

Zhu, Junyi; Steiner, Viktor

Conference Paper

A Joint Top Income and Wealth Distribution

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Zhu, Junyi; Steiner, Viktor (2020) : A Joint Top Income and Wealth Distribution, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/224651>

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.

A Joint Top Income and Wealth Distribution¹

Viktor Steiner

Junyi Zhu

Freie Universität Berlin

Deutsche Bundesbank

Preliminary draft – 2019.05.30

Abstract

Top distributions of income and wealth are still incompletely measured in many national statistics, particularly when using survey data. This paper develops the technique of incorporating the joint distributional relationship to enhance the estimation of these two top distributions. We leverage the bivariate parametric/non-parametric copula to extrapolate both income and wealth distributions from German PHF (Panel on Household Finance) data. The copula modelling potentially reduces the ad hocery in choosing the estimation domain as well as in the parametric specification (eg Pareto family) imposed by almost all the marginal approaches. One distinct feature of our paper is to complement the model fit with external validation. The copula estimate can help us to perform out-of-sample prediction on the very top of the tail distribution from one margin conditional on the characteristics of the other. The validation exercises show that our copula-based approach can approximate much closer to the top tax data and wealth “rich list” than those unconditional marginal extrapolations. The properness of copula and conditioning criterion seems to convince the asymmetric joint association between (labor) income and wealth (capital income) distributions as recently evidenced by other countries.

Keywords: income and wealth joint distribution, copula, heavy-tailed distributions, external consistency

JEL codes: C52, C46, C81, D31

¹ Results and opinions expressed in this paper are those of the authors and do not necessarily reflect views of the Deutsche Bundesbank. Email: [mailto: viktor.steiner@fu-berlin.de](mailto:viktor.steiner@fu-berlin.de) and junyi.zhu@bundesbank.de.

1. Introduction

The analysis of top income and wealth distributions on the basis of survey data is problematic due to severe under-coverage and measurement error. Respondents either under-report the figures or there is top-coding. People at the very top of the income and wealth distributions usually do not participate in such surveys and the survey administrators often do not have reliable information on the true statistical characteristics of non-respondents. Hence, most income and wealth surveys only sparsely cover the top given the nature of long tails. Researchers tackle this issue by extrapolation via conjecturing the wide observance of the univariate power-law distributions, in particular from the Pareto family (see, eg Vermeulen 2014, Eckerstorfer, Halak, Kapeller, Schütz, Springholz and Wildauer, 2015 and Jenkins, 2016).

While most of the empirical applications refer to the estimation of either the top income or wealth distribution, the analysis of the joint top income and wealth distributions has not been the focus of much research so far (see, eg Aaberge, Atkinson and Königs, 2018). In this paper we contribute to this evolving literature by introduce the information from the top joint distribution between income and wealth. By postulating that the under-covered top bivarite distribution follows the same structure (copula) as those observed in the data as well as a non-response or truncation model for the top, we extrapolate not only the marginal distributions but also the joint one. There are two main contributions. First, we apply the statistical concept of copulas to estimate the joint income and wealth distribution and compare the estimates from the copula-based (conditional) distributions to those from marginal (unconditional) Pareto distributions. One advantage of copula modelling is that there is no need to assume specific marginal distributions or to restrict the estimation domain to specific ranges, as it is the case for the Pareto family of distributions. Second, we use external information on the top income and wealth distributions to validate our empirical estimates of the joint income-wealth distribution. For this purpose, our benchmarks are the administrative income tax data (eg top 1000 percentiles within top 1%) and the list of the 300 wealthiest people in Germany ("rich list") from the Manager Magazin.

We compare copula-based and marginally extrapolated top distributions with the observed real data. First, we illustrate the mechanism on how the copula-based approach can differ from the marginal one. The affecting channels include the copula and conditioning characteristics from the margin different from the one to be extrapolated. Then, we explore the potential rationale on improving the marginal extrapolation by incorporating the joint association. Recent literature discusses such an asymmetry in the

joint income and wealth distribution: (top) income inequality drives (top) wealth inequality but the association from the other direction is weak. Specifically, high (labor) income earners are also rewarded by large capital return and many wealth-rich families are still living as rentiers. Using this feature, we justify the copula and conditioning characteristics chosen. Finally, the comparison is accomplished among candidate copulas and conditioning characteristics to validate the justification above.

By anchoring our exercise with the findings of economic relationship between these two dimensions (from other countries), there are three benefits: first, the results confirm the validity of introducing the joint association is not an arbitrary finding, which may approximate the true data generating process of single dimension better than marginal parameterization; second, we can discipline the search of approximating copula and conditioning characteristics in the future work and similar efforts on other data and countries (eg by ruling out the estimated copula and conditioning characteristics, possibly due to statistical fluctuation, incompatible with the economic relationship); third, we may explore the potential contribution to survey design (eg oversampling mechanism and trade-off between reaching the very top and minimally accessing the other correlated economic variables).

One problem with the use of the Pareto type I distribution in the literature is the strong dependence of the estimated Pareto parameter on the threshold value chosen. Given the relatively small group of people with high values of income or wealth observed in survey data, this threshold parameter is set relatively low, which is contradicted with estimates from tax data that includes also rich people. One approach to detect the threshold is to examine the inverted Pareto coefficient.² For type I Pareto model, inverted Pareto coefficient has a one-one mapping to the Pareto coefficient. Thus, its inverted Pareto curve is flat. Bach, Corneo and Steiner (2012) plots using German tax data (2005) the empirical inverted Pareto coefficient (see Figure 1). It seems the flattening starts at least at around 100,000 euros.³ This level is somewhere between P95 and P96 in the German gross household income distribution above which the survey data (eg GSOEP) covers little.⁴ However, studies using Pareto type I distribution usually adopt much lower thresholds.⁵

This problem may be less severe in studies that use the more flexible (Generalized) Pareto type II distributions that allows non-constant inverted Pareto coefficient (see

² Inverted Pareto coefficient is the ratio of the mean in the distribution above certain threshold and the threshold itself (see, eg Blanchet, Fournier and Piketty, 2017).

³ There may be a flattening for the wage and professional income above this threshold. The flattening in the other two income concepts is actually less obvious and the starting points are much higher, if at all.

⁴ See Bach et al. (2012).

⁵ See, eg Bartels and Metzger (2017; Appendix Table A.1).

Blanchet et al., 2017). Jenkins (2017) discusses the comparison between type I and II Pareto models and how to determine the optimal threshold in estimating Pareto parameters. Also, he suggests the alternative estimation approaches regarding the choice of survey and external validation data. As claimed by him using UK income data, (Generalized) Pareto type II distribution outperforms the Pareto I distribution in terms of goodness of fit.⁶ He also shows the evidence that the choice of optimal threshold for estimating the Pareto I model is not clear: estimates are sensitive to the choice of threshold and optimal threshold in the type I model has more variability across the years. The optimal threshold estimated for type II model can be much lower than that for type I model. This feature is particularly attractive for our application. It would be preferred to have larger sample when fitting two dimensional model than univariate one.

Atkinson (2017) shows Pareto coefficient in type I model might not be constant over the top distribution even for a cross section using UK historical data.⁷ Blanchet et al. (2017) follow Atkinson (2017) and show that the inverted Pareto coefficient converges from below when the percentile rank of income distribution is near one using US and French data. Using tax returns data, Bach et al. (2012) also show for Germany that the inverted Pareto coefficient looks more like U shaped than a flat curve and converges from below when income grows to the very high end. Based on these arguments, we estimate the Pareto II distribution for the use of both marginal and copula-based extrapolations.

On the other hand, the rank-rank association (copula) of income and wealth dimensions in a society may be rather rigid: the copula structure is probably quite stable regarding the ranks with respect to the richest individual (the king) in the society (or subsociety) no matter where we define the poorest observation in terms of income and wealth. As a result, introducing copula in the extrapolation can alleviate the concern on the threshold to pin down the fitting sample.

In section 2 we present our data as well as external information on top tails. We outline the methodology in section 3 concerning Pareto distributions, copula and difference between marginal and copula-based extrapolations. Section 4 outlines the estimation of

⁶ Jenkins (2017) calls the Generalized Pareto distribution $f(x) = \frac{1}{\sigma} \left[1 + \frac{\xi(x-\mu)}{\sigma} \right]^{-\frac{1}{\xi+1}}$ Pareto II (Lomax distribution) which results from the former under the restriction $\mu=0$. We follow his notation throughout the paper.

⁷ Fig 1 in his paper describes a vivid definition of the profile of Pareto coefficient in terms of *Baronial* or *regal* regimes: "...They are designed to recall the difference between the situation (as Pareto coefficient increases as rise in rank) where a monarch was surrounded by powerful barons whose resources were not dissimilar in scale, and a situation (as Pareto coefficient decreases as rise in rank) where the monarch had—for example by appropriating the income of the church or seizing mineral wealth—raced ahead...."

both top marginal and joint distributions from the survey data. Section 5 presents the external validation. Section 6 concludes the paper.

2. Data

Our estimation comes from the Panel on Household Finance (PHF), the German component of the Household and Consumption Survey (HFCS) of the European Central Bank. We pool income and wealth data for the two panel waves of 2009 and 2013 in order to increase sample size and the precision of estimates for the top of the respective distributions. Since PHF is multiply imputed, we average over five imputates to form the data used for the results presented. The top wealth tail distribution is drawn from the rich list collected by the Manger Magazin. Dalitz (2016), Vermeulen (2014) and Bach, Thiemann and Zucco (2015) have described both data sets in detail. The top income distribution is retrieved from the administrative tax data available from the Research Data Center of the Federal Statistical Office of Germany.⁸ This data allows building the distribution of gross income for all the tax units (spouse or single) who file the tax return and someone who do not file the tax return. We construct the same concept of gross income for the tax units from the PHF.

Table 1 demonstrates the comparison of the top 30% distributions of gross income for tax units from PHF (wave 1 - 2009) and income tax statistics (Est) 2008. Besides showing the percentile means, the last column provides the relative difference for each percentile. The PHF distribution can match with the Est one quite well, particularly in the top 10% distribution with the relative differences between two percentile means all below 5% except the top 1%.

3. Methodology

3.1 Power law and Pareto distributions

3.2 Copula

3.3 The relationship between copula-based and marginal extrapolations

The conditional distribution depends on the marginal distributions and copula specified as well as the conditioning characteristics. For simplicity and constrained by the limited

⁸ See <https://www.forschungsdatenzentrum.de/de/steuern/lest>.

information we have on the opposite dimension of the extrapolated subpopulation, we start from postulating a lower bound on the other margin as the conditioning characteristic. Namely, we impose the lowest income level of all the observations in the top rich list and the lowest wealth level for all the tax payers in the top income tax data.

Formally, let the copula be $C(u, v)$, where $[u, v]$ is a two dimensional random vector on the unit cube $[u, v]^2$. Then the copula-based probability density function of wealth w , conditional on the income y all above a lower bound y_b , $P(w|y > y_b) = \frac{P(w, y > y_b)}{P(y > y_b)} = \frac{P(w)P(y > y_b|w)}{P(y > y_b)} = \frac{P(w)P(U > u|V = v)}{P(U > u)} = \frac{P(w)}{1-u} \int_u^1 c(x, v) dx = \frac{P(w)}{1-u} \left[\frac{\partial}{\partial v} C(1, v) - \frac{\partial}{\partial v} C(u, v) \right] = P(w) \frac{1 - \frac{\partial}{\partial v} C(u, v)}{1-u}$, where U and V are the cumulative distributions of income and wealth respectively, $u = P(y \leq y_b)$, $v = P(W \leq w)$ and $c(u, v)$ is the copula density.⁹ u and v are calculated using the estimated marginal (Pareto) distributions.

As observed, the difference between the conditional probability $P(w|y > y_b)$ and unconditional one $P(w)$ depends on both $C(u, v)$ and u , ie the underlying copula and marginal distribution of the top wealth holders. In general, we consider that the top wealth or income list to be extrapolated and benchmarked with is a subsample of the whole population with income or wealth all above some lower bound.¹⁰ The alternative assumption is that there is a latent wealth random variable following the unconditional distribution and which is fully observable for the not so top part in the survey data but the top distribution of the population obeys the data generating process described by our conditional probability – ie social or economic rules preclude the top wealth holders/income earners with little income/wealth. Namely, the first assumes a (non) response rule and the second proposes a statistical model of data generation (truncation). Moreover, we assume that the copula $C(u, v)$ for the top distribution can be approximated by those estimated from some observed sample in the survey data (ie a parallel rank association assumption – copula structure is invariant in the data with or without top tails).

⁹ The corresponding conditional cumulative density function $P(V \leq v|U > u) = \frac{1 - ((1-v) + C(u, v))}{1-u} = \frac{v - C(u, v)}{1-u}$.

¹⁰ Given the construction of rich list and income tax data, it appears to be reasonable to treat our top lists as sampled from the whole population (eg due to obstacle and incomplete information of reaching all the richest population by a voluntary (journalism) effort, and tax evasion and institutional barrier to cover all the top income earners in the tax data).

3.4 Discovering the copula and conditioning characteristics

We can guide the choice of copula and conditioning criterion by empirical or theoretical evidence of the joint association between income and wealth. The properness of copula and conditioning criterion is postulated to comply with one feature of the joint association between (labor) income and wealth (capital income) recently discovered in quite a few of countries (see Aaberge et al., 2018, for US and Norway, Bach, Calvet and Sodini, 2017, for Sweden and Lieberknecht and Vermeulen, 2018, for US and France): (top) income inequality drives (top) wealth inequality but the association from the other direction is weak.¹¹ The main channel is the capital income: high (labor) income earners are also rewarded by high capital return and accumulate more wealth while many wealth-rich families are still living as rentiers. Therefore, the joint association of these bi-dimensional ranks should be more symmetrical for the (labor) income rich subpopulation and more asymmetrical for the wealth rich subpopulation.¹² In practice, if, a priori, we believe, for instance, the extrapolated tail for wealth is not thick enough under marginal approach and trust our data generating process postulated for the very top (ie the non-response /truncation mechanism in the copula-based approach) as well as the conditioning lower bound, it is more plausible that the (latent) copula is a symmetric than asymmetric one so that restricting the income of top wealth holders to be high enough can lead to more inequality in the top of wealth distribution. As far as the conditioning lower bound is concerned, a rather thick tail in the top wealth distribution implies that the lower bound for the income level among the wealthy in the rich list should also be high. On the other hand, the minimum wealth among the top income earners in the tax data might not be necessarily high if there is also a thick tail in the very top of income distribution. The below comparison exercises demonstrate the validity of these proposals.

¹¹ Bach et al. (2017) does not discuss the reverse link from wealth to income inequality. The copula contour curves shown before confirm this pattern: the very top contour curves are close to the negatively sloped straight lines (ie the negative diagonal pattern to the transition matrix, as with the class society, described in Aaberge et al (2013)) and those in the lower part look almost like the L-shaped lines, ie the association of the rank in both dimensions is little.

¹² Since the introduction of flat rate tax (Abgeltungsteuer) on capital income in 2009, the tax data in Germany does not cover capital income since then (ie the top 1% distribution from 2010 data we use). PHF also does not collect the capital income well. The aggregate coverage compared to national account is about 60%. Moreover, the share of capital income is about 20 to 30% in Germany. The distribution of labor income should dominate in our PHF total income measure. In the future, to be more clearly benchmarked, we should use only the labor income from PHF for extrapolation contrasting with the top 1% from 2010 income tax data or we can use the top 1% from 2007 income tax data when capital income is still covered (note the Est data is not available for the years between 2007 and 2010).

4. Estimating top marginal and joint distributions

We describe the fitting of Pareto II distributions on both top income and wealth distributions as well as the copula estimation.

4.1 Top marginal distributions

There are various procedures to determine the threshold value above which the Pareto coefficient can be assumed to stay fixed (see, eg Jenkins, 2017). We adopt the “more objective” approach to determine the optimal threshold: measuring distance between estimated parametric and observed distributions using the Kolmogorov-Smirnov (KS) statistic, ie the maximum distance between the empirical and simulated cumulative distribution function implied by the estimated Pareto distribution. The optimal one is supposed to minimize this distance. Alternatively, an intuitive approach is to plot the estimated parameters against thresholds and pick the one above which the estimate is flat. Besides the threshold where the model starts to apply, the Pareto II model has the other two parameters, a shape parameter ξ and the scale parameter σ , whose estimates are theoretically constant above the assumed threshold. Jenkins (2017) suggests that for UK data potential values of the threshold may range between the 95 to the 99 percentile of the income distribution.

Figure 2 plots the KS measure for the income observed for PHF between 35,000 and 290,000 euros, which is far below p90 (as a robustness check) and somewhat larger than p99.5. P95 to p99.5 is the common searching range and the optimal thresholds for UK income distribution lie within this range (Jenkins, 2017). The author also suggests this finding might not be universally applicable in other countries and variables.¹⁵ The minimum is reached at somewhere below or at p90. Figure 3 provides the curves of estimates against threshold. It justifies our choice by showing that the above area is most stable. The exact minimizer of KS is 70,271 euros. However, the inverted Pareto coefficient b based on the estimates using this threshold converges from above as rank p approaches to 100%.¹⁶ This also goes against the Figure 1. Our favoured threshold is 86,957 euros which is 7th smallest thresholds and has the KS measure being 0.0135. The KS measure for 70,271 is 0.0131.¹⁷

¹⁵ If our threshold searched becomes low and let our fitting sample to include some non-power-law distribution, KS statistic will become large than the optimal (minimum) one (See p.671 of Clauset, Shalizi and Newman (2009); generally, we should expect to see the U curves in Figure 2 and Figure 6 in our paper if we search the whole sample).

¹⁶ Blanchet et al. (2017) provides the formula for $b(p)$.

¹⁷ The 6th smallest threshold is 70,400 euros. And all the first six thresholds together with their parameter estimates can only deliver the downward converging inverted Pareto coefficient.

Figure 4 and Figure 5 are the inverted Pareto coefficients by income and probability for the top income when the Pareto II follows the most favoured estimate: threshold=86,957, shape=.42 and scale=34,241. They both converge from below which implies the tail is getting fatter as income becomes higher. It aligns with the behaviour of $b(p)$ in US, France and UK for the recent decades (see Fig 1 in Blanchet et al., 2017, and the discussion in Atkinson, 2017). Most interestingly, comparing with Figure 1, we can observe our inverted Pareto coefficient curve is very close to the grey counterpart (ie the coefficient for the wage and professional income in the income tax data) which lies between around 1.68 and 1.72 and converges at around 1.72. PHF does cover the distribution and aggregate of labor income very well when benchmarking with the income tax data and national account aggregate (coverage rate of 98%). But the coverage of capital income is just about 60%.¹⁸ It is reasonable to postulate that our gross income is closest to the wage and professional income concept in the tax data. Given these facts, this estimate of Pareto II distribution looks convincingly to be close to the true distribution.

Figure 6 plots the KS measure for the wealth observed for PHF between 100,000 and 5,000,000 euros, which is again far below p90 and somewhat larger than p99.5. The minimum 245,160 euros is located between somewhere around the half of p90 and somewhere a bit higher than p90. Figure 7 provides the estimated curves against the thresholds. For the modified sigma, this area does look most stable. For ξ , the most stable area lies between p90 and p99. But these estimates, at around 0.6, are close to that obtained when threshold is at about the half of p90. Given these results, we decide to take this exact minimizer 245,160 euros which can offer us more sample size in considering fitting a top joint distribution.¹⁹

Figure 8 and Figure 9 are the inverted Pareto coefficients by income and probability for the top wealth when the Pareto II follows the most favoured estimate: threshold=245,160, shape=.6 and scale=197,206. They both converge from above which implies the tail is getting thinner as wealth becomes higher. We do not have direct evidence from other countries or Germany to verify this profile. However, the Pareto indices (alpha in Pareto I distribution) are 1.52 for the Manager Magazin rich list and 1.47 for PHF when the most favoured MLE estimation is adopted for these two data

¹⁸ See Zhu (2014).

¹⁹ This argument applies when we perform the *net contribution benchmark* in the external validation which requires the same fitting sample for marginal estimation and estimating the copula. But as discussed below, we are actually looking for a copula approximate the true one in the very top (ie that with the margins covering top income tax data or wealth rich list). The most approximating copula does not have to come from the same sample for fitting marginal distribution. This identical sample restricted by the *net contribution benchmark* is simply one of the competing “approximate” samples.

separately (see the very first row in Table 3 of Dalitz, 2016). It implies the inequality is not so high in the top wealth distribution as that below the very top in the rich list for Germany. This is compatible with the profile of $b(p)$ for the most favoured estimate.

4.2 Top joint distributions

While estimating the parametric copula for the fitting samples proposed, we use `BiCopSelect()` from the `VineCopula` in R to choose the best copula using weights. The inputting marginal probabilities are either empirical one or parametrically specified. It fits a comprehensive set of copulas via maximum likelihood estimation. Then the Akaike Information Criteria are computed for all of them. The one with the minimum value is selected.²⁰ To check the goodness-of-fit, we compare the contour curves from both the empirical copula (from the fitting sample) with those from the estimated one. When the margins adopt the empirical distribution, the contour curves are compared in the unit square. When the margins are parametrically specified, they are plotted in the bivariate supports of fitting sample. For the description of each R packages (ie `VinceCopula` and `Copula`), we suggest to read Brechmann and Schepsmeier (2013), Schepsmeier, Stoeber, Brechmann, Graeler, Nagler, Erhardt and Killiches (2017) and Hofert, Kojadinovic, Maechler and Yan (2017). A comprehensive theoretical and computational coverage for the `Copula` package as well as copula modelling itself is available in Yan (2007) and Kojadinovic and Yan (2010).

5. External validity

5.1 Competing copula and marginal distributions

One straightforward standard to compare the copula-based and marginal approaches is to benchmark based on the most approximating specifications for both (*approximating copula benchmark*). The a priori candidate is the copula for the *full sample* suppose that the (social) rank relative to the top rich on these two dimensions are so steadfastly configured which is invariant with the position to define the very bottom (eg in our case, either the extreme poor in the survey data or the left end of top rich list's / top income earners' distribution). For the pedagogic assessment of the plausibility between symmetric and asymmetric copulas discussed before, we also supply a *high income* subsamples (all the observations with income higher than some threshold) and a *high wealth* subsample (all the observations with wealth higher than some threshold). The

²⁰ The package also allows using Bayesian Information Criteria when the models with higher number of parameters are to be penalized.

copula of the former should be more symmetric than that of the latter. Given the thresholds (parameters) estimated for our Pareto II marginal distributions, it is natural to use them as cut-off to form high income or high wealth subsamples.

We are comparing the extrapolation accuracy of marginal and copula-based approach. To gauge the net contribution of copula, we should estimate the copula from a sample whose marginal distributions follow the most favoured specifications proposed above. The natural strategy to locate such a fitting sample is to simultaneously apply the cutoffs from these two favoured marginal (Pareto II) specifications (*net contribution benchmark*). As described below, the copula is estimated based on the marginal cumulative distribution parametrically estimated instead of the empirical one as used in the previous cases under *approximating coupla benchmark*. This contending subsample (*high income – high wealth*) for copula estimation indeed yields the (almost) same marginal estimates since the degree of asymmetry is still not strong at the level of these two cutoffs.

Table 2 summaries these four cases as well as the accompanying marginal distribution specification. For each case, our exercise compares the distribution distance between the extrapolated targets – top rich list or top distribution from income tax data – and marginal and copula-based extrapolations. The extrapolated distribution is produced by inputting the exact income/wealth values from each point of these two top lists into the specification of either approach. We also show the comparison while adjusting the lower bound to be conditional on in order to test the symmetricity of joint association as argued above.

The copula is estimated using the weighted empirical cdfs under the cases of high income, high wealth and full sample. However, we adopt the two-stage estimation method called inference functions for margins (IFM) in high income-high wealth case. The IFM method firstly estimates the parameters of marginal distributions via a MLE and then estimates the copula parameters via a separate MLE using the parametric cdfs calculated based on the estimates in the first step. The second step is same as the semi-empirical estimation method used for the first three cases except the formation of marginal cdfs (See Yan, 2007, and Joe and Xu, 1996).

The copulas estimated in the samples defined in Table 2 are either (Student) t or BB8. BB8 copula is also named as Joe-Frank copula which is a mixture copula between single Joe and Frank copulas. Student t copula has high tail dependence on both ends near $[0, 0]$ and $[1, 1]$ and is symmetrical while BB8 (Joe-Frank) copula has zero tail dependence on both ends and is asymmetrical. As the degree-of-freedom parameter (ie

the 2nd parameter illustrated in Table 2) in t copula increases, the tail dependence drops.²⁶

The best copula chosen by BiCopSelect() is t for high income subsample and BB8 for high wealth. This outcome corroborates also in Germany the symmetricity/asymmetricity evidence on the joint association of these two dimensions as discussed previously.²⁷ Thus, the best fit is the t copula with heavy tail dependence. In contrast, the tail dependence disappears among the high wealth subpopulation because of wealthy rentiers. This is captured by the BB8 copula.

To illustrate the goodness-of-fit, we plot the contours of the empirical copula and fitted copulas together for each of these four cases in Figure 10, Figure 11, Figure 12 and Figure 13.²⁸ All the contour lines from both copulas seem to lie close to each other. Our model fits the data very well. Figure 14, Figure 15, Figure 16 and Figure 17 present the density of the fitted copula for these four cases. Overall, except the area around two extreme ends [0, 0] and [1, 1], the surfaces are upheaved in the direction of diagonal between [0, 0] and [1, 1] on the unit square (percentile ranks) of income and wealth with the densities sinking on two ends of the reverse diagonal between [0, 1] and [1, 0]. It is a sign of positive rank correlation between these two dimensions for the most part of the society except the very poor and very rich. Student-t copulas in (Figure 14, Figure 16 and Figure 17) are all featured with two spikes in the extreme ends [0, 0] and [1, 1] which is a sign of strong tail dependence. This pattern also confirms the inverse correlation between the tail dependence and the degree-of-freedom (2nd) parameter in t copula: the spikes in these two extreme ends shoot up most strongly when this parameter is smallest (ie. in the full sample case in Figure 16). BB8 copula does not have a mass in the lower end [0, 0] in Figure 16 since it does not have the tail dependence.²⁹

5.2 Comparison results

The competing specification is benchmarked with the real distribution of extrapolation targets (top rich list or top income tax list) through the P-P plot. One advantage of this distribution comparison tool is to minimize the disruption by skewness. The plots draw

²⁶ See Joe (1997), Brechmann et al. (2013, p.7-8), Buike (2018, p.6) and Trivedi, Pravin and Zimmer (2007, p.26).

²⁷ Namely, among the high income subpopulation, extreme income earners should be very likely to build up excessively high wealth. High wealth holders might not be income rich.

²⁸ As noted before, the first three cases use the unit square as margins since the inputs for copula estimation are the marginal CDFs. The last case is plotted on the support spaces of income and wealth distributions since they are directly the inputs for (IFM) estimation.

²⁹ But it has a mass/spike in the [1, 1] end. The spike is only a necessary condition of tail dependence.

the empirical cumulative distributions of the top lists against the extrapolated one at each of the observations in the top lists. The former lies on the x-axis and the latter on the y-axis.³⁰ In our setting, we plot two extrapolated distributions (distinguished by 1 and 2) and each one contains both the marginal and copula-based approaches as specified in Table 2. They are denoted as *marginal 1*, *copula 1*, *marginal 2* and *copula 2* respectively. This allows us to compare not only between marginal and copula-based approaches within cases (specifications) but also these approaches between cases. Moreover, we also vary the conditioning income or wealth lower bound to illustrate this additional channel in the copula-based approach. Our naïve lower bounds are the maximum values in PHF, ie 76.3 million euros for wealth and 2.9707 million euros for income.³¹ The exact adjustments of the conditioning bounds in the following examples are on an ad hoc basis for the current stage. Only the relative position (ie higher or lower) between contrasting settings is emphasized. We present these specifications under each p-p plot and particularly highlight the contrasting components in italics.³² Except the contrasting pairs, all the remaining components are kept the same between case 1 and 2.

5.2.1 Between copulas

In Figure 18, we compare the extrapolation on wealth rich list between the high income and high wealth cases where only the copulas (between t and BB8) are contrasted. The same is done on the top tax data in Figure 19. Copula-based extrapolation using high income (t) copula beats high wealth (BB8) copula for the top wealth distribution and it is the opposite for the top income extrapolation. High income (t) copula bears strong tail dependence which enhances the very top tail by restricting the other dimension to be high enough too and drives the extrapolated proportion curve to converge from above to the 45 degree line (ie the observed one).³³ High wealth (BB8) copula has zero tail dependence so that the extrapolated tail is still not thick enough as shown in Figure

³⁰ In the next stage, we can introduce the empirical / non-parametric copula instead.

³¹ This assumes the minimum wealth of all the top 1% income distribution (as in the tax data) is at least 76.3 million euros and the minimum income of all the observations in the wealth rich list is at least 2.9707 million euros. The real bound for the former should be lower than 76.3 million since PHF does cover some distribution above top 1% percentile. And the real bound for the latter should be higher than 2.9707 million euros since the maximum wealth in PHF is still much below the minimum in the rich list.

³² To be brief, we only highlight one occurrence of the contrasting components. For example, in Figure 18, only the copula specification in copula 1, *Student t (parameter 1=0.334, parameter 2=7.12)*, is italicized. It is contrasted with the copula specification in copula 2, BB8 (parameter 1=1.64, parameter 2=0.943).

³³ Since the bottom in the rich list is far above the Pareto II threshold we used (245,160), the left end of the distribution in the rich list is farther away from the bottom end [0, 0] of the copula structure imposed. Thus the similarly reinforced bottom tail dependence by t copula cannot affect the left tail of the extrapolated rich list.

18.³⁴ However, compared with high wealth (BB8) copula, illustrated in Figure 19, tail dependence introduced by high income (t) copula together with the unrealistic high conditioning wealth bound can also make the tail too thick and overbend the proportion curve.³⁵

For the top wealth list in Figure 18, copula-based extrapolation with high income (t) copula even outperforms the marginal approach. As claimed, the conditioning income lower bound 2.9707 million has positive probability to be true given the farther-out minimum wealth in the rich list. This superiority over the marginal approach cannot be observed for the high wealth (BB8) copula in Figure 19 for the top income distribution. The conditioning wealth lower bound being as high as 76.3 million is simply unrealistic.

Full sample case has a t copula estimated with parameters close to those in high income case. Consequently, the same evidence on high income (t) copula above can be applied to the full sample case.³⁶

High income – high wealth case has also a t copula estimated. The main difference from the t copula estimated for high income case is the far larger degree-of-freedom (2nd parameter being 30 vs 7.12 for high income case). Figure 20 presents its benchmark on the top rich list and Figure 21 displays the benchmark on the top tax data. Again, the similar discussion on the high income (t) copula applies. Furthermore, a higher degree-of-freedom estimate reduces the tail dependence (on both tails) in the t copula. Hence, the copula-based extrapolation in the high income – high wealth case looks closer to the real top tax data distribution than the high income case in Figure 21.

In the following exercises, we continue to use high income case. All the results should equally apply to the full sample and high income – high wealth cases.

5.2.2 Between conditioning points

The previous analysis proposes income lower bound for the top rich list may be quite high if we observe that the very top tail in the rich list is quite thick (eg compared to the marginal extrapolation) and believe the income inequality drives the wealth one. This is supported, relative to the starting lower bound 2.9707 million, by the Figure 22 when we raise the income lower bound to 5.9707 million in copula 1 and also Figure 23 when

³⁴ The proportion curve of the copula 2 (high wealth – BB8 copula) overlaps with that of the marginal extrapolations.

³⁵ High wealth (BB8) copula does also introduce more top tail thickness to bend down the proportion curve in relative to the marginal extrapolation. This can be attributed to the excessively high conditioning wealth lower bound. On the other hand, BB8 copula does have a mass in the top end [1, 1] even though the tail dependence is zero.

³⁶ For brevity, I only show a counterpart of Figure 18 in Figure 25

this lower bound becomes as low as 0.9707 million in copula 1. The first improves the extrapolation by skewing the distribution to right while the second worsens it.

The other claim is derived from the weak influence from wealth inequality on income one: tail thickness of the top tax data relative to marginal extrapolation may imply that wealth lower bound is not that high in reality. Figure 24 shows such a situation. When we lower the lower bound wealth in the top 1% tax data to 350,000, compared to the starting point 76.3 million, extrapolation based on high income (t) copula is much closer to the true distribution which even out performs the marginal extrapolation. Reducing the lower bound wealth simply contract the right top tail in the copula-based extrapolation. However, asymmetrical copula (eg BB8) seems to approximate the true copula closer than the symmetrical one (eg t) even while lowering the conditioning wealth (see Figure 26).

5.2.3 Most approximating settings

Raising conditioning income progressively does not enforce the copula-based extrapolation to infinitely converge to the true rich list distribution. Lowering conditioning wealth to be very small also does not bend the copula-based extrapolation unlimitedly approaching the true top tax data distribution. There seems to exist some “optimal” conditioning value such that the extrapolated distribution exactly overlaps the true one since we observe that changing the conditioning value further will make the extrapolation cut across the true distribution. Searching such “optimal” point is focusing on using the high income (t) copula which is representative of our main candidates – full sample and high income – high wealth cases. By trial and error, Figure 27 shows such an “optimal” conditioning income lower bound as roughly 20 million. Figure 28 presents the “optimal” conditioning wealth lower bound at about 0.7 million. Claiming the minimum income for the observations in the rich list as 20 million seems to be too exceptional.³⁷ Nevertheless, it is rather plausible that the top 1% income earners have at least 0.7 million wealth. Moreover, the tax data is more reliable than the rich list and approximation with top 1% distribution in Figure 28 is fairly accurate. The copula-based approach is reasonably justified by the reality.

6. Conclusion

This paper proposes a copula-based joint extrapolation for the top income and wealth distributions which are commonly under-covered in the survey data. When benchmarking with the empirical top rich distribution from administrative tax data and

³⁷ 20 million is almost twice the p999 within the top 1% tax data – 11.31 million. On the other hand, tax data does not cover capital income. We cannot completely rule out its plausibility.

wealth rich list, the copula-based extrapolation does always perform better than the marginal one when the copula is estimated under a range of plausible scenarios, especially the full sample / high income - high wealth cases, as long as the conditioning lower bound is sensibly set.

The adoption of joint association has a positive net contribution beyond the marginal approach when estimating the copula from the sample almost following the distributions specified in the marginal approach. The approximation towards the true distributions can be almost exact when some realistic conditioning lower bound is provided. For example, if credibly assuming the minimum wealth for the population in top 1% income distribution to be about 0.7 million, our copula-based extrapolation using a copula estimated from the whole PHF sample can almost exactly replicate the distribution of top 1% tax data.

In the next steps, non-parametric / empirical copula can be introduced to replace our current parametric ones. The exercise can also extend to use other parametric competitors of marginal distributions. Although our parallel rank association assumption has not been established empirically, there are evidences from other countries to support the statistical features of copula estimated (eg symmetricity and tail dependence). These properties are fundamental in explaining the merits of introducing joint association. The economic mechanism behind can at least qualitatively endorses these features should be widely consistent across the top subpopulation.⁴⁰ Nevertheless, we can test the validity of our parallel rank association assumption by examining the stability of copula structure, eg using cross-validation. Also, we can apply our methods in alternative data / countries.

The copula-based approach, or generally to introduce the joint association, can be more valuable in the situation when one margin is completely observed (to the very top; eg there is the administrative data of income or wealth and the survey simply draws the values from it while asking respondents the other dimension). Our exercises also justify the potential efforts to acquire the minimum information of the other measurement dimension. In a nutshell, to have a precise knowledge of the lower bound of income for the top missing wealth tail is cheap to achieve since it is simply one number.

⁴⁰ For example, high income subpopulation are very likely bear a symmetric copula for these two dimensions no matter whether the very top tails are covered or not.

References

- Aaberge, R., Atkinson, A. B., & Königs, S. (2018). From Classes to Copulas: Wages, capital, and top incomes. *The Journal of Economic Inequality*, 16(2), 295-320.
- Atkinson, A. B. (2017). Pareto and the upper tail of the income distribution in the UK: 1799 to the present. *Economica*, 84(334), 129-156.
- Bach, L., Calvet, L. E., & Sodini, P. (2017). From saving comes having? Disentangling the impact of savings on inequality. *Swedish House of Finance Research Paper*, 18-8.
- Bach, S., Corneo, G., & Steiner, V. (2012). Optimal top marginal tax rates under income splitting for couples. *European Economic Review*, 56(6), 1055-1069.
- Bach, S., Thiemann, A., & Zucco, A. (2015). The top tail of the wealth distribution in Germany, France, Spain, and Greece. Discussion paper 1502, Deutsches Institut für Wirtschaftsforschung, Berlin, 2015.
- Bartels, C., & Metzing, M. (2017). An integrated approach for a top-corrected income distribution. *The Journal of Economic Inequality*, 1-19.
- Blanchet, T., Fournier, J., & Piketty, T. (2017). Generalized Pareto Curves: Theory and Applications.
- Brechmann, E., & Schepsmeier, U. (2013). Cdvine: Modeling dependence with c-and d-vine copulas in r. *Journal of Statistical Software*, 52(3), 1-27.
- Buike, A, "Copula Modeling for World's Biggest Competitors", Master's Thesis, University of Amsterdam, 2018.
- Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," SIAM Review, vol. 51, no. 4, pp. 661–703, 2009.
- Dalitz, C. (2016). "Estimating Wealth Distribution: Top Tail and Inequality."
Technischer Bericht Nr. 2016-01, Hochschule Niederrhein, Fachbereich Elektrotechnik und Informatik

Eckerstorfer, P., Halak, J., Kapeller, J., Schütz, B., Springholz, F., & Wildauer, R. (2016). Correcting for the missing rich: An application to wealth survey data. *Review of Income and Wealth*, 62(4), 605-627.

Hofert, M., Kojadinovic, I., Maechler, M., & Yan, J. (2017). copula: Multivariate dependence with copulas. *R package version 0.999-17*, <https://cran.r-project.org/web/packages/copula/copula.pdf>

Jenkins, S. P. (2017). Pareto models, top incomes and recent trends in UK income inequality. *Economica*, 84(334), 261-289.

Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman & Hall, London.

Joe, H. and Xu, J. (1996). "The Estimation Method of Inference Functions for Margins for Multivariate Models." Technical Report 166, Department of Statistics, University of British Columbia.

Kojadinovic, I. (2017). Some copula inference procedures adapted to the presence of ties. *Computational Statistics & Data Analysis*, 112, 24-41.

Kojadinovic, I., & Yan, J. (2010). Modeling multivariate distributions with continuous marginals using the copula R package. *Journal of Statistical Software*, 34(9), 1-20.

Lieberknecht, P., & Vermeulen, P. (2018). Inequality and relative saving rates at the top (No. 2204). European Central Bank.

Roodman, D. (2016). EXTREME: Stata module to fit models used in univariate extreme value theory. Statistical Software Components archive. Boston, MA: Boston College. <https://ideas.repec.org/c/boc/bocode/s457953.html>

Schepsmeier, U., Stoeber, J., Brechmann, E. C., Graeler, B., Nagler, T., Erhardt, T., ... & Killiches, M. (2017). Package 'VineCopula', <https://cran.r-project.org/web/packages/VineCopula/VineCopula.pdf>

Sheather, S. J., & Jones, M. C. (1991). A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 683-690.

Trivedi, Pravin K., and David M. Zimmer. "Copula modeling: an introduction for practitioners." *Foundations and Trends in Econometrics* 1, no. 1 (2007): 1-111.

Vermeulen, P. (2018). How fat is the top tail of the wealth distribution?. *Review of Income and Wealth*, 64(2), 357-387.

Yan, J. (2007). Enjoy the joy of copulas: with a package copula. *Journal of Statistical Software*, 21(4), 1-21.

Zhu, J. (2014). Bracket Creep Revisited-with and without $r > g$: Evidence from Germany. *Journal of Income Distribution*, 23(3), 106-158.

**Table 1 German gross income for the top 30% by percentile mean: PHF (2009)
and Income tax statistics (Est, 2008)**

Percentile	PHF 09	Est 08	Relative difference (%: (PHF - Est)/Est*100)
71	40,769	46,517	-12
72	41,742	46,918	-11
73	42,591	48,414	-12
74	43,585	49,539	-12
75	44,728	49,563	-10
76	45,677	51,544	-11
77	46,481	51,697	-10
78	47,901	53,006	-10
79	49,634	54,557	-9
80	50,956	55,681	-8
81	52,409	56,771	-8
82	53,805	59,125	-9
83	55,064	59,680	-8
84	56,650	61,686	-8
85	58,259	63,085	-8
86	60,414	64,515	-6
87	62,663	67,368	-7
88	65,488	69,412	-6
89	68,027	71,715	-5
90	71,711	73,340	-2
91	74,938	76,061	-1
92	78,888	78,858	0
93	83,021	82,120	1
94	88,231	88,214	0
95	94,183	93,449	1
96	104,036	100,041	4
97	114,847	111,194	3
98	132,376	127,176	4
99	158,500	162,689	-3
100	277,084	531,812	-48

Note: Est 08 - income tax return data in 2008 from Income tax statistics (Einkommensteuerstatistik; source: Research Data Centre of the Federal Statistical Office of Germany); PHF 09 - wave one of the Panel on Household Survey (Bundesbank). Both are measured as the gross income for the tax unit population.

Table 2 Candidate copula estimates and the marginal distribution specification/estimates by different fitting samples

	Case	High income	High wealth	Full sample	High income - high wealth
Fitting sample	income cutoff in forming fitting sample	86,957	0		86,957
	wealth cutoff in forming fitting sample	0	245,160		245,160
Copula estimate	copula family	t	BB8	t	t
	1st parameter - copula	0.334	1.64	0.589	0.275
	2nd parameter - copula	7.12	0.943	6.2	30
	unweighted proportion of fitting sample (total sample size 8,026)	0.21	0.41	1	0.16
Marginal distribution (Pareto II)	threshold -income	86,957	86,957	86,957	87,000
	shape parameter - income	0.41	0.41	0.41	0.34
	scale parameter - income	34612	34612	34612	45208
	threshold - wealth	245,160	245,160	245,160	245,680
	shape parameter - wealth	0.60	0.60	0.60	0.63
	scale parameter - wealth	197203	197203	197203	370165

Note: Each case is listed in columns which corresponds to the fitting sample above the income and wealth cutoffs. The marginal distribution is used together with the copula estimate in each case to construct conditional probability. Only the couple for the case of high income-high wealth is also estimated under the same fitting sample as copula estimate. Note that, due to the joint restriction of cutoffs, the exact income and wealth lower bounds (ie threshold parameters specified in the Pareto II distribution) are 87,000 and 245,680 which are slightly higher than the cutoffs 86,957 and 245,160.

Figure 1 Empirical inverted Pareto curve for the couples' income in the German tax data (2005) - Fig 2 in Bach et al. (2012)

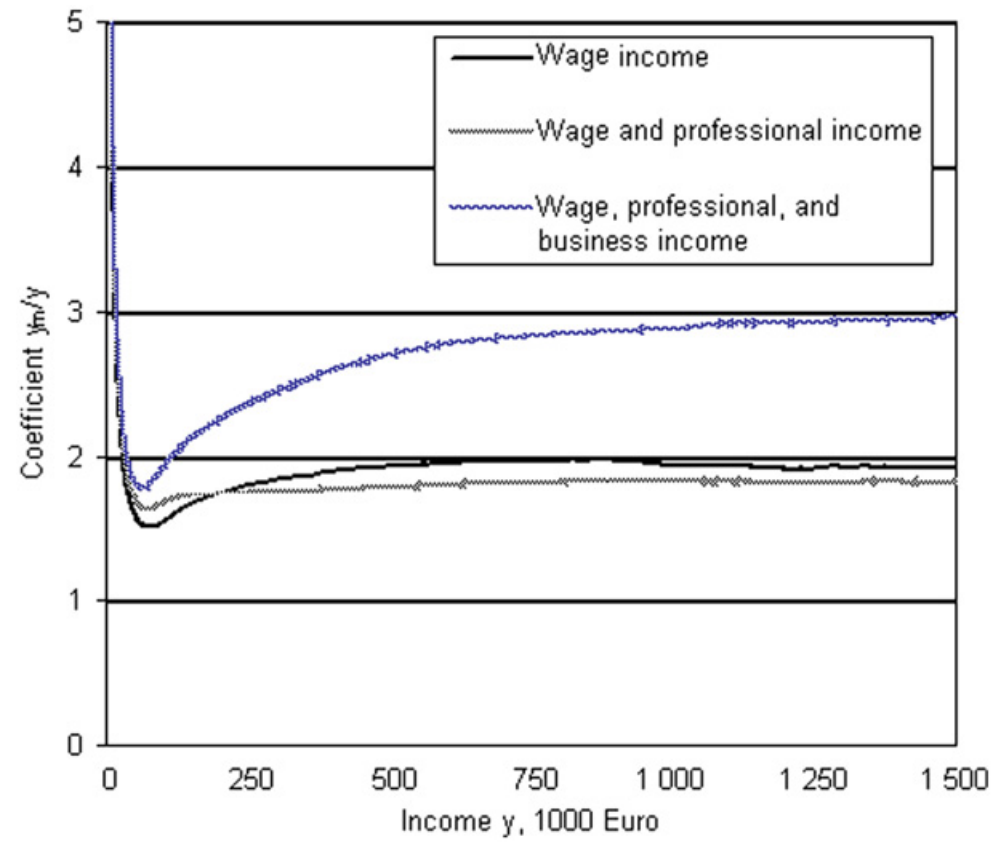


Figure 2 Goodness-of-fit criteria after Kolmogorov-Smirnov (KS) as a function of threshold for the averaged PHF income data (dashed lines: p90, p95, p99, p99.5)

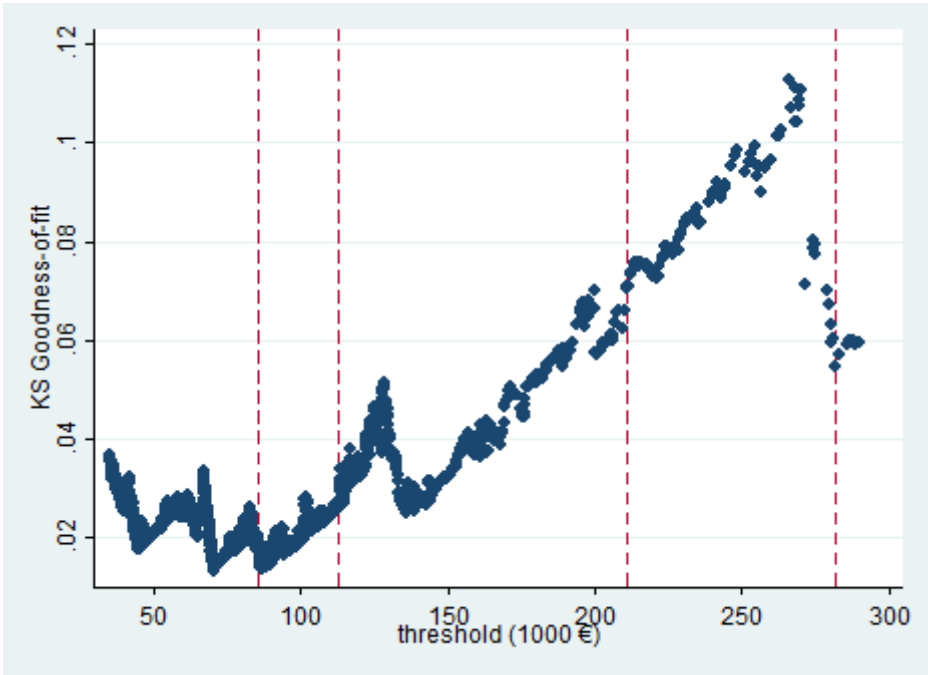


Figure 3 Pareto II parameter estimates by threshold for the averaged PHF income data (dashed lines: p90, p95, p99, p99.5)

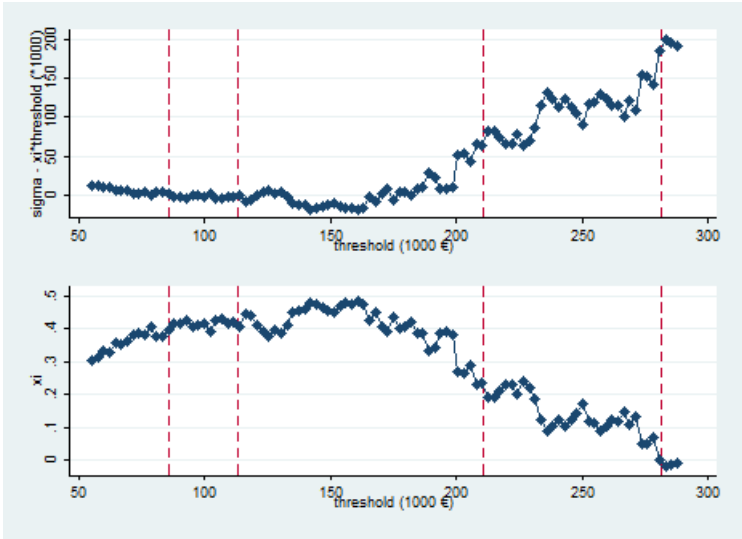


Figure 4 inverted Pareto coefficient w.r.t. income for the Pareto II distribution (threshold=86,957, shape=.42 and scale=34,241)

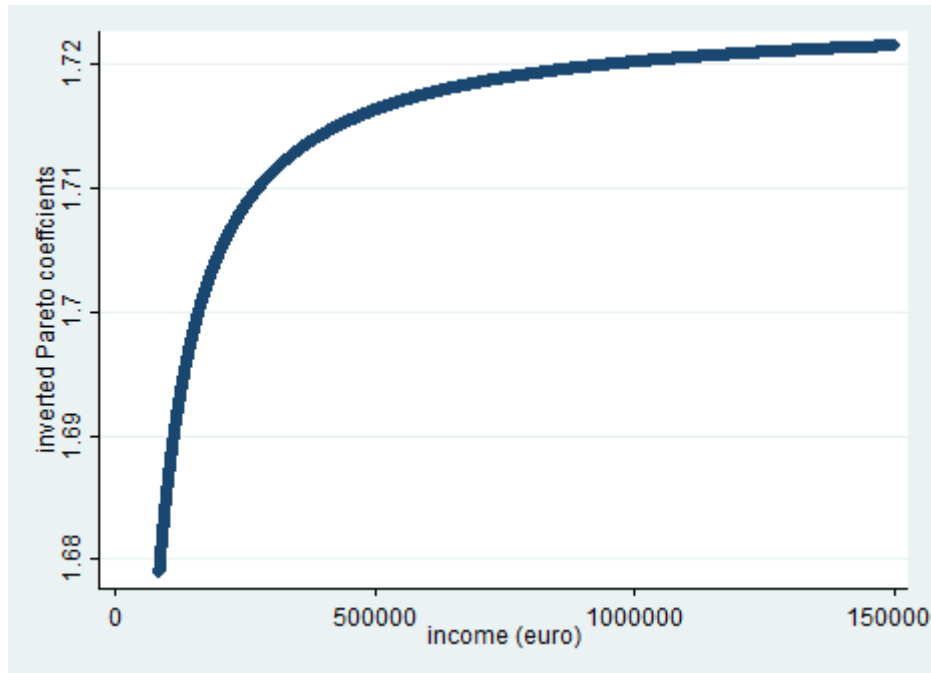


Figure 5 inverted Pareto coefficient w.r.t. probability for the Pareto II distribution (threshold=86,957, shape=.42 and scale=34,241)

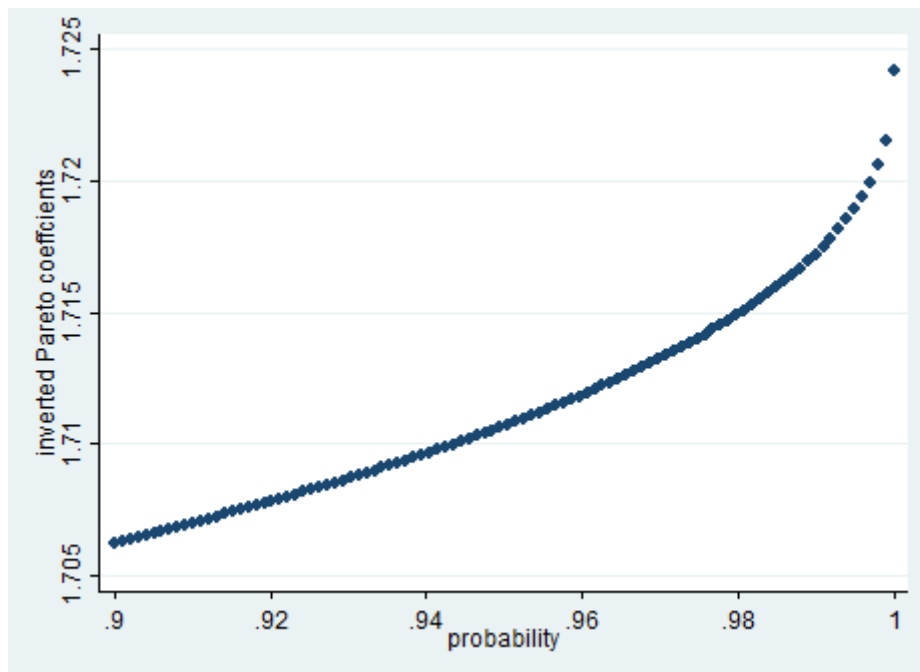


Figure 6 Goodness-of-fit criteria after Kolmogorov-Smirnov (KS) as a function of threshold for the averaged PHF wealth data (dashed lines: p90, p95, p99, p99.5)

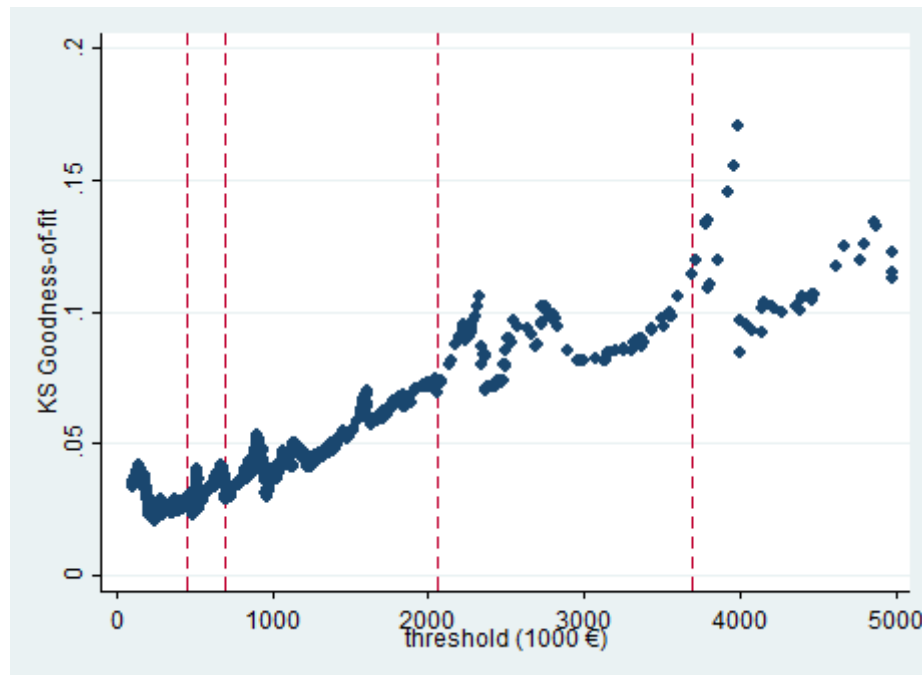


Figure 7 Pareto II parameter estimates by threshold for the averaged PHF wealth data (dashed lines: p90, p95, p99, p99.5)

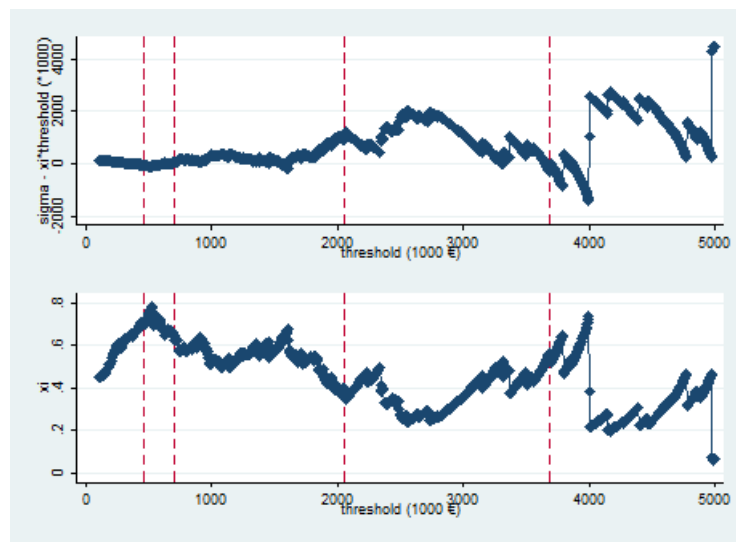


Figure 8 inverted Pareto coefficient w.r.t. wealth for the Pareto II distribution (threshold=245,160, shape=.6 and scale=197,206)

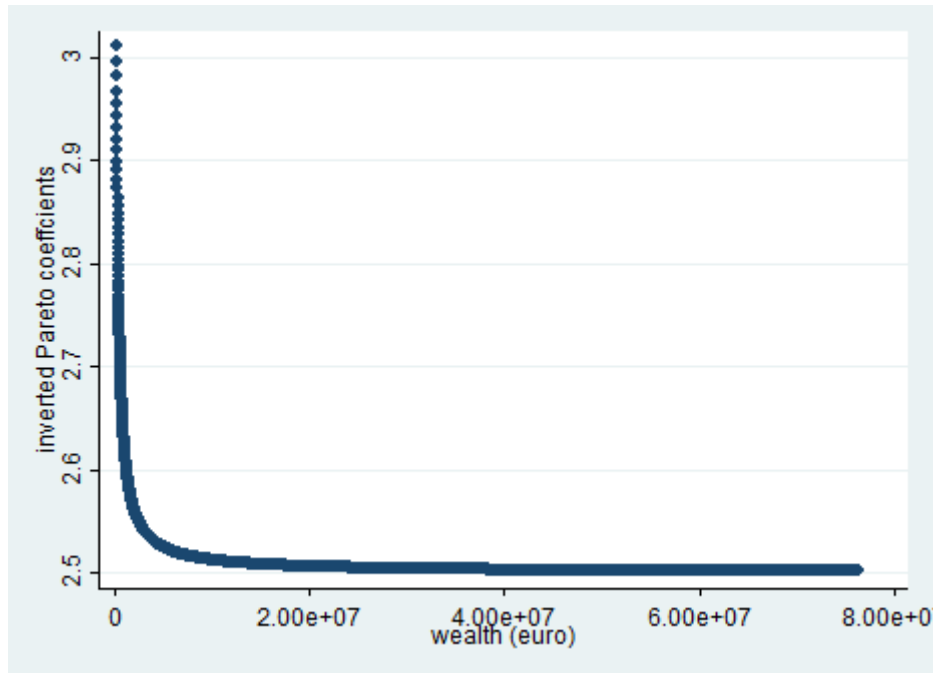


Figure 9 inverted Pareto coefficient w.r.t. probability for the Pareto II distribution (threshold=245,160, shape=.6 and scale=197,206)

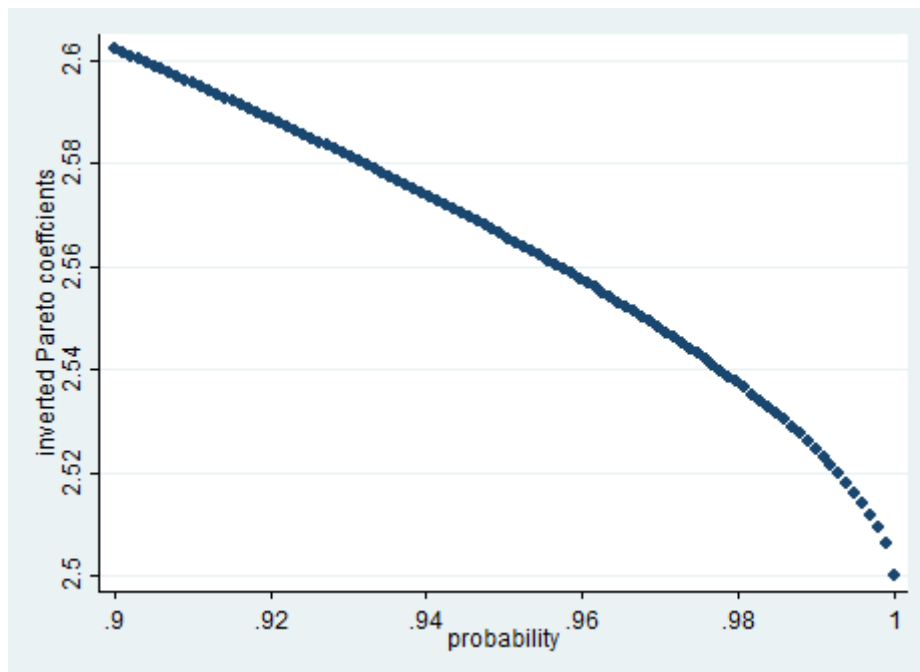


Figure 10 Contour curves of fitted and empirical copula over marginal CDFs of income and wealth – high income case (income cutoff = 86,957, wealth cutoff = 0, copula family, 1st and 2nd parameters – t , 0.334, 7.12); joint CDFs are marked above each contour curves

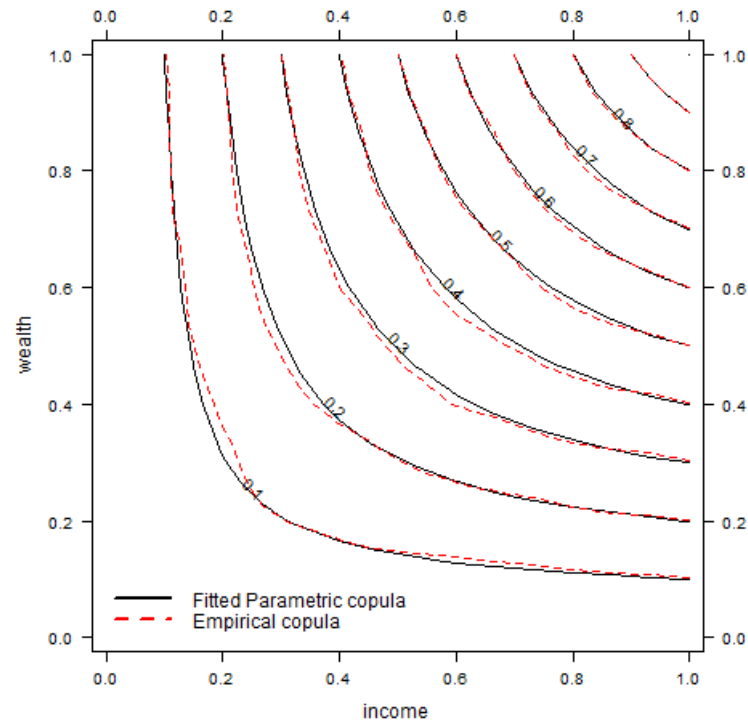


Figure 11 Contour curves of fitted and empirical copula over marginal CDFs of income and wealth – high wealth case (income cutoff = 0, wealth cutoff = 245,160, copula family, 1st and 2nd parameters – BB8, 1.64, 0.943); joint CDFs are marked above each contour curves

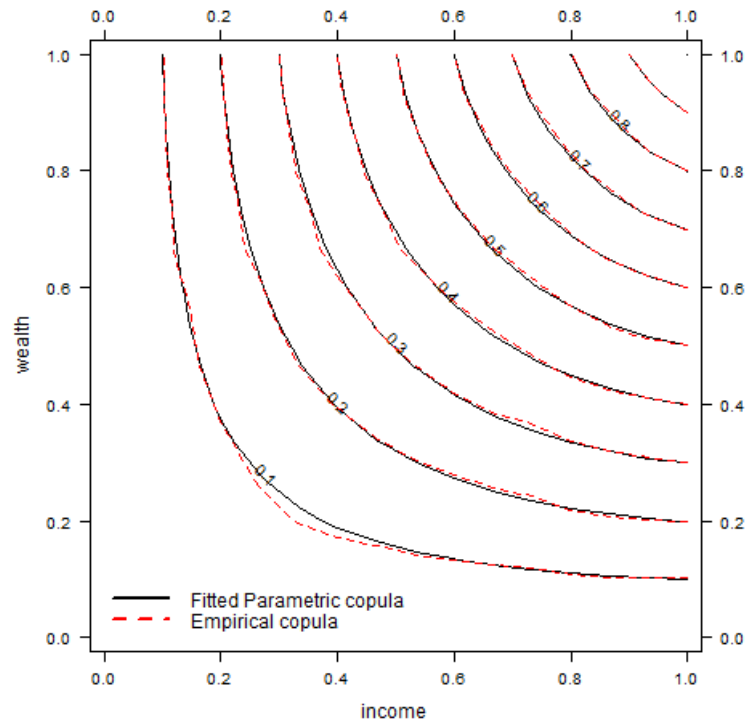


Figure 12 Contour curves of fitted and empirical copula over marginal CDFs of income and wealth – full sample case (copula family, 1st and 2nd parameters – t, 0.589, 6.2); joint CDFs are marked above each contour curves

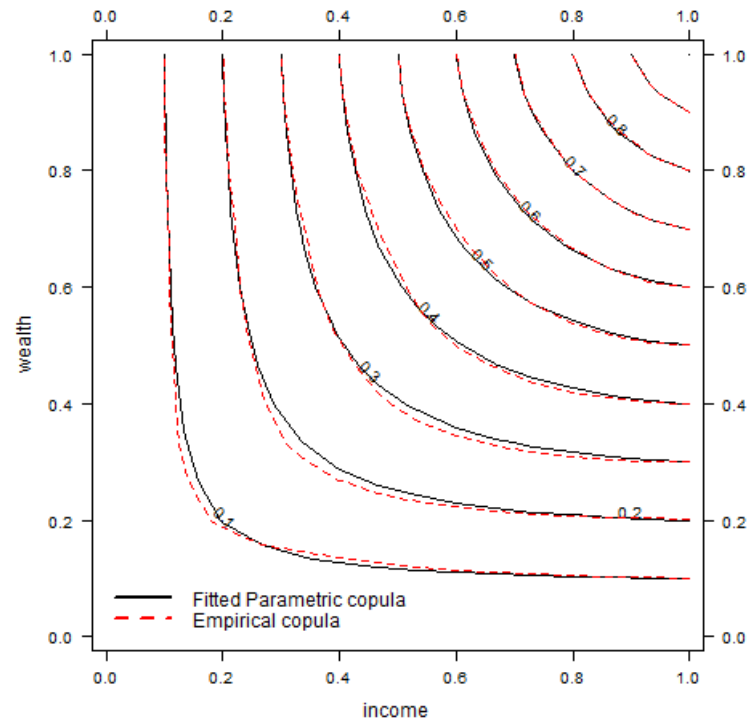


Figure 13 Contour curves of fitted and empirical copula over margins of income and wealth (unit: 1,000 euros; log-scaled axes) – High income – high wealth case (income cutoff = 86,957, wealth cutoff = 245,160, copula family, 1st and 2nd parameters – t , 0.275, 30) ; joint CDFs are marked above each contour curves

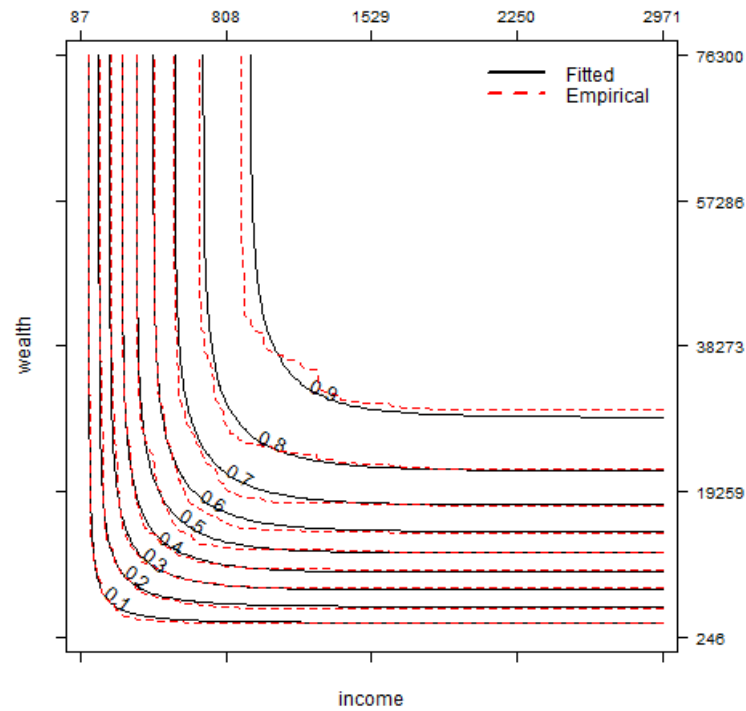


Figure 14 Density plot of fitted copula over marginal CDFs of income and wealth – high income case (income cutoff = 86,957, wealth cutoff = 0, copula family, 1st and 2nd parameters – t , 0.334, 7.12)

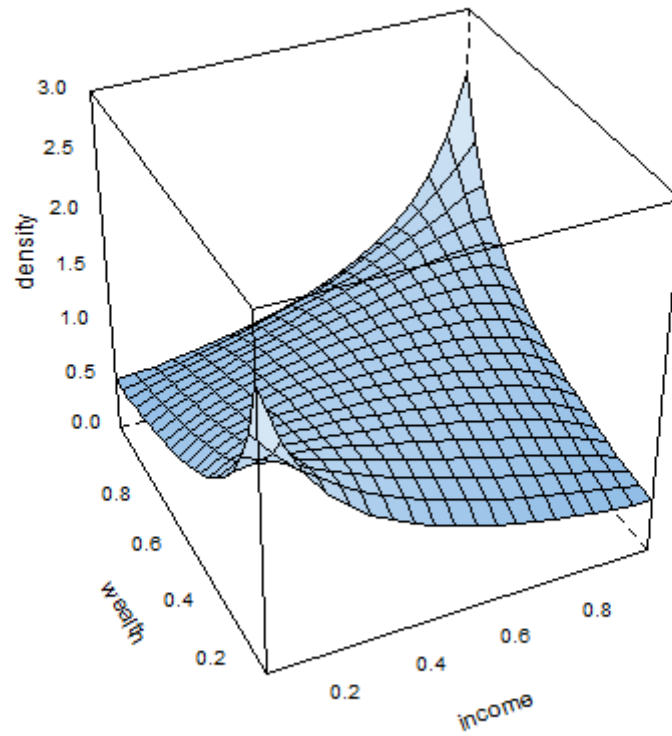


Figure 15 Density plot of fitted copula over marginal CDFs of income and wealth – high wealth case (income cutoff = 0, wealth cutoff = 245,160, copula family, 1st and 2nd parameters – BB8, 1.64, 0.943)

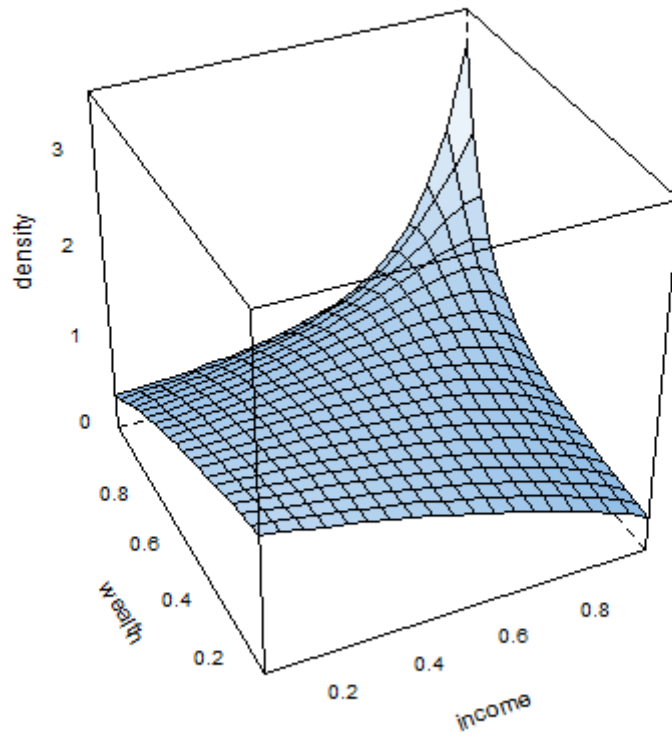


Figure 16 Density plot of fitted copula over marginal CDFs of income and wealth – full sample case (copula family, 1st and 2nd parameters – t, 0.589, 6.2)

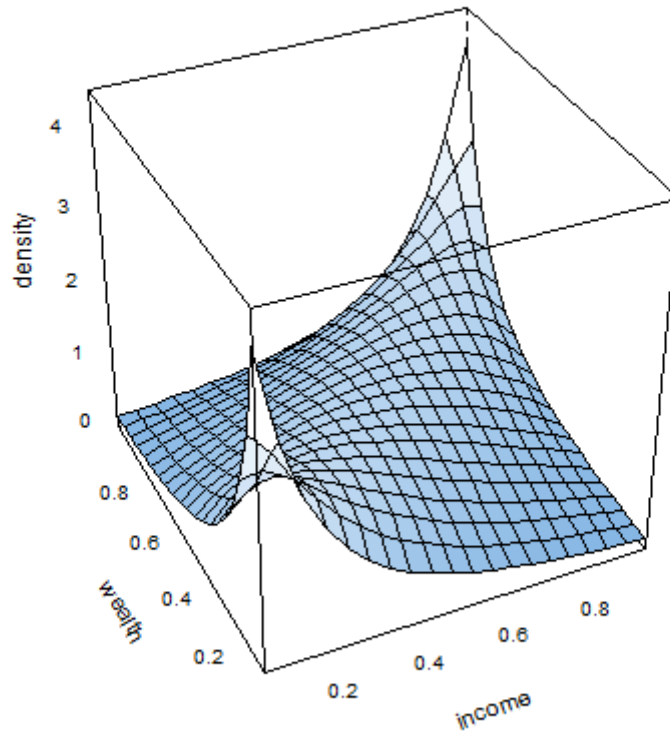
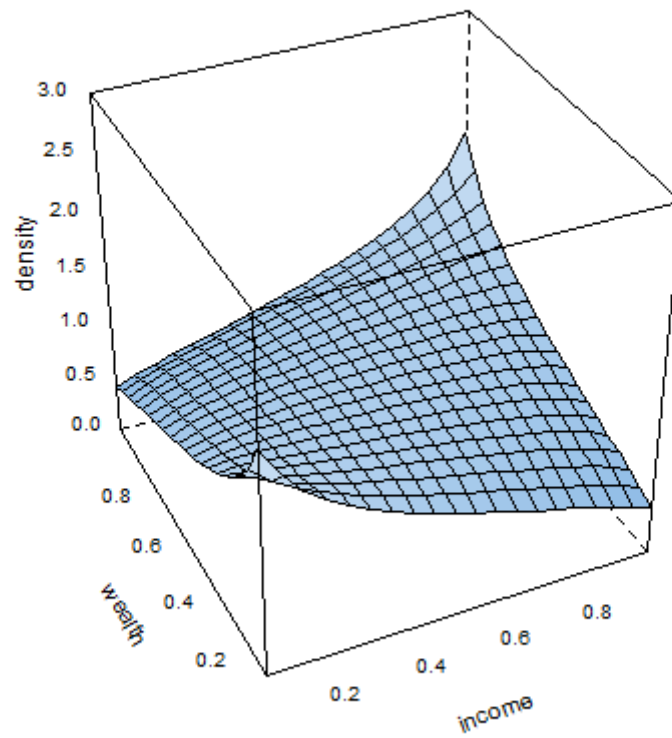


Figure 17 Density plot of fitted copula over marginal CDFs of income and wealth – high income – high wealth case (income cutoff = 86,957, wealth cutoff = 245,160, copula family, 1st and 2nd parameters – t, 0.275, 30)



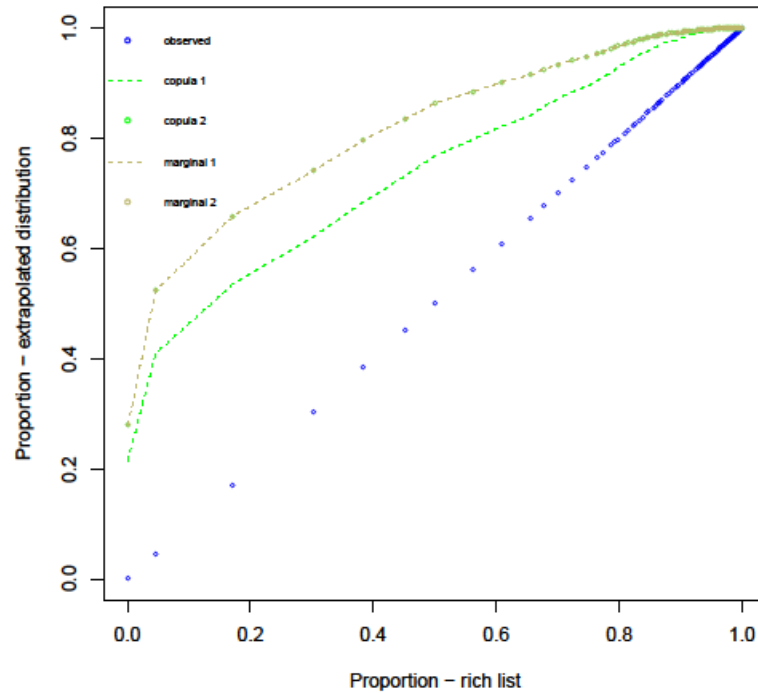


Figure 18 P-p plot benchmarking marginal and copula-based extrapolated distributions with top rich list – comparing high income (t) and high wealth (BB8) copulas:

1 – high income case and 2 – high wealth case

marginal 1 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

marginal 2 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 1: - *Student t* (parameter 1=0.334, parameter 2=7.12), conditioning income lower bound=2,970,700 and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 2: - BB8 (parameter 1=1.64, parameter 2=0.943), conditioning income lower bound=2,970,700 and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

Dots in unconnected lines (eg observed) correspond to the exact observations which reflect the dispersion of the sample.

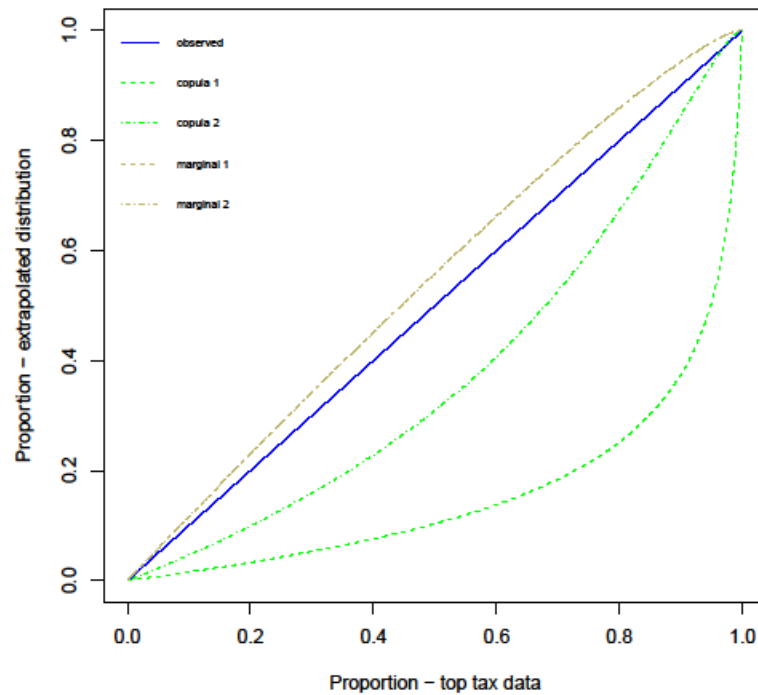


Figure 19 P-p plot benchmarking marginal and copula-based extrapolated distributions with top income tax data - comparing high income (t) and high wealth (BB8) copulas:

1 – high income case and 2 – high wealth case

marginal 1 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

marginal 2 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 1: - *Student t* (parameter 1=0.334, parameter 2=7.12), conditioning wealth lower bound=76,300,000 and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 2: - *BB8* (parameter 1=1.64, parameter 2=0.943), conditioning wealth lower bound=76,300,000 and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

It is not necessary to show the dispersion of this sample because they are 1000 percentile points within top 1%.

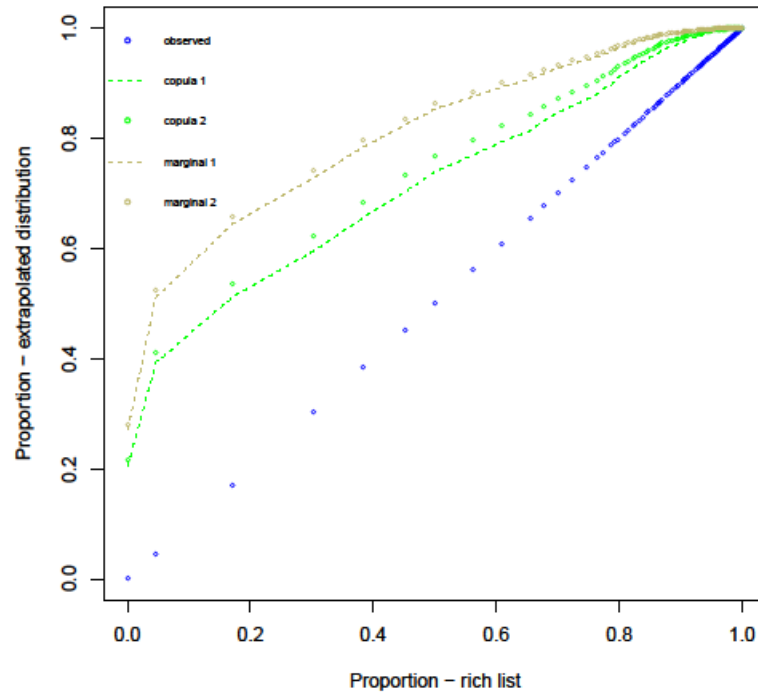


Figure 20 P-p plot benchmarking marginal and copula-based extrapolated distributions with top rich list – comparing high income-high wealth (t) and high income (t) copulas:

1 – high income-high wealth case and 2 – high income case

marginal 1 – wealth distribution - Pareto II (*threshold=245680, shape=.63, scale=370165*)

marginal 2 – wealth distribution - Pareto II (*threshold=245160, shape=.6, scale=197203*)

copula 1: - Student t (*parameter 1=0.275, parameter 2=30*), conditioning income lower

bound=2,970,700 and income distribution - Pareto II (*threshold=87000, shape=.34, scale=45208*)

copula 2: - Student t (*parameter 1=1.64, parameter 2=0.943*), conditioning income lower

bound=2,970,700 and income distribution - Pareto II (*threshold=86957, shape=.41, scale=34612*)

Dots in unconnected lines (eg observed) correspond to the exact observations which reflect the dispersion of the sample.

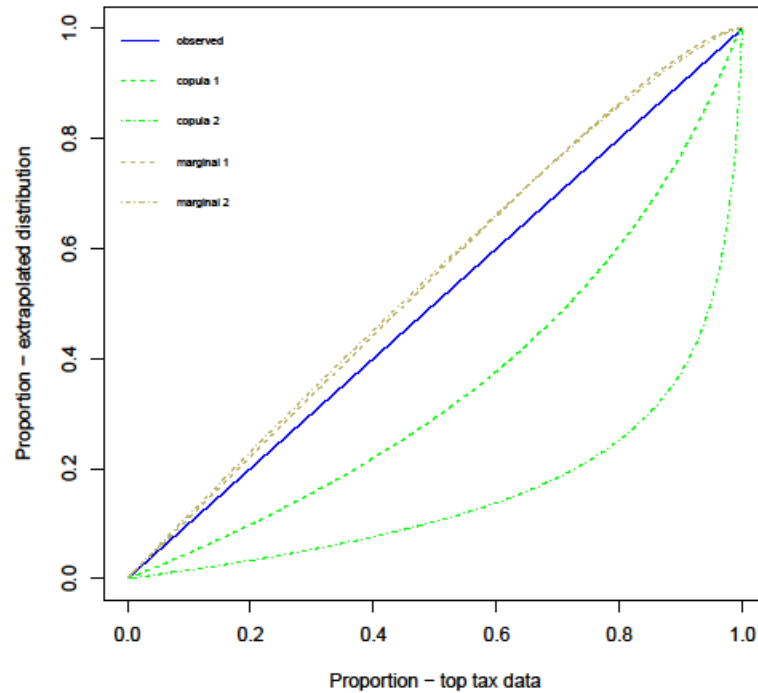


Figure 21 P-p plot benchmarking marginal and copula-based extrapolated distributions with top income tax data – comparing high income-high wealth (t) and high income (t) copulas::

1 – high income-high wealth case and 2 – high income case

marginal 1 – income distribution - Pareto II (*threshold=87000, shape=.34, scale=45208*)

marginal 2 – income distribution - Pareto II (*threshold=86957, shape=.41, scale=34612*)

copula 1: - Student t (*parameter 1=0.275, parameter 2=30*), conditioning wealth lower

bound=76,300,000 and wealth distribution - Pareto II (*threshold=245680, shape=.63, scale=370165*)

copula 2: - Student t (*parameter 1=1.64, parameter 2=0.943*), conditioning wealth lower

bound=76,300,000 and wealth distribution - Pareto II (*threshold=245160, shape=.6, scale=197203*)

It is not necessary to show the dispersion of this sample because they are 1000 percentile points within top 1%.

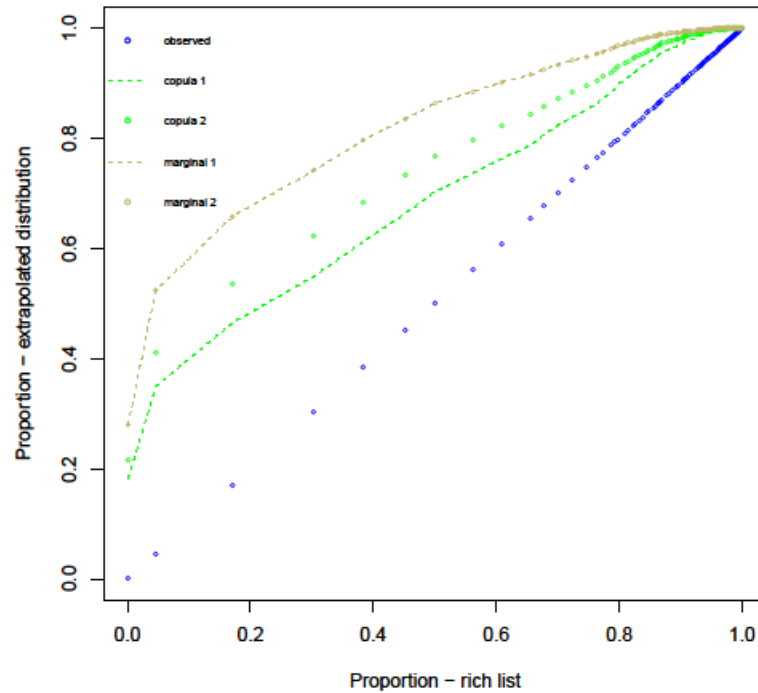


Figure 22 P-p plot benchmarking marginal and copula-based extrapolated distributions with top rich list – raising conditioning income:

1 – high income case and 2 – high income case

marginal 1 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

marginal 2 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 1: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning income lower bound=5,970,700* and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 2: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning income lower bound=2,970,700* and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

Dots in unconnected lines (eg observed) correspond to the exact observations which reflect the dispersion of the sample.

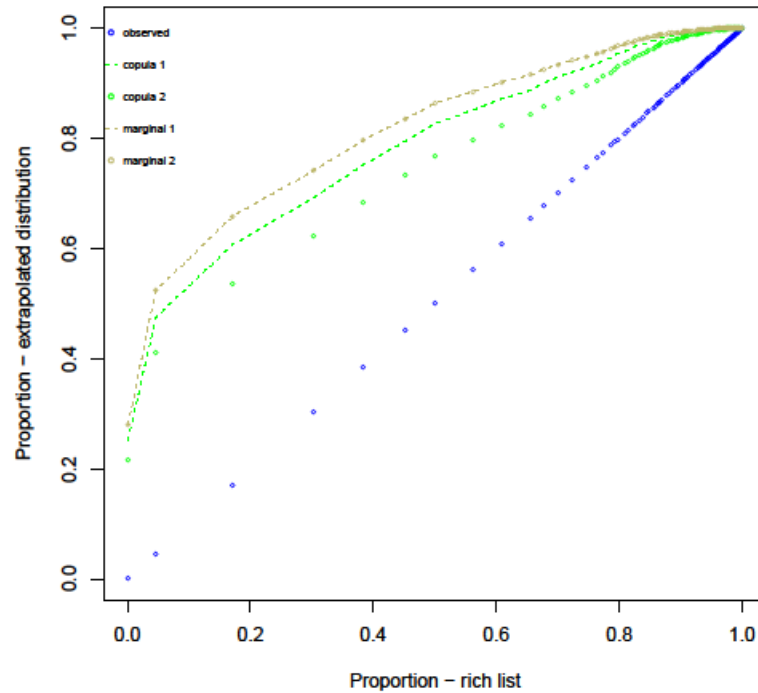


Figure 23 P-p plot benchmarking marginal and copula-based extrapolated distributions with top rich list - lowering conditioning income:

1 – high income case and 2 – high income case

marginal 1 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

marginal 2 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 1: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning income lower bound=970,700* and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 2: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning income lower bound=2,970,700* and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

Dots in unconnected lines (eg observed) correspond to the exact observations which reflect the dispersion of the sample.

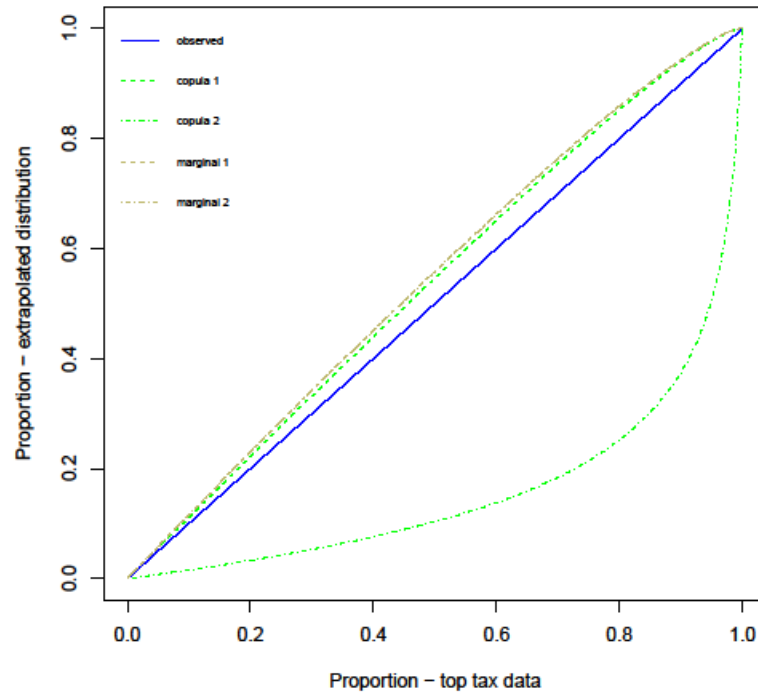


Figure 24 P-p plot benchmarking marginal and copula-based extrapolated distributions with top income tax data – lowering conditioning wealth:

1 – high income case and 2 – high income case

marginal 1 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

marginal 2 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 1: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning wealth lower bound=350,000* and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 2: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning wealth lower bound=76,300,000* and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

It is not necessary to show the dispersion of this sample because they are 1000 percentile points within top 1%.

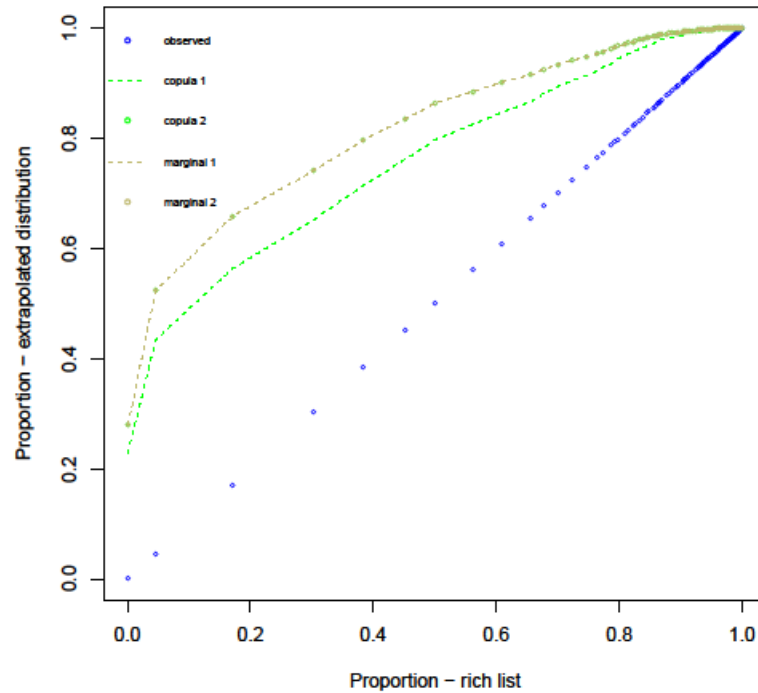


Figure 25 P-p plot benchmarking marginal and copula-based extrapolated distributions with top rich list – comparing full sample (t) and high wealth (BB8) copulas:

1 – full sample case and 2 – high wealth case

marginal 1 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

marginal 2 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 1: - *Student t* (parameter 1=0.589, parameter 2=6.2), conditioning income lower

bound=2,970,700 and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 2: - BB8 (parameter 1=1.64, parameter 2=0.943), conditioning income lower

bound=2,970,700 and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

Dots in unconnected lines (eg observed) correspond to the exact observations which reflect the dispersion of the sample.

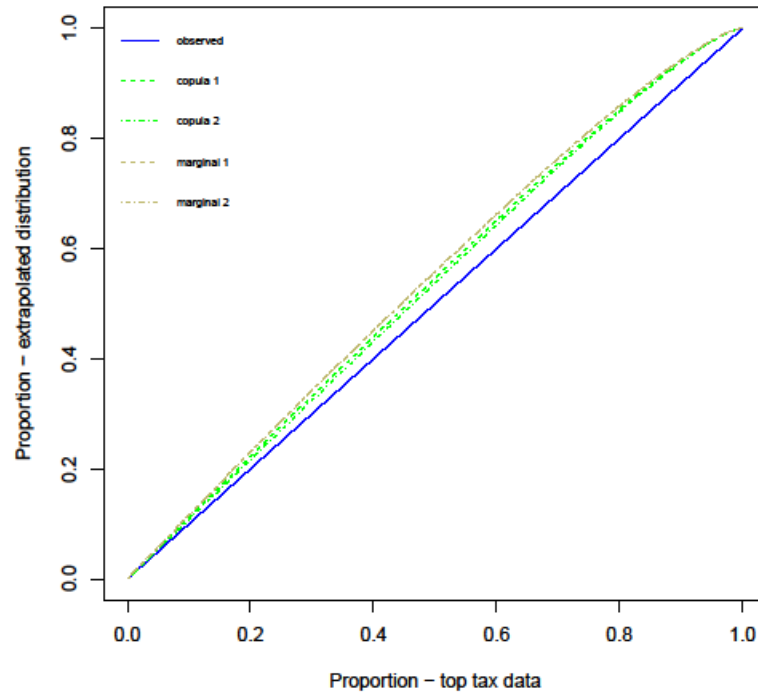


Figure 26 P-p plot benchmarking marginal and copula-based extrapolated distributions with top income tax data - comparing high income (t) and high wealth (BB8) copulas whiling lowering wealth:

1 – high income case and 2 – high wealth case

marginal 1 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

marginal 2 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 1: - *Student t* (parameter 1=0.334, parameter 2=7.12), conditioning wealth lower bound=350,000 and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 2: - *BB8* (parameter 1=1.64, parameter 2=0.943), conditioning wealth lower bound=350,000 and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

It is not necessary to show the dispersion of this sample because they are 1000 percentile points within top 1%.

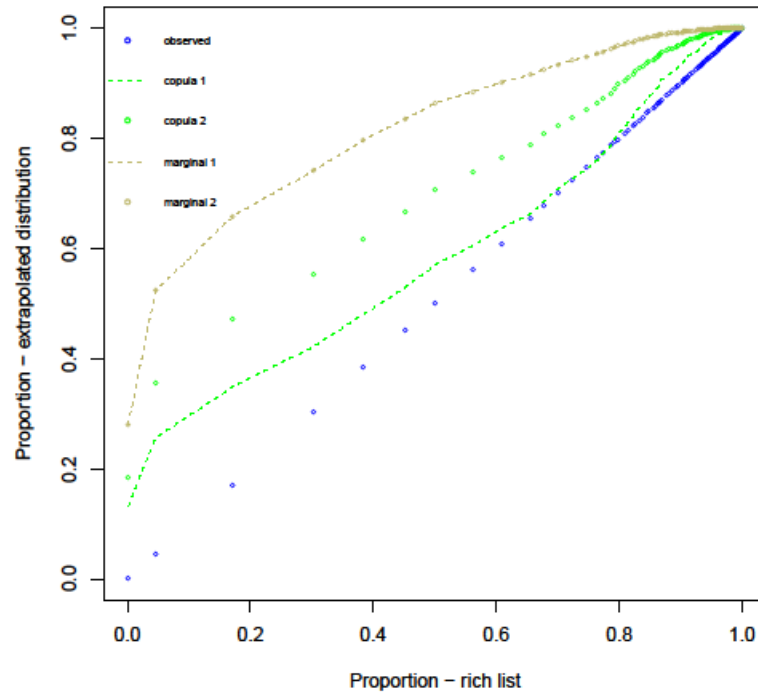


Figure 27 P-p plot benchmarking marginal and copula-based extrapolated distributions with top rich list – aggressively raising conditioning income:

1 – high income case and 2 – high income case

marginal 1 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

marginal 2 – wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 1: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning income lower bound*=20,000,000 and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 2: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning income lower bound*=5,970,700 and income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

Dots in unconnected lines (eg observed) correspond to the exact observations which reflect the dispersion of the sample.

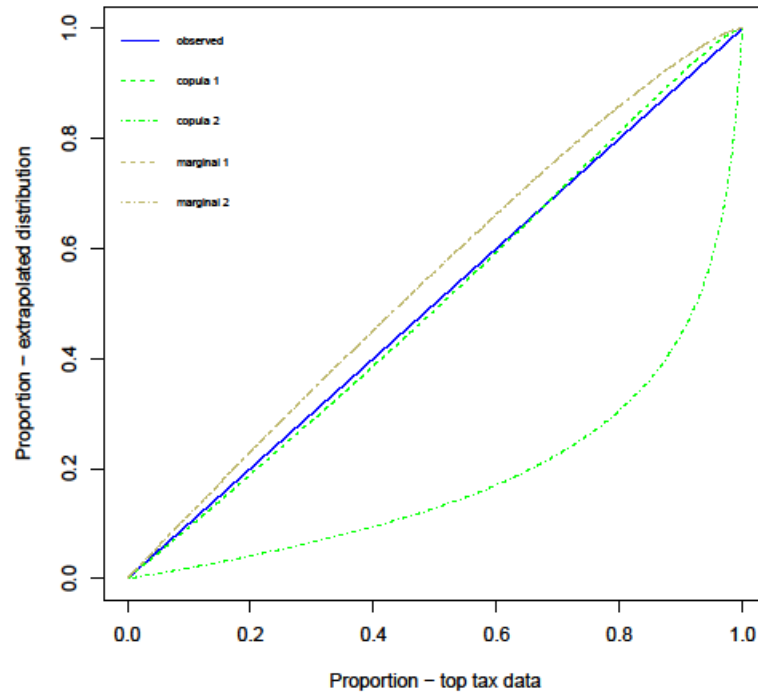


Figure 28 P-p plot benchmarking marginal and copula-based extrapolated distributions with top income tax data – aggressively lowering conditioning wealth:

1 – high income case and 2 – high income case

marginal 1 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

marginal 2 – income distribution - Pareto II (threshold=86957, shape=.41, scale=34612)

copula 1: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning wealth lower bound=700,000* and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

copula 2: - Student t (parameter 1=0.334, parameter 2=7.12), *conditioning wealth lower bound=76,300,000* and wealth distribution - Pareto II (threshold=245160, shape=.6, scale=197203)

It is not necessary to show the dispersion of this sample because they are 1000 percentile points within top 1%.