.2 Research Papers in Economics

The RP ranking system collects the bibliographic data of journal articles, books, working papers and other scientific media outlets. It contains around 2 million research items from more than 2300 journals and 4300 working paper series. Although the RP project offers a broad spectrum of services, in this paper we focus solely on author ranking that relies on data from more than 46 thousand registered researchers, see RePEc (accessed 22 Apr 2016).

The main idea of RP author ranking is to monthly publish a list of top 5% of researchers based on a average rank score. For each author this score is calculated based on a two step procedure. Firstly the authors are individually ranked within each of the 36 separate sub-rankings, excluding the w-index, a special case of the h-index. Secondly, a harmonic mean of the individual ranks represents this average rank score. For convenience, all sub-rankings are listed in the Table 4 in Appendix. In contrast to HB and GS, one should

note that within the RP system the top-ranked scientists receive the smallest score and vice versa. For more details, we refer to Zimmermann (2013) and the corresponding RP webpage.

Contrary to HB, all RP sub-rankings receive the same weight while providing the average rank score, although they per se may impose a weighting scheme. To boost the HB score, for instance, an author is considering the journal ranking list, whereas in order to improve the RP score, researchers have to work on other publication aspects, such as the number of citations, abstract views, etc. Since the HB ranks are available up to 2015 inclusive, the RP data have been collected for December 2015 for comparison purposes. The high-dimensional structure of RP ranks is introduced with the help of parallel coordinates plot in Figure 24 in Appendix.

2.3 Google Scholar

Different from HB and RP, GS concentrates on citation data, Hamermesh (2015). For every researcher it provides information about the number of citations per paper, the total number of citations, as well as the values of the h -index and the $i10$ -index. While calculating its metrics, as it is shown by Dilger and Müller (2011), GS takes all types of research publications into account. GS has a good coverage in social sciences, economics, finance and business administration, see also Harzing and Wal (2008), which makes it a desirable choice for our research purposes.

For our comparison through ranking scales we use GS number of citations. The parallel coordinates plot in Figure 5 gives an insight on the development of citation count for selected researchers over time. The upper x -axis numbers denote the maximum number of citations across all researchers for the corresponding years. As it can be inferred from these values, the maximum number of citations increases steadily over time, as does the citation count of top scientists. The relative position of best researchers does not show any significant temporal change. We attribute the visible decrease in 2015 to the time lag

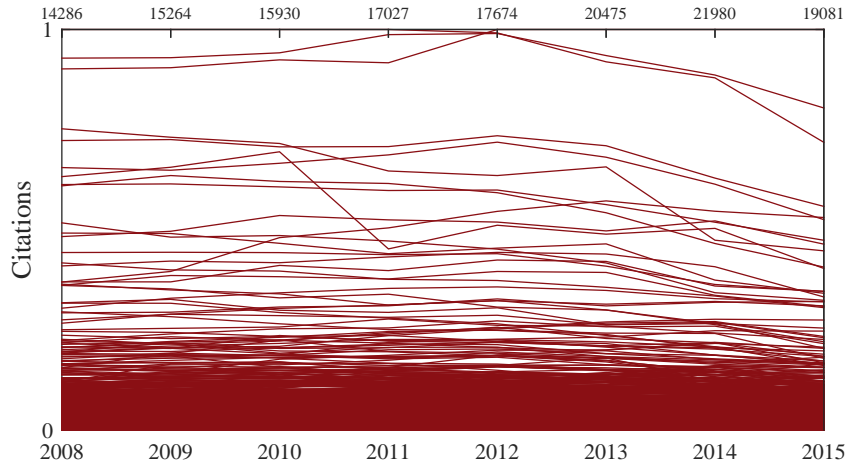



Figure 5: Parallel coordinates plot for standardized GS citations (y -axis) from 2008 to 2015 for selected 1357 researchers. Upper x -axis numbers show the maximum number of citations for the corresponding year.  ARRpcpgscit

effect (time between publishing of the research work and its citing by other scientists).

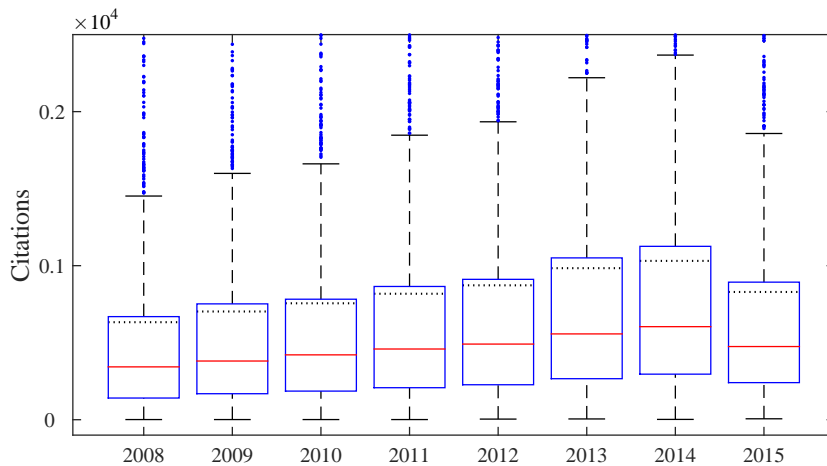



Figure 6: Enlarged boxplot for GS citations from 2008 to 2015 for selected 1357 researchers. The red lines denote the median, whereas the dotted lines represent the mean.  AARRboxgscit

A positive trend of the number of citations is observed over the years on the enlarged lower part of the boxplot in Figure 6. The medians are slightly increasing with similar distributions of citations for all years. The largest values in the sample are outliers and are depicted as blue circles. All GS boxplots indicate relatively small spread of data (interquartile range) that can be seen from the box lengths.

3 Data

Based on the academic ranking systems introduced in the previous section, we present the database used in this paper and provide the corresponding summary statistics in Subsection 3.1 and in Subsection 3.2 the methodology for treatment of missing HB ranking data for the selected scientists.

3.1 Data and Descriptive Statistics

Our paper aims to compare ranking data from HB, RP and GS scales. Using information from one ranking system can be inadequate for comparison of individuals. The HB ranking, for example, does not account for the information about monographs, although in German speaking countries this art of publication remains quite popular, as discussed in Wohlrabe (2013). For researchers of advanced age thus a lower rank may be expected. As there exists a relatively strong correlation between HB and country-specific RP rankings, according to Wohlrabe (2011), it is reasonable to consider more ranking systems in performance evaluation.

Our paper correspondingly considers HB (data from 2014 – 2015), RP (December 2015) and GS data (December 2015). In order to take into account both economic and business sciences, we select two main HB rankings with data available for 500 scientists: (i) the VWL LW in 2015 for 250 persons and (ii) BWL LW in 2014 for 250 persons. In December 2015, 2304 researchers have been listed in RP top 5% author ranking. From them, 1027 had a GS profile with corresponding GS scores. The detailed descriptive statistics is shown in Table 5, which is for convenience included in the Appendix.

A more detailed view on the data merging results is depicted in the mosaic plot, Figure 7. Consider the 500 scientists in HB. There are 122 persons that have also a RP score, but not a GS profile. Similarly, 260 persons have a HB and GS, but no RP ranking data. Finally, there are 84 researchers for which the HB, RP and GS data are available.

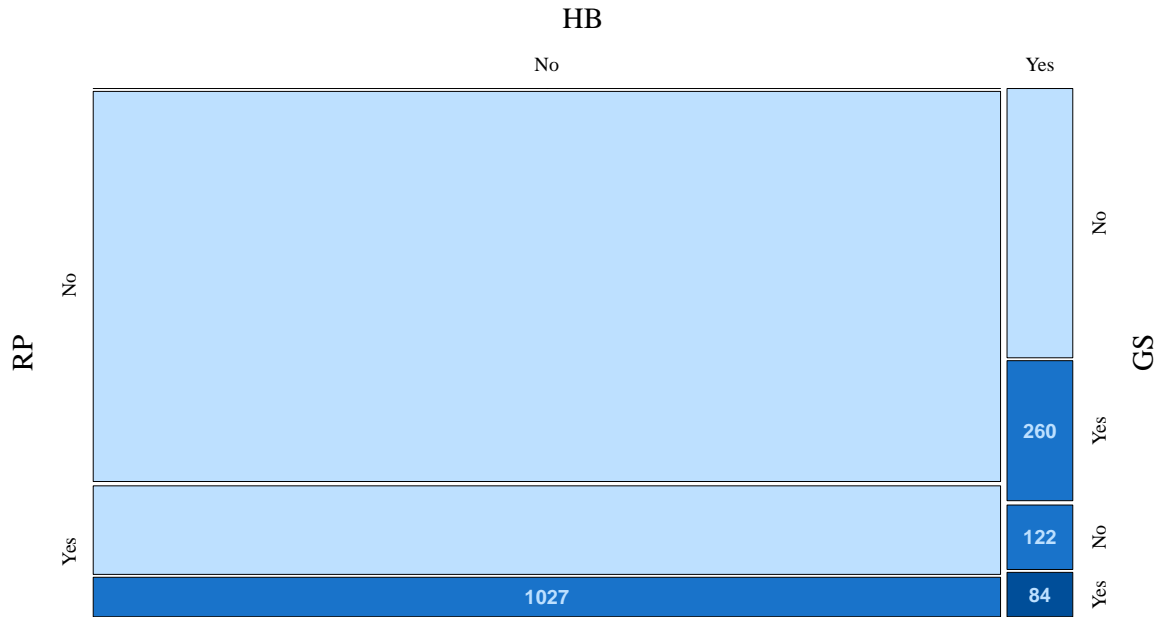



Figure 7: Mosaic plot for the number of researchers, when merging of HB, RP and GS rankings takes place or not (Yes/No). The number of GS profiles is quite large and is here shown only figurative. Ranking data of researchers in 2015 (RP, GS and HB VWL LW) and 2014 (HB BWL LW) have been used.  ARRmosage

In our research it is convenient to utilize the data of these 84 scientists, if not specified differently.

3.2 Predicting the Handelsblatt Common Score

Recall from Section 2.1, there are 500 researchers that are included either in the VWL (250 persons) or the BWL (250 persons) ranking list. As it follows from the mosaic plot in Figure 7, 84 researchers have ranks in all three ranking systems (HB, RP and GS). A closer look at the data indicates that these 84 HB researchers belong either to VWL LW or to BWL LW rankings. By predicting and imputing the unobserved rankings, here the VWL LW scores can be compared with the BWL LW scores of researchers. For this purpose, we transform the HB sub-rankings into one unique rank, the so-called *HB Common Score*. For its computation we consider linear as well as quantile (median) regression.

3.2.1 Quantile Regression

One can describe a relationship between two variables through a linear regression (LR) model, see, e.g., Härdle and Simar (2015)

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, \quad i = 1, \dots, n, \quad (1)$$

where β_0 denotes the intercept and β_1 depicts the regression line slope with ε_i denoting the error term. Here n stands for the sample size, i.e., in our case the number of scientists. The response observation of individual i is denoted by y_i , and x_i is the i -th value of the regressor.

Referring to Koenker and Bassett (1978) and Koenker and Hallock (2001), we introduce the quantile regression (QR) model related to the linear regression (1) as

$$y_i = \beta_0^{(\tau)} + \beta_1^{(\tau)} x_i + \varepsilon_i^{(\tau)}, \quad i = 1, \dots, n, \quad (2)$$

where $\tau \in (0, 1)$ denotes the quantile level and the error $\varepsilon_i^{(\tau)}$ has τ -quantile zero. For instance, setting $\tau = 0.5$, results in median quantile regression.

In the estimation of the linear regression model, the estimates of the unknown intercept and the slope parameter are found by least square minimization

$$\left(\widehat{\beta}_0, \widehat{\beta}_1\right) = \arg \min_{\beta_0, \beta_1} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2. \quad (3)$$

In median quantile regression the estimates of the parameters are obtained by the minimization of the sum of absolute residuals

$$\left(\widehat{\beta}_0^{(0.5)}, \widehat{\beta}_1^{(0.5)}\right) = \arg \min_{\beta_0^{(0.5)}, \beta_1^{(0.5)}} \sum_{i=1}^n |y_i - \beta_0^{(0.5)} - \beta_1^{(0.5)} x_i|. \quad (4)$$

The linear model (1) defines the relation between dependent variable and regressors, as an average through the conditional mean function. On the contrary, the quantile regression (2) offers a broader perspective, since it models the conditional τ -th quantile, providing the possibility to depict the interconnections at various points, see Koenker (2015) and Baum (2013). For instance, for $\tau = 0.5$ the conditional median function is resulting, a functional that is of bounded influence, i.e. robust w.r.t. outliers.

The analysis of data with thick tail, non-normal errors may not only become challenging, but also biasing for the linear model, and indeed our data has many outliers. For the sake of brevity we refer to the figures presented. In summary, the quantile regression is robust in the presence of outliers and therefore we use median quantile regression for further analysis.

3.2.2 Parameter Estimation and Estimation Quality

Consider for convenience the top 100 VWL LW (y_i) as well as the top 100 BWL LW (x_i) scores, $i = 1, \dots, 100$, and then fit the equation (4). Denote the estimated model parameters by $\hat{\beta}_0^{(0.5)}$ and $\hat{\beta}_1^{(0.5)}$. Then the estimated HB Common Scores (VWL LW), using the analyzed $n = 100$ pairs (y_i, x_i) , are found by

$$\hat{y}_i = \hat{\beta}_0^{(0.5)} + \hat{\beta}_1^{(0.5)} x_i, \quad i = 1, \dots, 100 \quad (5)$$

Empirical results show an excellent explanatory performance. Figure 8 includes the scatterplot with the imposed fitted median quantile regression line, as well as the Quantile–Quantile (QQ) plot of the associated residuals versus the quantiles of the normal distribution. The fitted least squares regression line lies below the fitted quantile regression line. The least squares approach provides a prediction of mean VWL LW score given the average values of the BWL LW score. The quantile regression gives a forecast of the τ -th quantile of VWL LW; in our case the median. Note that here the fit of least squares is

affected by the outliers, thus by relatively high and low scores.

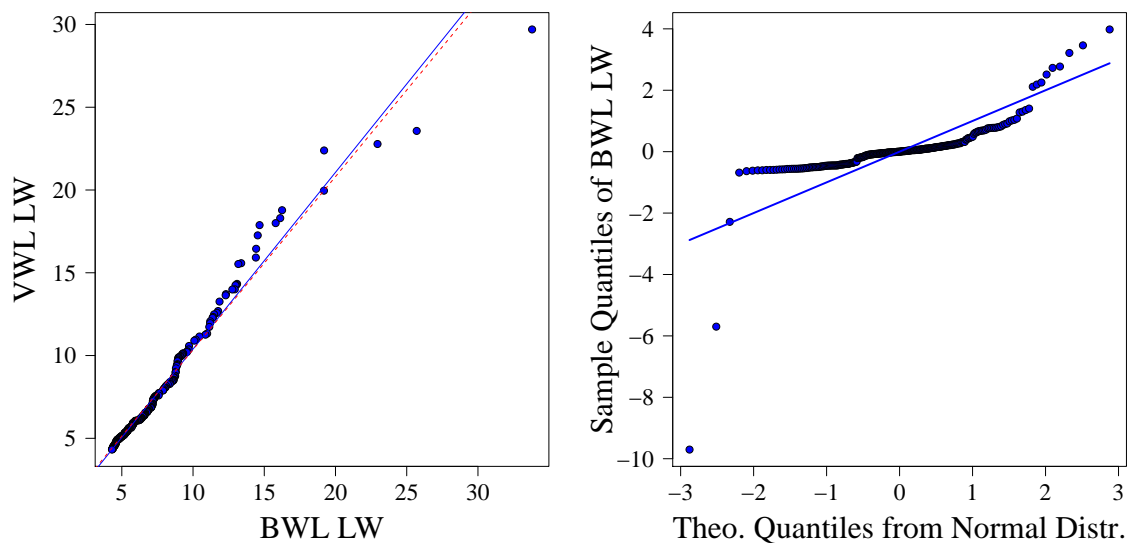



Figure 8: Scatterplot and Quantile Regression Fit (left) of the HB on VWL LW vs BWL LW for a sample of 100 top researchers within these rankings. Superimposed on the plot is the 0.50 quantile regression line (solid blue) and the least squares estimate of the conditional mean function (dashed red line). The coefficient of determination of the median regression equals to 0.93. A QQ plot (right) of the same sample of data versus a normal distribution.  ARRqrqhb

The estimation quality is illustrated by the QQ plot, where empirical quantiles from the quantile median regression function are depicted on y -axis and x -axis fits theoretical quantiles from a given distribution, as introduced in Ricci (2005). The QQ plot (right plot in Figure 8) includes the residuals from the quantile regression that are plotted against the theoretical quantiles of the normal distribution.

The estimated parameters of the quantile regression analysis, summarized in Table 1, are used for the HB Common Score estimation. The goodness-of-fit results are presented in Table 2. The standard error (SE) of the estimates in Table 1 reflects a relatively low variability. The coefficient of determination from Table 2, as a goodness-of-fit measure, is equal to 0.93. This means that the variation of the VWL LW is quite convincingly explained by the quantile regression, i.e., there exist a strong dependence of VWL LW with respect to BWL LW. The quantile regression analysis consequently enables us to predict the HB Common Score for researches that have different HB scores.

| | | Est. | SE | t | p -value |
|--------|-----------------------------|-------|------|-------|------------|
| BWL LW | $\widehat{\beta}_1^{(0.5)}$ | -0.28 | 0.21 | -1.37 | 0.1725 |
| | $\widehat{\beta}_0^{(0.5)}$ | 1.07 | 0.04 | 27.71 | 0.0000 |

Table 1: Estimated regression model parameters (Est.) for rankings between VWL LW (dependent variable) and BWL LW (explanatory variable) for HB researchers. We provide the standard error of estimates (SE), the test statistic t for testing whether the null hypothesis that the true parameter equals 0 and the associated p -value.

| | MSE | r^2 |
|--------|--------|--------|
| BWL LW | 0.9976 | 0.9308 |

Table 2: Mean squared error (MSE) and coefficient of determination of the regression model for rankings between VWL LW (dependent variable) and BWL LW (explanatory variable) for HB researchers.

The so-called HB Common Score is represented either by the existing VWL LW score or the predicted one. For 250 researchers of VWL LW their existing scores depict HB Common Scores. The remaining 250 VWL LW scores have been predicted from BWL LW and now represent the HB Common Score (HB CS). In total, 500 HB Common Scores are associated to the 500 researchers.

4 Score Comparison

As was implied in Section 3.2, the HB Common Score can be used for further comparison with other ranking scales in Section 4. In Subsection 4.1 we investigate the relations between HB, RP and GS. The influence of age on ranking scores is examined in Subsection 4.2, whereas Subsection 4.3 shows the results of research fields analysis.

4.1 HB, RP and GS

From now on when we write HB, RP or GS we mean HB Common Score, RP average rank score and GS citations, if it is not stated differently.

For visualizing the data distribution we produce histograms and corresponding boxplots in Figure 9 for HB, RP and GS.

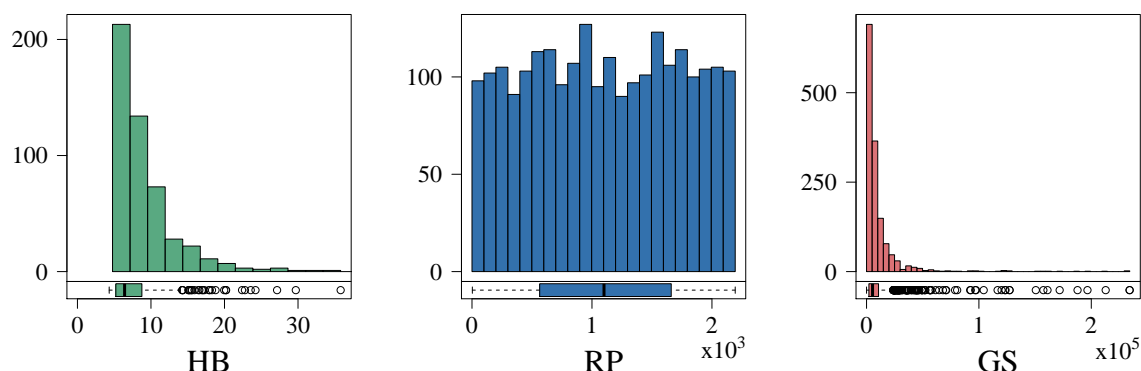



Figure 9: Histogram of HB (500 observations, Common Score), RP (2304, total score $\times 10^3$) and GS (1357, citations $\times 10^5$) rankings for December 2015.  ARRhismer

The distributions of HB and GS scores of researchers are asymmetric, right-skewed and single-peaked. The heavy tails stretching away from the peaks indicate the presence of many outliers that fall outside the overall pattern, here associated with extreme values. We have a concentration of data in the left part and a long tail to the right. This represents the vast majority of scientists with lower ranks, with few persons possessing very high ranks. In the RP scores distribution, in contrast, one can identify multiple peaks close together, slight symmetry, but no apparent pattern. This can be explained by the structure of RP average rank score, as it is calculated from 36 another sub-rankings that was discussed in Section 2.2.

The data seem to be centered for HB around 6.41 points, for RP at 1100 scores and for GS at 5332 citations, as depicted at the boxplots on the lower part of Figure 9. This implies that roughly half of ranks are given to the scientists possessing ranks less than the mentioned levels.

The parallel coordinates plot in Figure 10 shows a moderate and positive dependence between HB, RP and GS scores. The RP is here rescaled in order to show all the best individuals as a maximum value on the upper scale. The three quartiles (25%, 50% and 75%) indicate a considerable number of outliers that influence the results. This can be

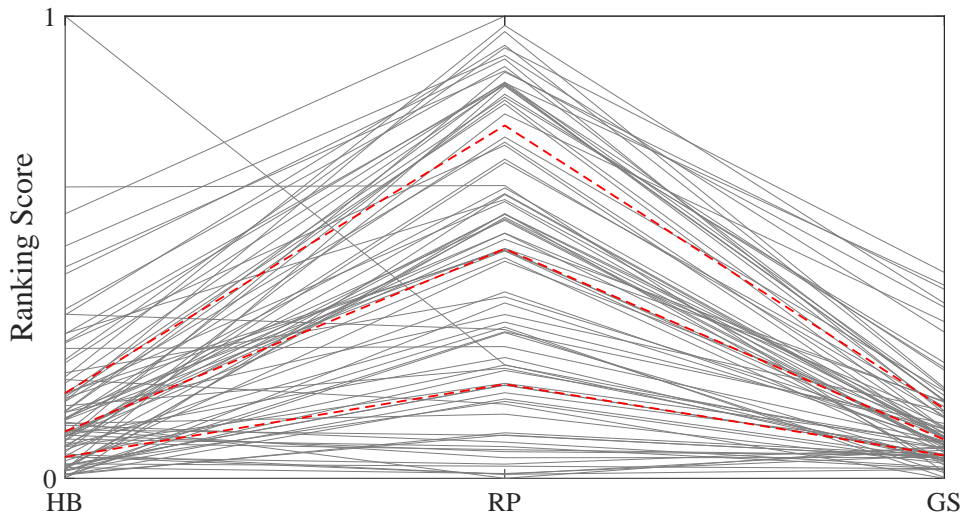


Figure 10: Parallel coordinate plot for three variables (HB, RP and GS) on 84 researchers for December 2015. RP values are rescaled. Red lines denote the three quartiles (25%, 50% and 75%).

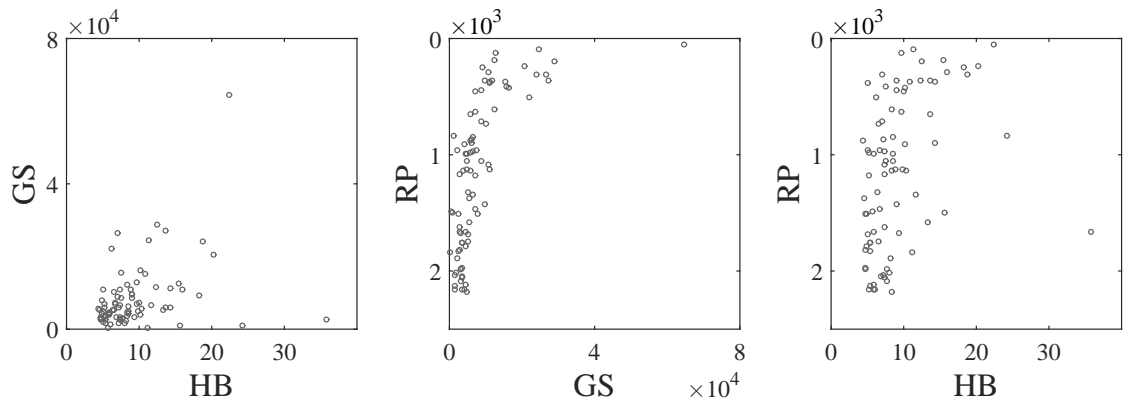


Figure 11: Scatterplots of HB, RP and GS rankings for 84 researchers in December 2015. The correlation equals to 0.44 for HB–GS plot, -0.64 for GS–RP and -0.46 for HB–RP.



confirmed, while removing the extreme scores from HB and GS. The result is shown in Figure 25 in Appendix.

Figures 11 and 12 present two dimensional scatterplots and corresponding density estimate for pairs of HB and GS, RP and GS, as well as HB and RP. The contour lines indicate the height of the density. Some separate distributions could be identified in this two dimensional space, but they have visible overlappings.

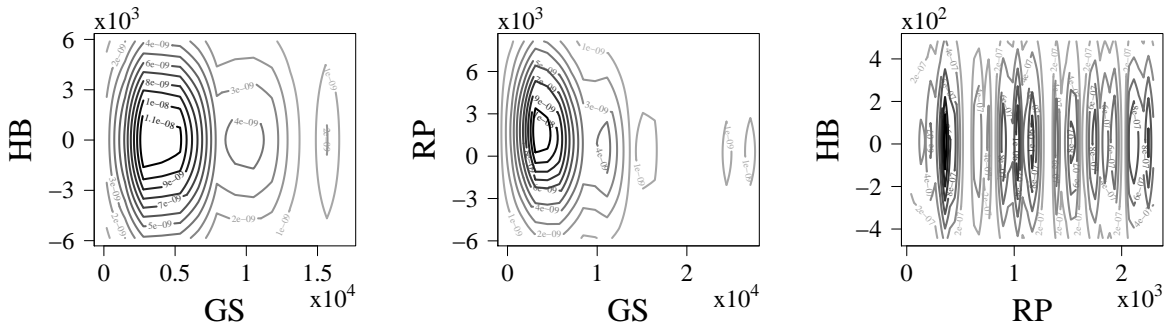



Figure 12: Contour plots of the kernel density of HB and GS (left), GS and RP (center), as well as HB and RP (right). The contour lines indicate the height of the density. Here the data of 84 researchers from 2015 is analyzed.  ARRdenmer2d

The relationship between HB, RP and GS scores is further analyzed with correlation matrix in Figure 13 for the full data frame consisting of 42 factors. Here we use the HB Common Score and also include the age of researchers as an additional factor. The descriptive statistics is introduced in Table 6 in Appendix.

The correlation plot reveals that a lot of variables indicate a strong linear relationship. In particular, the correlation between GS citations and other variables varies mainly moderate to strong. The HB Common Score shows in the most cases moderate correlation. The visible clusters that characterize RP data corresponding to the groups of rankings introduced in Table 4. The negative correlation between RP average rank and other variables is due to the difference in scales, as explained in Section 2.2.

One can notice that the RP and GS citations and h -index show a very strong correlation. These pairwise relations are additionally explored through the scatterplots and hexagon

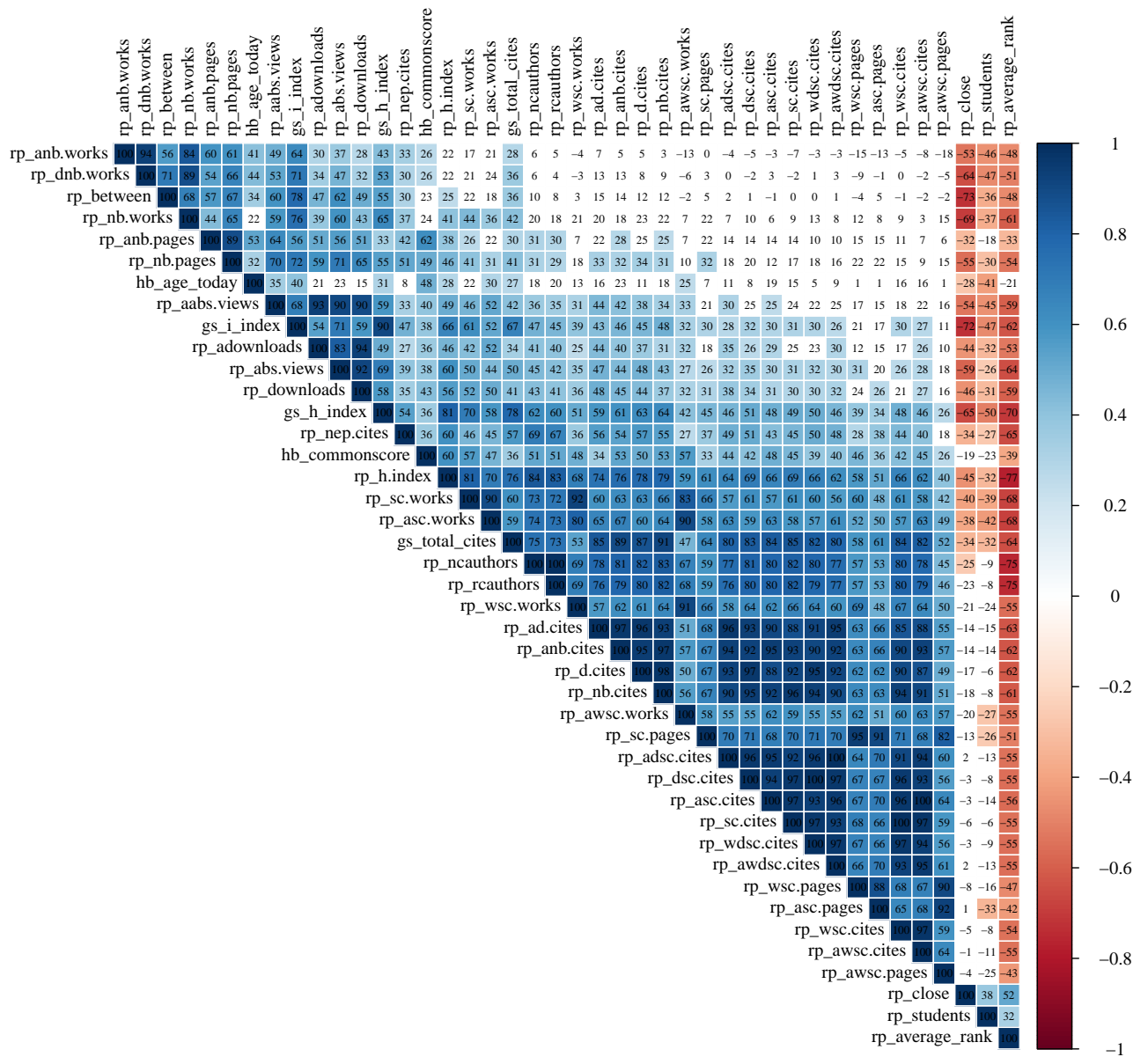


Figure 13: Correlation matrix of 42 factors of HB, RP and GS for 84 researchers in December 2015. The color depicts the strength of correlation: from positive (blue) to negative (red).



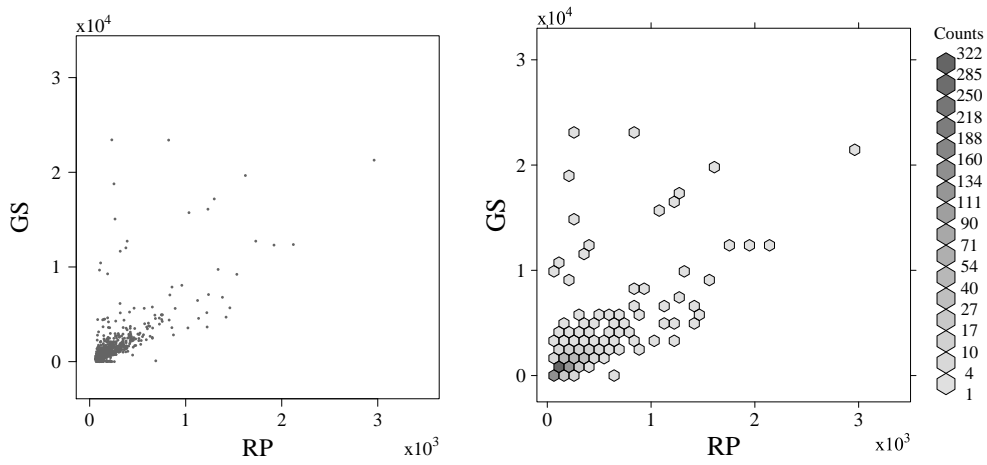



Figure 14: Scatterplot (left) and hexagon plot (right) of RP and GS citations for 1024 researchers in December 2015. Correlation coefficient equals to 0.70.  ARRhexcit

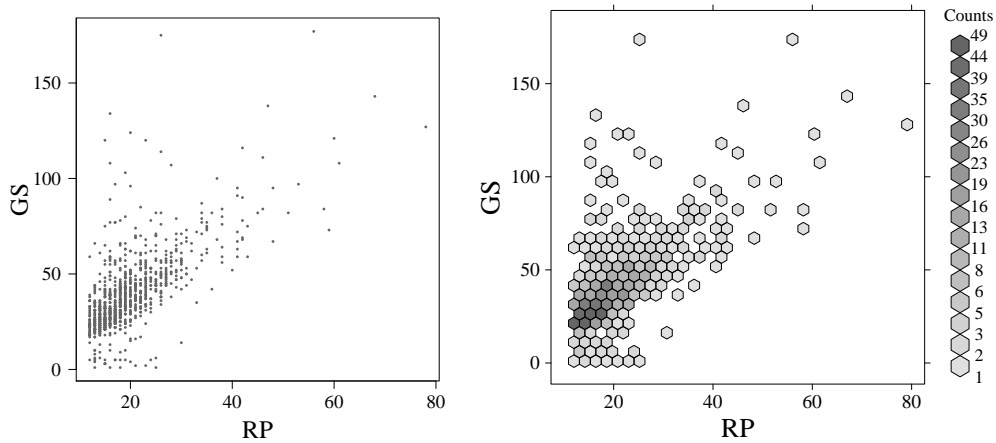



Figure 15: Scatterplot (left) and hexagon plot (right) of RP and GS h-index for 928 researchers equals to 2015. Correlation coefficient equals to 0.68.  ARRhexhin

plots on the Figures 14 and 15 correspondingly. The Figures indicate a positive linear relationship between the two. However, there are some outliers that do not follow the trend.

4.2 Influence of Age

Our research question is to study whether the age influences the rankings of scientists. As the age data is available only for 458 persons from HB, we also reduce the full datasets within RP and GS to the top 458 observations, see Section 2. The scatterplots, hexagonplots and boxplots in Figures 16 – 20 show the relations between age and ranking scores in a more detailed way.

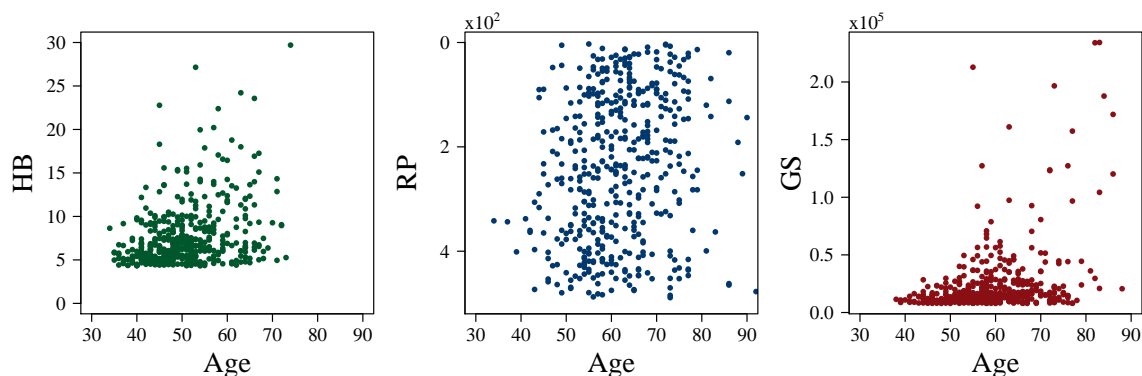


Figure 16: Scatterplots for age and ranking scores of HB, RP and GS for Top-458 scientists within each ranking system for December 2015.

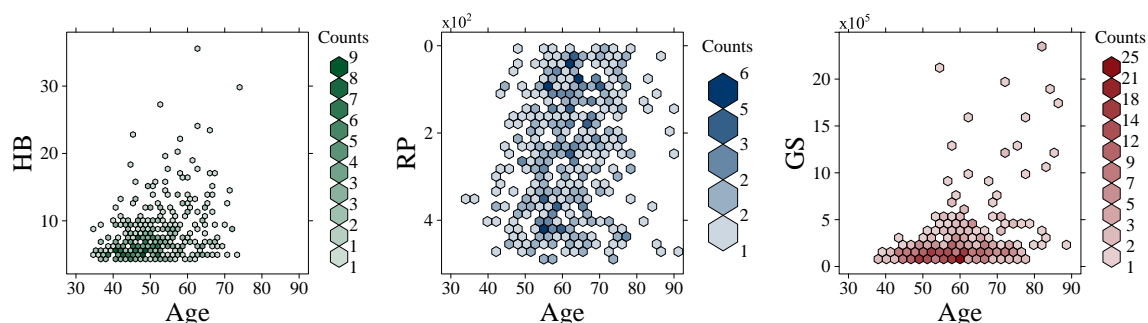


Figure 17: Hexagon plots for age and ranking scores of HB, RP and GS for Top-458 scientists within each ranking system for December 2015.



From Figures 16 – 17 one can make several observations. First, there exists a positive

relationship between age and HB ranks. For RP it is difficult to identify any pattern of data points. Here it is important to note that some RP rankings are standardized with respect to age, see Table 4 in Appendix. At the same time it seems to be a very weak association between age and GS.

For research aggregate we divide ranking scores of scientists into nine groups with respect to their age with the step of five years, starting from persons younger than 36 years old and summing up with ones older than 70 years. The overall patterns of response for the age groups are described on the boxplots in Figures 18 – 20.

The notable high box length of ranks from RP age groups gives an indication of the high sample variability. On the other hand, the comparatively short boxplots from GS age groups indicate that GS researchers have only slight difference on the introduced scale. In the same way, the boxplots of HB are comparatively tall. This suggests that top 458 of HB scientists have relatively different ranking scores. Almost all age groups of HB moreover indicate the presence of heavy tails in the direction of higher ranks, as in some cases the length of the whiskers exceeds the length of the boxes.

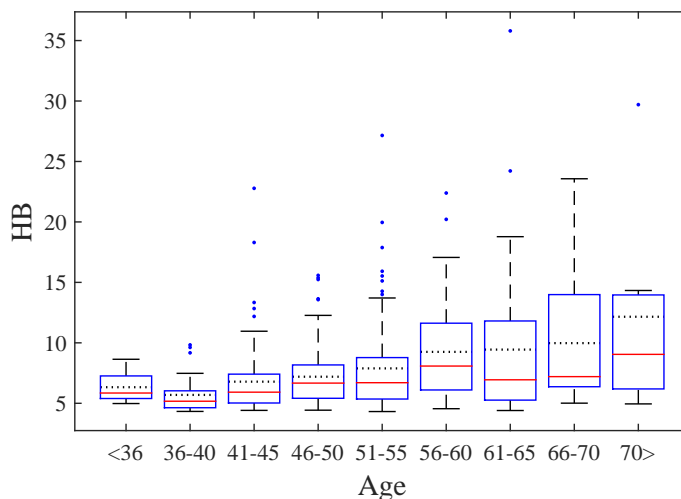



Figure 18: Boxplots for age and ranking scores of HB for Top-458 scientists within each ranking system for December 2015. The red lines denote the median, whereas the dotted lines introduce the mean.  ARRboxage

A further analysis enables us to summarize, that ranks of younger researchers are increas-

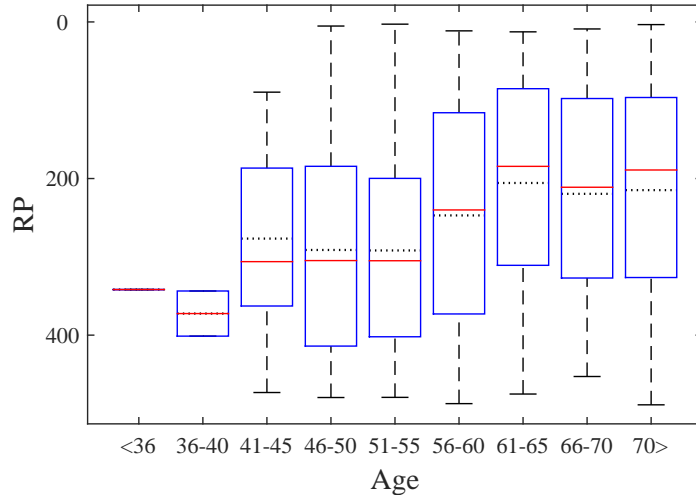



Figure 19: Boxplots for age and ranking scores of RP for Top-458 scientists within each ranking system for December 2015. The red lines denote the median, whereas the dotted lines introduce the mean. For comparison purposes the scale is inverted.  ARRboxage

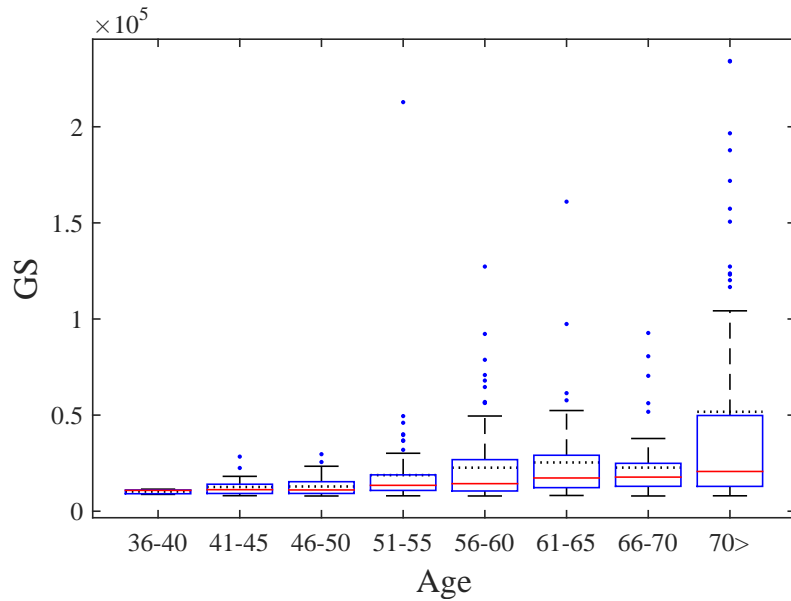



Figure 20: Boxplots for age and ranking scores of GS for Top-458 scientists within each ranking system for December 2015. The red lines denote the median, whereas the dotted lines introduce the mean.  ARRboxage

ing, whereas the middle-age group has relative consistency or slight decline and then the next growth, now between scientists of advanced age, could be observed. One of the possible explanations of this observation could originate from a scientific path in academia. In order to get a position in a university, young researchers are motivated to write more papers and produce other significant outputs. While the middle-age researchers, who usually have stable position, are more concentrated on teaching, long-term projects and other duties. The slight increase of ranking scores of older persons could be explained by experience in writing papers, acknowledgment among scientific community, enlarged research network they are working within and other reasons. As a result, this lead to a higher number of works, citations, indexes, number of papers downloads, etc.

The relative comparison of three academic rankings through the age groups in a four-dimensional plot (HB, RP and GS scores and age) is represented by a mosaic plot in Figure 21. We consider three academic rankings with the Top-458 researchers from each one. Here HB, RP and GS scores are shown by green, blue and red colors respectively. The width of columns represents the number of persons within each age group, whereas the colored dot represent zero.

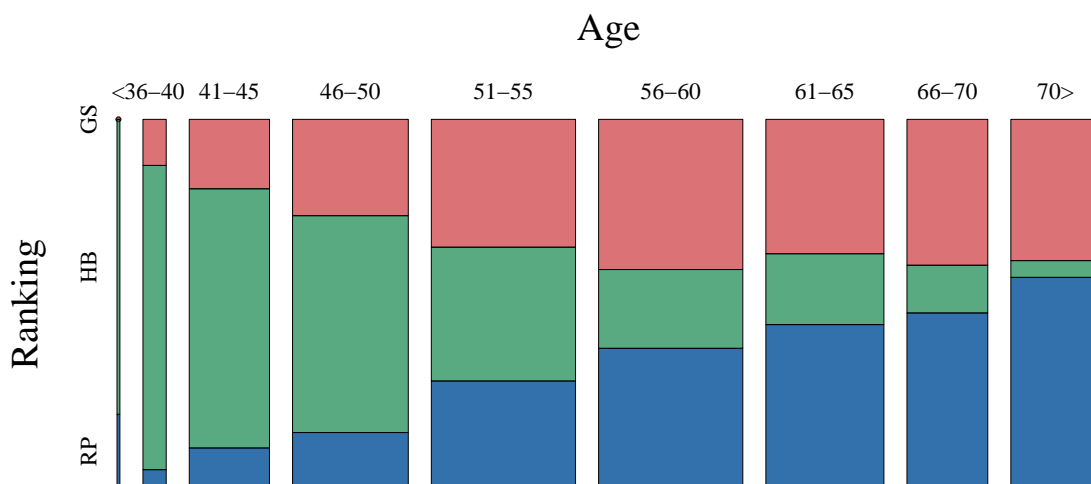



Figure 21: Mosaic plot of HB (green), RP (blue) and GS (red) scores for Top-458 scientists within each ranking system for December 2015. The width of the columns represents the number of persons within each age group.  ARRmosagegr

This plot shows that the majority of younger extraordinary researchers belong to the HB

group. Among middle-aged ones the slight domination of GS over RP system is visible. At the same time the scientists of advanced age are mostly located in RP and partly in GS.

4.3 Research Fields

The field of research of every scientist provided by HB and GS enable us to enrich our data set by adding this factor and to perform a comparative analysis. From 500 researchers in HB only 448 persons have information about subject fields. This constrain forces to reduce the GS dataset, by taking 448 best ones, thus enabling the comparison. From RP we also select top 448 persons, though from merged data with GS or HB, see Figure 7. As a result, the RP scientists that originally do not have any information about areas of research receive these from their GS profiles or HB ranking system. Therefore, we receive a dataset that contains best 448 scientists within each of the discussed ranking system with their main research field.

In order to analyze the influence of research area on ranking scores, all researchers were divided into 19 groups of subject fields according to the well-known in economic sciences Journal of Economic Literature (JEL) classification, see JEL (accessed 22 Apr 2016). The explanation of JEL codes is given in Table 8 in Appendix.

| | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | Z | Total |
|----|----|---|----|----|----|----|----|----|---|----|---|----|----|---|----|---|----|---|---|-------|
| GS | 86 | 3 | 53 | 32 | 43 | 67 | 46 | 13 | 5 | 22 | 0 | 13 | 12 | 0 | 29 | 0 | 13 | 9 | 2 | 448 |
| HB | 1 | 2 | 49 | 73 | 49 | 39 | 59 | 1 | 6 | 10 | 3 | 48 | 67 | 1 | 24 | 0 | 8 | 4 | 4 | 448 |
| RP | 72 | 2 | 50 | 41 | 68 | 73 | 42 | 14 | 4 | 26 | 1 | 13 | 2 | 1 | 22 | 1 | 7 | 6 | 3 | 448 |

Table 3: Frequency Table for JEL codes and ranking scores of HB, RP and GS for Top-448 scientists within each ranking system for December 2015.

A distribution of scores of researchers within research areas (JEL codes) and corresponding ranking systems can be clearly seen on the comparative histograms in Figure 22. The generated from our data set frequency Table 3, as well as the corresponding percentage values on the Table 9, show that more than 16% of selected HB researches come from mi-

croeconomics. They are followed by the scientists from business field, financial economics, mathematical and quantitative methods, as well as macroeconomics and monetary economics, with over 10% within each research area.

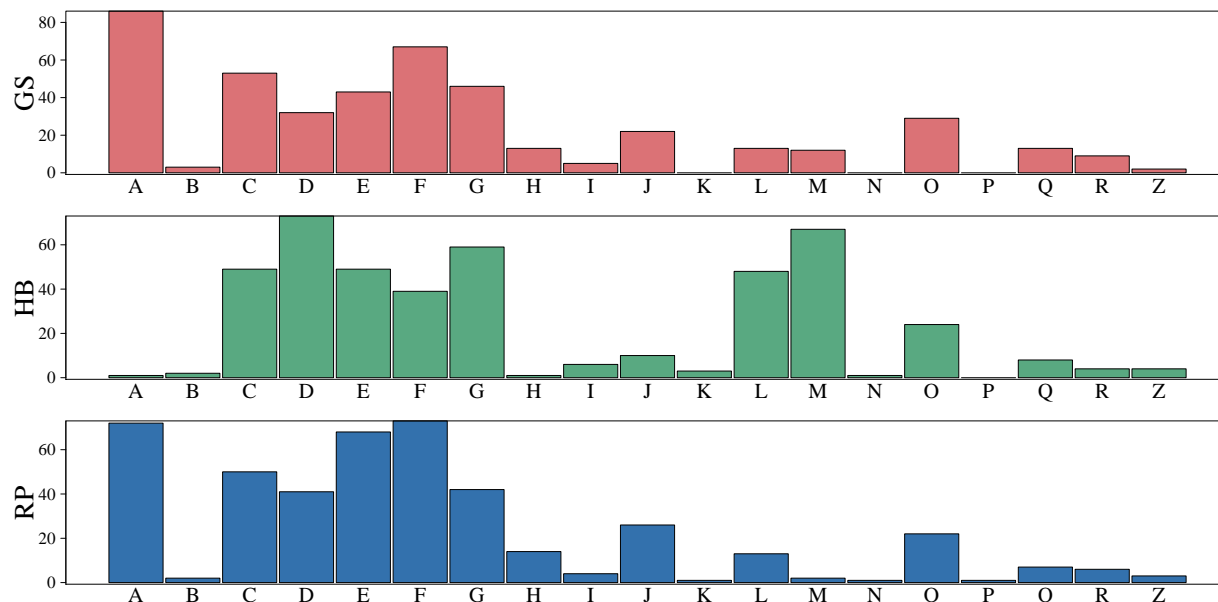



Figure 22: Comparative histogram for JEL codes and ranking scores of GS (upper/red), HB (middle/green) and RP (lower/blue) for Top-458 scientists within each ranking system for December 2015. 

A distinct difference is introduced by RP, where international economics, as well as general economics and teaching hold the leading positions with over 16% each one. These are accompanied by macroeconomics and monetary economics along with mathematical and quantitative methods with over 15% and 11% respectively. In the same manner the dominant research area of GS is presented by general economics and teaching with more than 19% of researchers. Besides, international economics produces above 14% of GS, while mathematical and quantitative methods, as well as financial economics, make 11% and 10% correspondingly. However, the mathematical and quantitative methods is the only research field among compared ones which has over 10% across all three ranking systems.

To sum up, we present a mosaic plot in Figure 23 that gives us an advantage of relative simultaneous comparison of ranking systems through the subject fields in a four dimen-

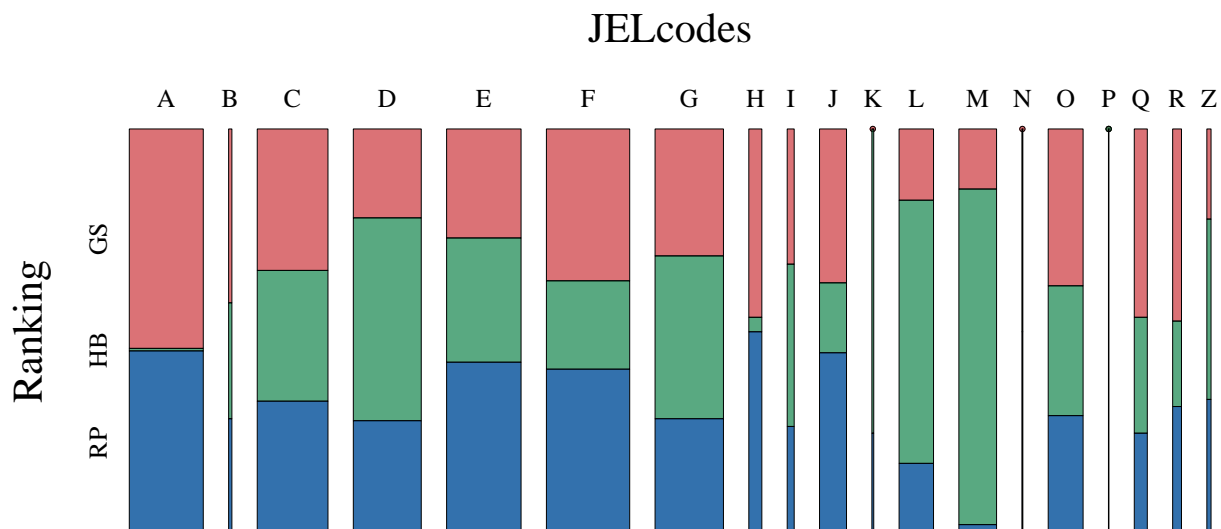



Figure 23: Mosaic plot of JEL codes and ranking scores of GS (upper), HB (middle) and RP (lower) for Top-458 scientists within each ranking system for December 2015. The width of the columns represents the number of persons within each research area and dots represent zero.  ARRmossub

sional space. Width of columns, illustrating the aggregated amount of people within each research area, brings us to the following important conclusions. Since F column is a widest one, the largest number of researchers possessing the leading positions among HB, RP and GS fulfill their research in international economics. Such fields as macroeconomics and monetary economics, general economics and teaching, mathematical and quantitative methods, microeconomics and financial economics illustrate a little difference. On the contrary the presence of scientists from economic systems, economic history, as well as law and economics in the top positions of discussed ranking systems is rather uncommon.

5 Conclusions

Summing up the results, it can be concluded that the quantile median regression model successfully predicts the missing ranking data in HB. It has been demonstrated that there exist different correlation structures between underlying sub-rankings of HB, RP and GS. Moreover, it was shown that there is a considerable improvement of young researchers within HB for the last several years.

The results obtained show that the academic ranking variation is sensitive to the age. Particularly, the rank of younger and advanced-age scientists is marginally changing (increasing) stronger than that of middle-age researchers. In addition, individuals from such research field as mathematical and quantitative methods possess the leading positions across all three discussed ranking systems. Whereas, persons from microeconomics, international economics, as well as general economics and teaching introduce the dominant part within HB, RP and GS respectively. Summarizing this contribution we have shown that the comparison of academic ranking scales reveals hidden information across and inside the ranking systems.

References

- Baum, C. (2013). Quantile regression, *Lecture notes on Applied Econometrics, Boston College*.
- Combes, P.-P. and Linnemer, L. (2010). Inferring Missing Citations - A Quantitative Multi-Criteria Ranking of all Journals in Economics, *HAL-SHS* (00520325).
- Dilger, A. and Müller, H. (2011). Ein Forschungsleistungsranking auf der Grundlage von Google Score, *Diskussionpapier des Instituts für Organisationsökonomik*.
- Fitzenberger, B. and Wilke, R. (2005). Using Quantile Regression for Duration Analysis, *ZEW, Mannheim* (Discussion Paper, No. 05-65).
- Forschungsmonitoring, P. (accessed 14 Oct 2015). Verein für Socialpolitik, <http://htmldb-hosting.net/pls/htmldb/f?p=193:4:3942442368081138::NO::..>
- GitHub (accessed 22 Apr 2016). ARR Project, <http://www.github.com/QuantLet/ARR>.
- Guo, M., Zhou, L., Huang, J. and Härdle, W. (2013). Functional data analysis of generalized regression quantiles, *Statistics and Computing* **25**(2): 189–202.
- Gygli, S., Müller-Dröge, H. and Schläpfer, J. (2014). HB BWL-Ranking 2014 - Methodik und Zeitschriftengewichte.
- Hamermesh, D. (2015). Citations in Economics: Measurement, Uses and Impacts, *NBER Working Paper Nov 2015*(21754).
- Härdle, W. and Simar, L. (2015). *Applied Multivariate Statistical Analysis*, 4th ed., Springer Verlag, Berlin Heidelberg, ISBN 978-3-662-45170-0, e-ISBN 978-3-662-45171-7 (580 p).
- Härdle, W. and Song, S. (2010). Confidence bands in quantile regression, *Econometric Theory* **26**(4): 1180–1200.
- Harzing, A. and Wal, R. (2008). Google Scholar as a new source for citation analysis, *Ethics in science and environmental politics* **8**: 61–73.

- JEL (accessed 22 Apr 2016). JEL (Journal of Economic Literature) Classification System, <https://www.aeaweb.org/econlit/jelCodes.php?view=jel>.
- Koenker, R. (2005). *Quantile Regression*, Quantile Regression, Econometric Society Monograph, Cambridge University Press, Cambridge, ISBN 0-521-84573-4, ISBN 0-521-60827-9.
- Koenker, R. (2015). Quantile Regression in R: A Vignette, *www.cran.r-project.org* (Version: March 25, 2015).
- Koenker, R. and Bassett, G. (1978). Regression Quantiles, *Econometrica* **46**(1): 33–50.
- Koenker, R. and Hallock, K. (2001). Quantile Regression, *Journal of Economic Perspectives* **15**(4): 143–156.
- Oberschelp, A. and Jaeger, M. (2010). Leistungsvergleiche als Instrument der Hochschulsteuerung: Ansätze, organisatorischer Kontext und Unterstützung des Steuerungshandelns, *Bibliotheksdienst* **49**(5): 475–494.
- QuantNet (accessed 05 Mai 2016). QuantNet, <http://quantlet.de/d3/ia/>.
- RePEc (accessed 22 Apr 2016). Research Papers in Economics, <http://repec.org>.
- Ricci, V. (2005). Fitting Distributions with R, *Release 0.4 – 21 February 2005*.
- Schläpfer, F. (2011). Reformbedarf bei der Rating-Agentur für Ökonomen, *Neue Züricher Zeitung* **198**: 23.
- Schläpfer, F. and Schneider, F. (2010). Messung der akademischen Forschungsleistung in den Wirtschaftswissenschaften: Reputation vs. Zitierhäufigkeiten, *Perspektiven der Wirtschaftspolitik* **11**(4): 325–339.
- Spokoiny, V., Wang, W. and Härdle, W. (2013). Local quantile regression, *Journal of Statistical Planning and Inference* **143**(7): 1109–1129.
- Wohlrabe, K. (2011). Das Handelsblatt- und das RePEc-Ranking im Vergleich, *ifo Schnelldienst* **17**: 66–71.

Wohlrabe, K. (2013). Einige Anmerkungen zum Handelsblatt-Ranking 2013, *ifo Schnelldienst* **23**: 79–83.

Zimmermann, C. (2013). Academic Rankings with RePEc, *Econometrics* **1**(3): 249–280.

6 Appendix

| | Criteria | Description |
|----|-------------------------------|--|
| | <i>Number of Works</i> | |
| 1 | NbWorks | Simple Count |
| 2 | DNbWorks | Count of distinct Works |
| 3 | ANbWorks | Count of distinct Works divided by number of authors on each work |
| 4 | ScWorks | Count with simple impact factor weights |
| 5 | AScWorks | Count with simple impact factor weights divided by number of authors on each work |
| 6 | WScWorks | Count with recursive impact factor weights |
| 7 | AWScWorks | Count with recursive impact factor weights divided by number of authors on each work |
| | <i>Citation Counts</i> | |
| 8 | NbCites | Simple citation count |
| 9 | ANbCites | Citation count divided by number of authors on each work |
| 10 | ScCites | Citation count with simple impact factor weights |
| 11 | AScCites | Citation count with simple impact factor weights divided by number of authors on each work |
| 12 | WScCites | Citation count with recursive impact factor weights |
| 13 | AWScCites | Citation count with recursive impact factor weights divided by number of authors on each work |
| 14 | DCites | Citation count discounted by age |
| 15 | ADCites | Citation count discounted by age and divided by number of authors on each work |
| 16 | DScCites | Citation count with discounted impact factor weights |
| 17 | ADScCites | Citation count with discounted impact factor weights divided by number of authors on each work |
| 18 | WDSccites | Citation count with recursive discounted impact factor weights |
| 19 | AWDSccites | Citation count with recursive discounted factor weights divided by number of authors on each work |
| 20 | HIndex | h-index |
| 21 | WIndex (-) | Wu-index |
| 22 | NCAuthors | Count of citing registered authors |
| 23 | RCAuthors | Rank weighted count of citing registered authors |
| | <i>Journal Page Counts</i> | |
| 24 | NbPages | Simple page count |
| 25 | ScPages | Page count divided by number of authors on each work |
| 26 | WSCPages | Page count with simple impact factor weights |
| 27 | ANbPages | Page count with simple impact factor weights divided by number of authors on each work |
| 28 | AScPages | Page count with recursive impact factor weights |
| 29 | AWScPages | Page count with recursive impact factor weights divided by number of authors on each work |
| | <i>Journal Page Counts</i> | |
| 30 | AbsViews | Total abstract views in the past 12 months |
| 31 | AAbsViews | Total abstract views per author in the past 12 months |
| 32 | Downloads | Total downloads in the past 12 months |
| 33 | ADownloads | Total downloads per author in the past 12 months |
| | <i>Co-Authorship Networks</i> | |
| 34 | Close | The average number of degrees of separation through co-authorship with all other registered authors |
| 35 | Betweenn | The frequency the author appears on the shortest path through co-authorship between any two other registered authors |
| | <i>Others</i> | |
| 36 | Students | Average rank across all other criteria for the set of graduates registered in RePEc and listed in RePEc Genealogy |
| 37 | NEPCites | Citation breadth across fields |

Table 4: Subrankings of RP. W-index is not taken into account while calculating the average rank score.

| | Count | Mean | St.dev | Median | Min | Max |
|--------------------|-------|---------|---------|--------|-----|----------|
| HB | | | | | | |
| VWL2015 LW | 250 | 7.6 | 3.8 | 6.2 | 4.3 | 29.7 |
| VWL2015 CR | 100 | 3.1 | 1.1 | 2.7 | 2.2 | 8.8 |
| VWL2015 U40 | 100 | 2.7 | 0.9 | 2.4 | 1.8 | 6.2 |
| BWL2014 LW | 250 | 7.4 | 3.6 | 6.4 | 4.3 | 33.8 |
| BWL2014 CR | 100 | 4.3 | 1.3 | 3.9 | 3.0 | 9.2 |
| BWL2014 U40 | 100 | 3.7 | 1.5 | 3.3 | 2.1 | 9.3 |
| RP | | | | | | |
| Average Rank Score | 2304 | 1107.0 | 631.7 | 1100.0 | 2.8 | 2194.0 |
| GS | | | | | | |
| Total Cites | 1438 | 10190.0 | 19831.2 | 5332.0 | 0.0 | 234200.0 |

Table 5: Descriptive statistics for HB, RP and GS values. Count means the number of observations, mean is the average of values, St.dev - standard deviation, max and min - maximum and minimum values.

| Code | Research field |
|------|---|
| A | General Economics and Teaching |
| B | History of Economic Thought, Methodology, and Heterodox Approaches |
| C | Mathematical and Quantitative Methods |
| D | Microeconomics |
| E | Macroeconomics and Monetary Economics |
| F | International Economics |
| G | Financial Economics |
| H | Public Economics |
| I | Health, Education, and Welfare |
| J | Labor and Demographic Economics |
| K | Law and Economics |
| L | Industrial Organization |
| M | Business Administration and Business Economics / Marketing / Accounting / Personnel Economics |
| N | Economic History |
| O | Economic Development, Innovation, Technological Change, and Growth |
| P | Economic Systems |
| Q | Agricultural and Natural Resource Economics / Environmental and Ecological Economics |
| R | Urban, Rural, Regional, Real Estate, and Transportation Economics |
| Y | Miscellaneous Categories |
| Z | Other Special Topics |

Table 8: JEL Classification System

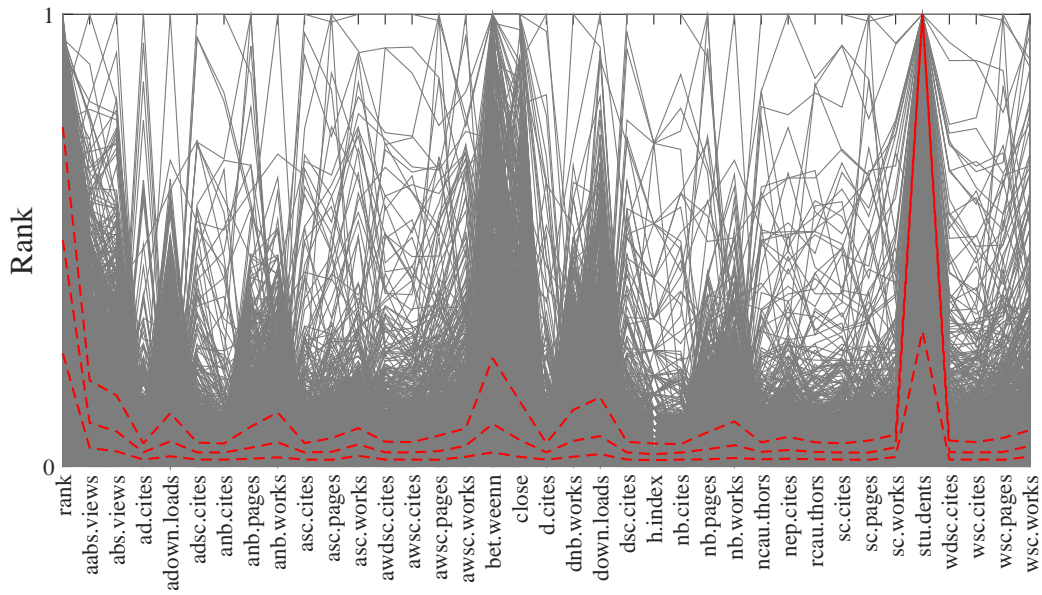



Figure 24: Parallel coordinates plot for RP average rank score and ranks from 36 sub-rankings for 2304 researchers in December 2015. The red dashed lines denote three quartiles (25%, 50% and 75%) that are shown for all individuals.  ARRpcprp

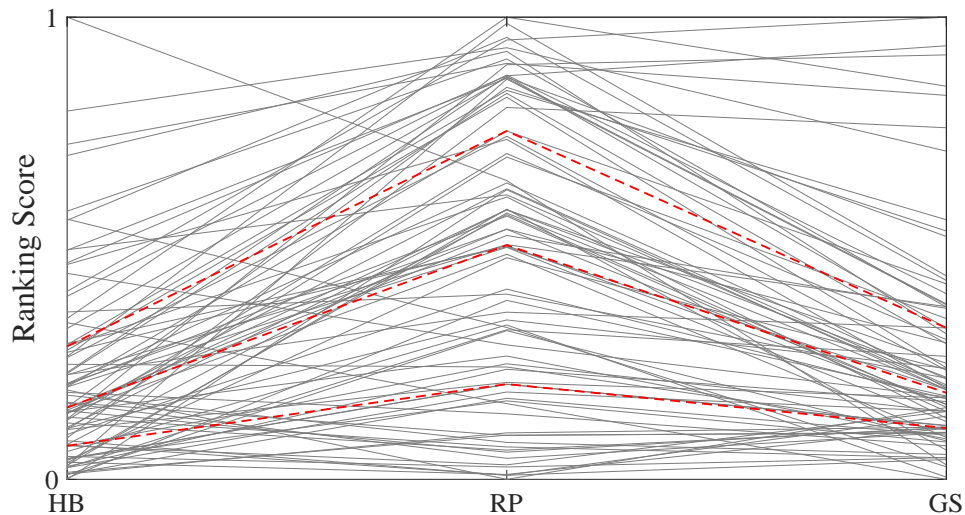


Figure 25: Parallel coordinate plot for three variables (HB, RP and GS) on 82 researchers. Two outliers from HB and GS are removed. Red lines denote the three quartiles (25%, 50% and 75%). RP values are rescaled.

| | Count | Mean | St.dev | Median | Min | Max |
|--------------------|-------|---------|---------|---------|--------|----------|
| HB | | | | | | |
| Age | 936 | 47.3 | 9.5 | 45.0 | 29.0 | 75.0 |
| Common Score | 500 | 7.6 | 3.8 | 6.4 | 4.3 | 35.8 |
| RP | | | | | | |
| Average Rank Score | 2304 | 1107.0 | 631.7 | 1100.0 | 2.8 | 2194.0 |
| Aabs-Views Score | 1435 | 2640.0 | 2544.7 | 1861.0 | 1052.0 | 36870.0 |
| Abs-Views Score | 1529 | 4447.0 | 3494.7 | 3323.0 | 1860.0 | 44760.0 |
| Ad-Cites Score | 1922 | 299.6 | 304.0 | 200.4 | 98.9 | 3378.0 |
| Adownloads Score | 1410 | 738.6 | 685.9 | 520.6 | 287.0 | 7766.0 |
| Adsc-Cites Score | 1874 | 852.7 | 880.0 | 570.2 | 244.3 | 10300.0 |
| Anb-Cites Score | 1936 | 1321.0 | 1432.8 | 856.7 | 404.5 | 16800.0 |
| Anb-Pages Score | 1415 | 877.2 | 430.8 | 754.3 | 463.5 | 4486.0 |
| Anb-Works Score | 1319 | 109.3 | 58.6 | 92.3 | 55.8 | 903.7 |
| Asc-Cites Score | 1890 | 13320.0 | 15007.4 | 8274.0 | 3405.0 | 162100.0 |
| Asc-Pages Score | 1680 | 13610.0 | 9677.1 | 10600.0 | 5414.0 | 115800.0 |
| Asc-Works Score | 1823 | 1381.0 | 1010.8 | 1046.0 | 555.8 | 10210.0 |
| Awsc-Cites Score | 1821 | 180.0 | 186.9 | 118.8 | 48.7 | 2081.0 |
| Awsc-Cites Score | 1835 | 685.0 | 785.3 | 420.1 | 162.3 | 8311.0 |
| Awsc-Pages Score | 1614 | 682.8 | 500.7 | 524.3 | 250.8 | 5334.0 |
| Awsc-Works Score | 1718 | 79.8 | 63.1 | 58.4 | 28.3 | 592.5 |
| Between Score | 1148 | 10.8 | 9.3 | 7.9 | 3.6 | 94.7 |
| Close Score | 1223 | 4.6 | 0.2 | 4.6 | 4.0 | 4.8 |
| D-Cites Score | 1889 | 500.8 | 494.8 | 342.8 | 162.5 | 5878.0 |
| Dnb-Works Score | 1343 | 128.5 | 66.1 | 111.0 | 68.0 | 1091.0 |
| Downloads Score | 1444 | 1273.0 | 992.3 | 950.0 | 511.0 | 10950.0 |
| Dsc-Cites Score | 1840 | 1444.0 | 1468.3 | 956.3 | 418.9 | 17640.0 |
| H-Index Score | 2017 | 19.4 | 7.4 | 17.0 | 12.0 | 78.0 |
| Nb-Cites Score | 1951 | 2113.0 | 2275.9 | 1385.0 | 640.0 | 29620.0 |
| Nb-Pages Score | 1521 | 1211.0 | 581.4 | 1046.0 | 658.0 | 6722.0 |
| Nb-Works Score | 1456 | 185.8 | 94.1 | 161.0 | 97.0 | 1288.0 |
| Ncauthors Score | 1898 | 1113.0 | 844.1 | 834.0 | 425.0 | 7787.0 |
| Nep-Cites Score | 1764 | 82.1 | 6.9 | 82.3 | 69.2 | 93.9 |
| Rcauthors Score | 1897 | 854.7 | 633.8 | 645.2 | 326.8 | 5722.0 |
| Sc-Cites Score | 1889 | 21610.0 | 24319.3 | 13500.0 | 5548.0 | 313000.0 |
| Sc-Pages Score | 1762 | 19410.0 | 13171.2 | 15450.0 | 8056.0 | 167500.0 |
| Sc-Works Score | 1884 | 2025.0 | 1402.1 | 1567.0 | 851.8 | 14870.0 |
| Students Score | 1093 | 814.1 | 575.2 | 711.2 | 4.3 | 2202.0 |
| Wdsc-Cites Score | 1787 | 306.5 | 313.3 | 201.6 | 83.1 | 3580.0 |
| Wsc-Cites Score | 1834 | 1114.0 | 1271.3 | 697.3 | 265.4 | 15220.0 |
| Wsc-Pages Score | 1681 | 980.2 | 678.3 | 782.2 | 377.1 | 7587.0 |
| Wsc-Works Score | 1791 | 116.9 | 90.2 | 87.7 | 43.8 | 1007.0 |
| GS | | | | | | |
| Total Cites | 1438 | 10190.0 | 19831.2 | 5332.0 | 0.0 | 234200.0 |
| H Index | 1438 | 32.9 | 20.2 | 29.0 | 0.0 | 177.0 |
| I Index | 1438 | 66.0 | 69.4 | 46.0 | 0.0 | 814.0 |

Table 6: Descriptive statistics for 42 factors of HB, RP and GS values. Count means the number of observations, mean is the average of values, St.dev - standard deviation, max and min - maximum and minimum values. 36

| | Count | Mean | St.dev | Median | Min | Max |
|-------|-------|---------|---------|---------|--------|----------|
| HB | | | | | | |
| <36 | 4 | 6.3 | 1.6 | 5.8 | 5.0 | 8.6 |
| 36-40 | 33 | 5.7 | 1.5 | 5.2 | 4.3 | 9.8 |
| 41-45 | 97 | 6.8 | 2.8 | 5.9 | 4.4 | 22.8 |
| 46-50 | 117 | 7.2 | 2.4 | 6.7 | 4.4 | 15.6 |
| 51-55 | 90 | 7.9 | 3.9 | 6.7 | 4.3 | 27.1 |
| 56-60 | 53 | 9.3 | 4.0 | 8.1 | 4.6 | 22.4 |
| 61-65 | 39 | 9.4 | 6.2 | 6.9 | 4.4 | 35.8 |
| 66-70 | 18 | 10.0 | 5.3 | 7.2 | 5.0 | 23.6 |
| >70 | 7 | 12.2 | 8.5 | 9.0 | 5.0 | 29.7 |
| RP | | | | | | |
| <36 | 1 | 341.8 | – | 341.8 | 341.8 | 341.8 |
| 36-40 | 2 | 372.4 | 40.8 | 372.4 | 343.6 | 401.3 |
| 41-45 | 15 | 276.7 | 117.8 | 306.1 | 89.7 | 473.3 |
| 46-50 | 30 | 291.2 | 140.0 | 304.8 | 5.2 | 479.7 |
| 51-55 | 72 | 291.8 | 123.0 | 305.0 | 2.8 | 479.5 |
| 56-60 | 94 | 247.1 | 142.5 | 240.1 | 11.4 | 487.5 |
| 61-65 | 90 | 205.7 | 137.9 | 184.4 | 12.7 | 475.2 |
| 66-70 | 66 | 219.5 | 129.3 | 211.1 | 9.0 | 452.8 |
| >70 | 88 | 214.8 | 147.2 | 189.1 | 3.4 | 489.0 |
| GS | | | | | | |
| <36 | 0 | – | – | – | – | – |
| 36-40 | 5 | 10240.0 | 1182.5 | 10840.0 | 8758.0 | 11470.0 |
| 41-45 | 26 | 12600.0 | 4745.0 | 11200.0 | 8075.0 | 28400.0 |
| 46-50 | 52 | 12860.0 | 5179.4 | 11070.0 | 7924.0 | 29670.0 |
| 51-55 | 86 | 18780.0 | 22906.8 | 13460.0 | 8012.0 | 212800.0 |
| 56-60 | 101 | 22640.0 | 20020.8 | 14340.0 | 7932.0 | 127300.0 |
| 61-65 | 74 | 25360.0 | 22591.8 | 17290.0 | 8190.0 | 161000.0 |
| 66-70 | 55 | 22680.0 | 17533.9 | 17740.0 | 7931.0 | 92730.0 |
| >70 | 59 | 51730.0 | 61926.0 | 20680.0 | 8022.0 | 234200.0 |

Table 7: Descriptive statistics for HB, RP and GS values through age groups. Count means the number of observations, mean is the average of values, St.dev - standard deviation, max and min - maximum and minimum values.

| | A | B | C | D | E | F | G | H | I | J |
|----|-------|-------|-------|-------|-------|-------|-------|------|------|--------|
| HB | 0.22 | 0.45 | 10.94 | 16.29 | 10.94 | 8.71 | 13.17 | 0.22 | 1.34 | 2.23 |
| RP | 16.07 | 0.45 | 11.16 | 9.15 | 15.18 | 16.29 | 9.38 | 3.12 | 0.89 | 5.80 |
| GS | 19.20 | 0.67 | 11.83 | 7.14 | 9.60 | 14.96 | 10.27 | 2.90 | 1.12 | 4.91 |
| | K | L | M | N | O | P | Q | R | Z | Total |
| HB | 0.67 | 10.71 | 14.96 | 0.22 | 5.36 | 0.00 | 1.79 | 0.89 | 0.89 | 100.00 |
| RP | 0.22 | 2.90 | 0.45 | 0.22 | 4.91 | 0.22 | 1.56 | 1.34 | 0.67 | 100.00 |
| GS | 0.00 | 2.90 | 2.68 | 0.00 | 6.47 | 0.00 | 2.90 | 2.01 | 0.45 | 100.00 |

Table 9: The percentage values for the frequency Table 3 for JEL codes and ranking scores of HB, RP and GS for Top-448 scientists within each ranking system for December 2015.

SFB 649 Discussion Paper Series 2016

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Downside risk and stock returns: An empirical analysis of the long-run and short-run dynamics from the G-7 Countries" by Cathy Yi-Hsuan Chen, Thomas C. Chiang and Wolfgang Karl Härdle, January 2016.
- 002 "Uncertainty and Employment Dynamics in the Euro Area and the US" by Aleksei Netsunajev and Katharina Glass, January 2016.
- 003 "College Admissions with Entrance Exams: Centralized versus Decentralized" by Isa E. Hafalir, Rustamdjan Hakimov, Dorothea Kübler and Morimitsu Kurino, January 2016.
- 004 "Leveraged ETF options implied volatility paradox: a statistical study" by Wolfgang Karl Härdle, Sergey Nasekin and Zhiwu Hong, February 2016.
- 005 "The German Labor Market Miracle, 2003 -2015: An Assessment" by Michael C. Burda, February 2016.
- 006 "What Derives the Bond Portfolio Value-at-Risk: Information Roles of Macroeconomic and Financial Stress Factors" by Anthony H. Tu and Cathy Yi-Hsuan Chen, February 2016.
- 007 "Budget-neutral fiscal rules targeting inflation differentials" by Maren Brede, February 2016.
- 008 "Measuring the benefit from reducing income inequality in terms of GDP" by Simon Voigts, February 2016.
- 009 "Solving DSGE Portfolio Choice Models with Asymmetric Countries" by Grzegorz R. Dlugosz, February 2016.
- 010 "No Role for the Hartz Reforms? Demand and Supply Factors in the German Labor Market, 1993-2014" by Michael C. Burda and Stefanie Seele, February 2016.
- 011 "Cognitive Load Increases Risk Aversion" by Holger Gerhardt, Guido P. Biele, Hauke R. Heekeren, and Harald Uhlig, March 2016.
- 012 "Neighborhood Effects in Wind Farm Performance: An Econometric Approach" by Matthias Ritter, Simone Pieralli and Martin Odening, March 2016.
- 013 "The importance of time-varying parameters in new Keynesian models with zero lower bound" by Julien Albertini and Hong Lan, March 2016.
- 014 "Aggregate Employment, Job Polarization and Inequalities: A Transatlantic Perspective" by Julien Albertini and Jean Olivier Hairault, March 2016.
- 015 "The Anchoring of Inflation Expectations in the Short and in the Long Run" by Dieter Nautz, Aleksei Netsunajev and Till Strohsal, March 2016.
- 016 "Irrational Exuberance and Herding in Financial Markets" by Christopher Boortz, March 2016.
- 017 "Calculating Joint Confidence Bands for Impulse Response Functions using Highest Density Regions" by Helmut Lütkepohl, Anna Staszewska-Bystrova and Peter Winker, March 2016.
- 018 "Factorisable Sparse Tail Event Curves with Expectiles" by Wolfgang K. Härdle, Chen Huang and Shih-Kang Chao, March 2016.
- 019 "International dynamics of inflation expectations" by Aleksei Netsunajev and Lars Winkelmann, May 2016.
- 020 "Academic Ranking Scales in Economics: Prediction and Imputation" by Alona Zharova, Andrija Mihoci and Wolfgang Karl Härdle, May 2016.

SFB 649, Spandauer Straße 1, D-10178 Berlin
<http://sfb649.wiwi.hu-berlin.de>

This research was supported by the Deutsche
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

