UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SÓCIO-ECONÔMICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Adriano Filipe da Silva Maia

Three essays on interoccupational earnings distribution

Florianópolis
2022

Adriano Filipe da Silva Maia

Three essays on interoccupational earnings distribution

Tese submetida ao Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina para a obtenção do título de doutor em Economia
Orientador: Prof. Eraldo Sérgio Barbosa da Silva, Dr.

Florianópolis

2022

Ficha de identificação da obra elaborada pelo autor através do Programa de Geração Automática da Biblioteca Universitária da UFSC

Maia, Adriano Filipe da Silva
Three essays on interoccupational earnings distribution / Adriano Filipe da Silva Maia ; orientador, Eraldo Sérgio Barbosa da Silva, 2022.
80 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Sócio-Econômico, Programa de Pós-Graduação em Economia, Florianópolis, 2022.

Inclui referências.

1. Economia. 2. Distribuição de renda. 3. Lei de potência. 4. Profissões escaláveis e não escaláveis. 5. Distribuição de Pareto. I. da Silva, Eraldo Sérgio Barbosa . II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Economia. III. Título.

Adriano Filipe da Silva Maia

Three essays on interoccupational earnings distribution

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Raul Yukihiro Matsushita, Dr.
UnB

Prof. Guilherme de Oliveira, Dr.
UFSC

Prof. Newton Carneiro Affonso da Costa Junior, Dr.
UFSC

Prof. Marcelo Henrique Romano Tragtenberg, Dr.
UFSC

Certificamos que esta é a **versão original e final** da tese que foi julgada adequada para obtenção do título de doutor em Economia.

ICP-Edu
Documento assinado digitalmente
Guilherme Valle Moura
Data: 29/03/2022 15:45:08-0300
CPF: 013.234.686-92
Verifique as assinaturas em https://v.ufsc.br

____

Coordenação do Programa de Pós-Graduação

ICP-Edu
Documento assinado digitalmente
Eraldo Sergio Barbosa da Silva
Data: 28/03/2022 14:09:00-0300
CPF: 278.930.814-49
Verifique as assinaturas em https://v.ufsc.br

____

Prof. Eraldo Sérgio Barbosa da Silva, Dr.
Orientador

Florianópolis, 2022.

Este trabalho é dedicado a todas as pessoas que acreditam que a ciência tem o poder de mudar o mundo.

AGRADECIMENTOS

Primeiramente, gostaria de agradecer ao Criador por me conduzir pelo caminho da luz e me guiar nos momentos-chave de tomada de decisão.

Agradeço à toda minha família, especialmente meus pais que nunca pouparam esforços para que eu tivesse acesso à melhor educação.

Agradeço também à minha querida mulher pelo companheirismo e ajuda na superação das dificuldades acadêmicas.

Aos estimados amigos pelo aprazível convívio, que me proporcionou boas risadas e bastante aprendizado.

Aos meus orientadores e a todos os demais professores que contribuíram direta e indiretamente na sólida formação adquirida ao longo dessa árdua jornada de doutorado.

“A bad book is as much of a labor to write as a good one; it comes as sincerely from the author’s soul.” (Aldous Huxley)

RESUMO

Técnicas de pesquisa oriundas da física estatística têm se mostrado uma ferramenta de análise bastante poderosa para explicar a dinâmica da distribuição de renda. No entanto, o atributo da escalabilidade das profissões não vem sendo explorado nessa agenda de pesquisa, embora apresente um enorme potencial para explicar as diferenças salariais. Foi sugerido que as ocupações em que alguém é pago por hora não são escaláveis, enquanto as ocupações escaláveis permitem que se ganhe mais dinheiro sem um aumento equivalente no trabalho e no tempo. Espera-se que as ocupações não escaláveis tenham baixa variância de renda, enquanto as escaláveis mostrem grandes desigualdades de renda. Este estudo examina as evidências para esta distinção sugerida usando microdados de rendimentos pessoais para doze ocupações candidatas a ambos os tipos, escaláveis ou não. De modo geral, encontramos que as caudas superiores de todas as distribuições decaem como leis de potência. No primeiro ensaio, testamos a eficácia de um modelo de distribuição de renda de duas classes que concilia a lei de Pareto para os indivíduos com maior renda e a distribuição log-normal para o restante da amostra. No segundo ensaio, computamos o ponto de corte ótimo para caracterizar o extrato superior dos rendimentos para as diferentes ocupações. No terceiro ensaio, utilizamos a abordagem não paramétrica das curvas de Pareto generalizadas para caracterizar a desigualdade de rendimentos em todos os níveis da distribuição entre as diferentes profissões. Em síntese, não podemos rejeitar a distinção sugerida entre ocupações escaláveis e não escaláveis através de três diferentes modelos teóricos.

Palavras-chave: Distribuição de renda. Profissões escaláveis. Lei de potência. Distribuição de Pareto. Econofísica.

RESUMO EXPANDIDO

Introdução

Nas últimas décadas, a agenda que analisa problemas econômicos a partir de modelos derivados da física vem ganhando importância. Nesse sentido, técnicas de pesquisa oriundas da mecânica estatística vem sendo cada vez mais aplicadas a temas de interesse relacionados a economia e finanças. O termo econofísica foi então cunhado para fazer referência a essas novas abordagens que buscam integrar assuntos econômico-financeiras a conceitos de física.
Uma das principais circunstâncias que ensejou o surgimento e a difusão da econofísica foi a vinculação dos modelos clássicos de econometria à distribuição normal (Gaussiana) e suas propriedades e a limitação que isso provocava na capacidade explicativa de eventos extremos. Um dos tópicos mais recorrentes nas pesquisas em econofísica é a evidenciação em diversos fenômenos das chamadas leis de potência, que apresentam distribuição de Pareto para as caudas em detrimento do decaimento exponencial característico da distribuição normal.
A dinâmica da distribuição de renda, por exemplo, tem se mostrado um campo de estudo bastante fértil sob o arcabouço da econofísica. No entanto, o atributo da escalabilidade das profissões não vem sendo explorado nessa agenda de pesquisa, embora apresente um enorme potencial para explicar diferenças salariais. Segundo Taleb (2010), algumas atividades podem ser "escaláveis" no tempo, enquanto outras, não. Uma profissão escalável é aquela na qual não se é pago por hora e, portanto, não se está sujeito às limitações do volume de trabalho. Atividades cuja remuneração depende do tempo e esforço empregado muito dificilmente podem tornar as pessoas muito ricas. Nesse tipo de profissão, dita não escalável, a renda do trabalho depende mais de esforços contínuos do que da qualidade das decisões tomadas. Nas ocupações não escaláveis, espera-se um tipo moderado de aleatoriedade e uma distribuição aproximadamente Gaussiana dos rendimentos. Em contrapartida, as ocupações escaláveis caracterizam-se por um tipo intenso de aleatoriedade em que não existe um membro típico. Esse tipo de atividade produz o efeito *winner-takes-almost-all*. Taleb (2010) sugere que as atividades escaláveis estão associadas à geometria Mandelbrotiana e às leis de potência, em detrimento da "curva em formato de sino". Neste caso, uma única observação pode afetar desproporcionalmente o todo. Como as caudas desta distribuição são mais grossas do que as da distribuição Gaussiana, nenhum evento extremo precisa ser excluído e tratado como "outlier". Assim, o presente trabalho se apoiará no conceito de escalabilidade como um fator explicativo para as acentuadas diferenças nos rendimentos do trabalho entre determinadas profissões. A desigualdade na distribuição das rendas entre as ocupações será avaliada a partir de técnicas de pesquisa oriundas da mecânica estatística.

Objetivos

O presente trabalho tem como objetivo geral identificar, descrever e analisar as distribuições de probabilidade da renda do trabalhador brasileiro a partir de características de escalabilidade das profissões. Desse modo, busca-se verificar se ocorre diferença estatística significativa entre os rendimentos de algumas ocupações escaláveis e não escaláveis. Além disso, propõe-se averiguar a presença de leis de potência nas caudas superiores e estimar os parâmetros correspondentes. De maneira geral, o trabalho investiga se as profissões escaláveis apresentam cauda superior mais grossa na distribuição de renda se comparadas com as profissões não escaláveis.

No primeiro ensaio, o objetivo específico é avaliar a eficácia de um modelo de distribuição de renda de duas classes que concilia a lei de Pareto para indivíduos com renda mais alta e a distribuição log-normal para o restante da amostra. No segundo ensaio, o objetivo é calcular o ponto de corte ótimo para caracterizar o estrato superior de rendimentos para as diferentes ocupações e compará-los. No terceiro ensaio, ampliamos o escopo de análise da cauda superior para toda a distribuição de renda através da abordagem não paramétrica das curvas de Pareto generalizadas.

Metodologia

O artigo propõe uma investigação sobre a distribuição da remuneração do trabalhador formal brasileiro, devidamente tabulada por grupos de ocupação, a partir de microdados da Relação Anual de Informações Sociais (RAIS) do ano de 2017. Para agrupar as profissões, foi utilizada a Classificação Brasileira de Ocupações (CBO).

Portanto, partindo da base da RAIS para o ano de 2017 consideramos os microdados de renda nominal média mensal para 12 diferentes ocupações. A partir da base do estado do Rio de Janeiro, coletamos dados para motoboy, escriturário de banco, dentista, arquiteto, ator e jogador de futebol. A partir da base de São Paulo, foram coletados dados de renda para advogado, professor de português do ensino fundamental, designer de moda, locutor de rádio e televisão, artista visual e músico intérprete instrumentista.

Convém destacar que os dados para as ocupações candidatas a escaláveis são indiscutivelmente conservadores. Como a RAIS é um registro administrativo e de âmbito nacional, suas informações revelam as características do mercado de trabalho formal a partir de declarações enviadas pelos empregadores brasileiros. Desse modo, os superastros do futebol que atuam em clubes de outros países, por exemplo, não são considerados nos registros. Da mesma maneira, para os jogadores de futebol que atuam no mercado brasileiro não são computadas rendas oriundas de outras fontes, como contratos publicitários.

Busca-se, a partir desses dados, derivar as distribuições de probabilidade da renda do trabalhador brasileiro a partir de características de escalabilidade das profissões e compará-las. Para tal, será apresentado um conjunto de técnicas estatísticas que permitem diagnosticar e caracterizar a lei de potência, bem como métodos para calcular seus parâmetros.

Resultados e discussão

No primeiro ensaio, o modelo de distribuição de renda de duas classes, que concilia a lei de Pareto para os indivíduos de maior renda e a distribuição log-normal para o restante da amostra, mostrou-se compatível com a categorização interocupacional proposta por Taleb (2010). Nela, espera-se que ocupações não escaláveis apresentem baixa variância de rendimentos, enquanto as escaláveis se caracterizem por grandes desigualdades salariais. Os resultados encontrados nos permitem concluir que, a considerar o modelo de distribuição de renda de duas classes, não podemos rejeitar a hipótese de que ocupações não escaláveis são mais igualitárias que as escaláveis para a porção superior dos dados a partir do ponto de corte ótimo. As distribuições de rendimentos de ocupações escaláveis decaem como lei de potência e apresentam caudas mais pesadas - menores expoentes de Pareto - do que ocupações não escaláveis.

No segundo ensaio, calculamos os expoentes de Pareto das distribuições de renda de doze profissões selecionadas usando OLS e ML. Nosso objetivo foi testar a hipótese de que ocupações não escaláveis são mais igualitárias que as escaláveis. Concluímos que não podemos rejeitar tal hipótese para a porção dos dados de renda entre a mediana e o ponto de corte ótimo. No entanto, existe outro regime de lei de potência acima dos pontos de corte

ótimos onde esses resultados são invertidos. Esse resultado geral diferenciado surge porque os pontos de corte computados dos rendimentos extremos apresentam alta variância e, portanto, uma dinâmica de concentração bastante distinta entre as ocupações.
Os resultados do terceiro ensaio alinham-se às conclusões do segundo ensaio de que na região que caracteriza o topo dos rendimentos as ocupações escaláveis são mais igualitárias do que as não-escaláveis.

Considerações finais

Em vista dos resultados, evidencia-se uma dinâmica de concentração bastante distinta entre os tipos de profissões. Desse modo, a categorização das ocupações de não escalável para escalável depende não apenas do threshold de rendas mais altas escolhido, mas também do regime de escala (scaling) nas caudas. Além de revelar um mecanismo distinto de concentração de rendimentos entre os tipos de profissões, o presente trabalho também pode ser entendido como um novo olhar sobre a dimensão macroeconômica. Com base nas caracterizações intra e interocupacionais da distribuição de rendimentos, fornecemos novos elementos e perspectivas para abordar a desigualdade de renda em nível agregado.

Palavras-chave: Distribuição de renda. Profissões escaláveis. Lei de potência. Distribuição de Pareto. Econofísica.

ABSTRACT

Research techniques derived from statistical physics have proved to be a very powerful analysis tool to explain the dynamics of income distribution. However, the attribute of scalability of professions has not been explored in this research agenda, although it has huge potential to explain salary differences. It has been suggested that occupations where one is paid by the hour are not scalable, while scalable occupations allow one to make more money without an equivalent increase in labor and time. Non-scalable occupations are expected to have low income variance, whereas scalable ones show large income inequalities. This study examines the evidence for this suggested distinction using personal earnings microdata for twelve candidate occupations of both types, scalable and not. Generally, we find the upper tails of all distributions decay as power laws. In the first essay, we tested the effectiveness of a two-class income distribution model that reconciles Pareto's law for individuals with higher incomes and the log-normal distribution for the rest of the sample. In the second essay, we computed the optimal cut-off point to characterize the upper stratum of earnings for the different occupations. In the third essay, we used the non-parametric approach of generalized Pareto curves to characterize earnings inequality at all levels of the distribution between different occupations. In summary, we cannot reject the suggested distinction between scalable and non-scalable occupations through three different theoretical models.

Keywords: Income distribution. Scalable professions. Power law. Pareto distribution. Econophysics.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ABNT Associação Brasileira de Normas Técnicas

CBO Classificação Brasileira de Ocupações

IBGE Instituto Brasileiro de Geografia e Estatística

ML Maximum Likelihood

KS Kolmogorov-Smirnov

OLS Ordinary Least Squares

RAIS Relação Anual de Informações Sociais

p90 Percentil 90

p95 Percentil 95

CONTENTS

1 INTRODUCTION

1.1 THEME

In recent decades, different approaches to income distribution have played a fundamental role in economic models. At the same time, the research agenda that analyzes economic problems from concepts derived from physics is gaining importance. In this sense, statistical mechanics instruments have been increasingly applied to research topics in economics and finance.

The dynamics of income distribution, for example, has proved to be a very fertile field of study under the framework of econophysics. However, the attribute of scalability of professions has not been explored in this research agenda, although it has enormous potential to explain salary differences.

Thus, the present thesis will be based on the concept of scalability as an explanatory factor for the marked differences in labor earnings between certain professions. The inequality in the distribution of income between occupations will be evaluated using research techniques from statistical mechanics.

1.2 OBJECTIVES

1.2.1 General Objectives

The thesis has as general objective to identify, describe and analyze the probability distributions of the Brazilian worker's income from the scalability characteristics of the professions. Thus, we seek to verify whether there is a statistically significant difference between the income of scalable and non-scalable occupations.

Furthermore, it is proposed to investigate the presence of power laws in the upper tails and estimate the corresponding parameters. In general, the work investigates whether scalable professions have a thicker upper tail in the income distribution compared to non-scalable professions.

1.2.2 Specific Objectives

In its first essay, the thesis evaluates the effectiveness of a two-class income distribution model that reconciles Pareto's law for individuals with higher incomes and the log-normal distribution for the rest of the sample.

In the second essay, we computed the optimal cut-off point in order to characterize the upper stratum of earnings for the different occupations and compare them.

In the third essay, we broaden the scope of analysis from the upper tail to the entire income distribution through the non-parametric approach of the generalized Pareto curves.

2 THEORETICAL REFERENCE

Economists have long historically and empirically analyzed income and wealth inequalities. Several prominent researchers such as Smith, Marshall, Pareto, Friedman and Kuznets have made great contributions to consolidate the study of the dynamics of the distribution of earnings and assets as one of the main fields of research in social sciences and political economy.

With regard to the positive perspective of income distribution, the theoretical approach of classical economists focused mainly on the functional distribution of income, that is, on the way in which incomes are distributed among the factors of production.

In order to carry out an analysis with a high level of aggregation, the classical approach to the functional distribution of income assumed the premise that production factors are homogeneous. In fact, the assumption of homogeneity was admittedly a very strong theoretical abstraction, especially when it came to the distribution of income from the labor factor. It was always very evident that salaries were very different between different occupations. In principle, there could be two reasons for this.

First, wage differences can be caused by competitive forces. Second, they can be caused by the absence of competition, whether by particular restrictions or government regulations.

Smith's (1776) view of wage differentials became known as the theory of compensating variances. His most general idea is that wage incomes will reflect the particular circumstances of each of the different professions. Thus, for any particular job class, these circumstances will determine whether the salary will be above or below the average for all professions.

In this sense, Smith (1776) mentioned some elements that could cause inequality between wages. The first one would be the ease or difficulty inherent in this work. Furthermore, the attributes of honor and social repulsion of occupations would also play a key role.

According to the eminent author, some professions are particularly honorable and, as honor is related to reward, salaries are relatively lower. Other professions, on the contrary, embody the general feeling of disgrace, are carried out in unhealthy environments or carry a high degree of danger. In these cases, the effect on wages is the opposite, that is, the corresponding earnings are relatively higher.

For Smith (1776), the great variability of wages is also related to how difficult and expensive it is to learn the profession, with the constancy or inconstancy of employment and with the amount of trust placed in the worker.

The last cause of wage inequality according to Smith (1776) is the probability of success in the profession. In his best-known example, if an individual practices to become a shoemaker, he will almost certainly be able to make a living making shoes. However, if someone is brought up to become a lawyer, only one in 20 will be able to do well enough to live up to that occupation.

In Smith's (1776) view, professions such as lawyer function as a kind of lottery. So, as there are few winning tickets, they should receive very high prizes. However, wage differences of this nature would, in fact, answer less than rational considerations of probabilities, given that most people, especially young people, have a tendency to overestimate the probability of success.

In Mill's (1848) view, Smith's theory only has some explanatory power for the case of perfect competition with jobs of the same category and filled by similar people. However, this case would be far from the reality of labor markets.

Thus, Mill (1848) believed that Smith's hypothesis that wages tended to increase with the net disadvantages associated with different occupations was wrong. For him, otherwise, the difficulties and the corresponding income maintained an inverse relationship between them. Thus:

> "The really exhausting and the really repulsive labours, instead of being better paid than others, are almost invariably paid the worst of all, because performed by those who have no choice… The more revolting the occupation, the more certain it is to receive the minimum of remuneration, because it devolves on the most helpless and degraded, on those who from squalid poverty, or from want of skill and education, are rejected from all other employments." (MILL, 1848; 1965, p. 475)

The marginalist revolution, which gave rise to the neoclassical school, enunciated the new foundations of conventional economic theory. Unlike classical authors, marginalist economists based their analyzes on the individual behavior of economic agents. This evolution in the conception of economic ideas was only possible thanks to advances in the use of optimization theories and other tools of differential calculus.

In this sense, the equality between the value of marginal productivity and the price of the corresponding factor of production, arising from the problem of maximizing profit, becomes one of the main foundations of neoclassical economics. Walras (1874-1877; 1954), one of the main exponents of marginalism, also emphasized that a theory about the average wage rate – which, according to him, was the core of classical thought – would not be very useful. According to the eminent author, the analysis of wages should be based on a disaggregated view of the labor market. Thus, earnings from work should reflect specific conditions of occupation.

Like Walras (1874-1877; 1954), Marshall (1890; 1920) also believed that analyzes of the general wage rate were misleading since:

> "[…] in fact there is no such thing in modern civilization as a general rate of wages. Each of a hundred or more groups of workers has its own wage problem, its own set of special causes, natural and artificial, controlling the supply-price, and limiting the number of its members; each has its own demand-price governed by the need that other agents of production have of its services." (MARSHALL, 1890; 1920, p. 533)

In this way, a contrasting perspective can be seen between the marginalist view and the classical authors. While the classical school based its discussion on the general rate of wages (though it later added an ad hoc discussion of wage differentials), the marginalists analyzed wage incomes within multiple (albeit interrelated) labor markets.

In the 20th century, the neoclassical ideas were rescued by economists who would give rise to the theory of human capital, according to which items such as education and technical courses are also considered capital. Thus, spending on education and professional training can be classified as investments, and the decision-making criteria are similar to those for investments in physical capital.

Although the theory of human capital emerged with the work of Schultz, its conceptions were extensively developed and popularized by Becker. In this sense, it is fair to say that Becker (1964) inaugurated an extremely influential new field of research, even addressing issues related to income distribution. For the author, the quality of labor reflects the stock of cognitive skills, which can be improved through investment in human capital, thus making work more productive and improving workers' earnings.

According to Becker and Chiswick (1966), at the individual level, the amount to be invested in human capital is determined by the intersection between the supply and demand curves, or more specifically, between the marginal benefit and marginal cost curves. Empirically, it is expected that the supply and demand curves show great variety among individuals.

While the position of the supply curves may reflect parents' income and wealth and access to capital markets, the different demand curves may represent individual characteristics, such as the potential for skill development and risk behavior pattern.

The balance of risk in the distribution of income between occupations represented a fundamental element in compensating wage differences. In choosing between a safe and a risky occupation (shoemaker and lawyer, respectively, in Smith's classic example), the expected wage income in the risky occupation would have to be higher than in the safe profession to compensate individuals for their additional risk burden.

To the extent that individuals correctly assessed probabilities, these *ex ante* expectations would be translated into *ex post* income inequality. Thus, the salary of lawyers would tend to have a higher average, although with greater variability compared to the earnings of shoemakers.

The possibility of formally modeling individual choice in risk scenarios was driven by the axioms of expected utility theory developed by von Neumann and Morgenstern (1947). The theory of income distribution was one of the first fields of research to adopt this theoretical tool.

The seminal contributions in this field were given by Friedman (1953), who interpreted income distribution as the result of rational choice under uncertainty. In this sense, as much as individuals have *ex ante* equal opportunities, the income lotteries in which they engage incur an unequal distribution of results, as some will benefit *ex post* with high incomes and others will be included in low-income groups.

For Friedman (1953), the format of the utility function will generate a given income distribution, consistent with the observed patterns. The renowned author also argued that individuals, as members of a democratic society, will be motivated to introduce redistributive mechanisms that protect them against the socioeconomic consequences of the most adverse outcomes.

According to this view, therefore, both income inequality and redistributive policies emerge as a result of the free choice of individuals in a situation of equal opportunities and reflect their attitude towards risk. The less risk-averse individuals are, the greater will be the income inequality in society.

Although the marginalist revolution presented a set of quite consistent theoretical propositions regarding the personal distribution of income, the end of the 19th century saw the emergence of a more inductive view of the phenomenon, based no longer on an a priori theory, but on statistical inference. The pioneering contribution of this new conception can be attributed to Pareto (1897), whose work caused much discussion and controversy for several decades after its initial publication.

The prototype that later came to be known as Pareto's law did not derive from a theoretical model. Instead, it was based on a detailed study of income statistics for various countries and periods. Data analysis led Pareto to the hypothesis that all statistical distributions of income have a common form which can be characterized as follows.

Suppose we make a list of all the society's incomes, from lowest to highest. From the median of income, we know that 50% of income earners have an income above the median. We then move to an income level 1% higher than the median and ask what percentage of the population has an income above that level.

Obviously, the percentage is less than 50, but how much less? Pareto found the response to be 1.5%; in other words, as the income level increases by 1%, the number of individuals with income above that level falls by 1.5%. In general, in mathematical terms Pareto wrote his law as $\log N = \log A - \alpha \log y$, where N is the number of individuals with a minimum income of y and A is a parameter that reflects the size of the population. α is the Pareto constant that he estimated to be approximately equal to 1.5.

The relationship has the interesting property that the average income of those whose income is greater than y will be equal to $\alpha / (\alpha-1)$ times y. Thus, assuming again that $\alpha = 1.5$, the average income of those with income above 10,000 units of money should equal 30,000 units of money.

The tradition established by Pareto's work of looking for regularities or empirical laws in the distribution of income was continued by several later writers.

The seminal work by Gibrat (1931), for example, inaugurates the view that income data fit sufficiently well with the lognormal distribution.

Roy (1950) also claimed that the observed income distributions could be reasonably approximated by the lognormal distribution. According to him, "there must be some rational explanation for the fact that all income distributions have similar shapes" (ROY, 1950, p. 490). He tried to discover this explanation by studying several industrial cases where workers performed a standard and identical task and where individual output was easy to measure. To the extent that people are paid according to output, this result can go a long way in explaining the distribution of earnings in terms of the distribution of individual skills.

Roy (1951) studies the theoretical case of a "primitive" society in which people can choose to work in two or more occupations and where their abilities differ between occupations. He then discussed how different skill correlations give rise to different statistical distributions of earnings (always assuming earnings are proportional to output), emphasizing the central role played by the lognormal distribution.

Recently, econophysicists discovered that only upper stratum incomes follow the Paretian distribution, with the lower stratum being distributed as a two-parameter lognormal (CLEMENTI; GALLEGATI, 2005). Mandelbrot (1960) also points out that Pareto's law applies only asymptotically to the extremes of the upper tail of distributions. In this sense, the Pareto Type I model fits well with higher income data (ATKINSON et al., 2014).

The various statistical approaches to the study of income distribution involve some controversy. From attempts to rationalize the observed income distribution, these approaches use some stylized facts or assumptions about income generation to explain the observed patterns. The controversy of this type of analysis is no longer supported by works such as Atkinson (1970) and Sen (1973), since it has been established that any particular index of inequality is implicitly based on some ethical judgment about the nature of inequality.

When it comes to the need to reduce inequalities, we certainly do not seek to eliminate all differences in terms of economic results. Differences in economic rewards are, in a sense, inherent in market economies. In recent literature, the factors that determine the discrepancy of economic gains are divided into inequality of opportunities and inequality of results.

Inequality of opportunities is related to socioeconomic circumstances (social class in which the family is inserted and inheritance, for example) and are beyond personal control. When factors associated with these socioeconomic circumstances do not interfere with the achievement of rewards, we say that there is equality of opportunity. In this way, equality of opportunity is an *ex ante* concept, as all agents start from the same point and dispute the allocation of scarce resources with equal chances of success.

The inequality of results is associated with the degree of effort made by the individual in the demand for economic rewards. It is, therefore, an *ex post* concept and imputed to the agent, in the sense that he is responsible for his diligence.

However, it is essential to differentiate between competitive and non-competitive equal opportunities. According to Atkinson (2015), non-competitive equality must ensure that individuals have the same opportunity to carry out their independent life projects. All people should thus have a chance of becoming a successful lawyer.

Alternatively, competitive equality of opportunity states that individuals must have an equal chance of becoming a successful lawyer or a famous football athlete. Thus, it is verified the presence of unequal rewards *ex post*, which leads us to a more detailed investigation regarding the role of inequality in results.

According to Atkinson (2015), “it is the existence of a highly unequal distribution of prizes that makes us give so much importance to ensuring that the race is fair. And the prize structure is, to a large extent, socially constructed.” In this way, our socioeconomic apparatus determines whether the winners in a given market earn R$ 2,000 or R$ 60,000 per month.

However, the distribution of rewards between different types of occupation does not follow a purely stochastic process. Several factors can influence the great inequality in the division of premiums between professions. Among these factors, we can highlight the scalability attribute.

According to Taleb (2010), some activities can be “scalable” in time, while others cannot. A scalable profession is one in which you are not paid by the hour and therefore not subject to workload limitations. In the author's conception:

> “A scalable profession is good only if you are successful; they are more competitive, produce monstrous inequalities, and are far more random, with huge disparities between efforts and rewards—a few can take a large share of the pie, leaving others out entirely at no fault of their own. One category of profession is driven by the mediocre, the average, and the middle-of-the-road. In it, the mediocre is collectively consequential. The other has either giants or dwarves—more precisely, a very small number of giants and a huge number of dwarves.” (TALEB, 2010, p. 28)

Activities whose remuneration depends on the time and effort employed can hardly make people very rich. In this type of profession, which is said to be non-scalable, income from work depends more on continuous efforts than on the quality of decisions taken.

In non-scalable professions, moderate or type 1 randomness is perceived, in which the most typical member is mediocre. In this type of activity, earnings have an approximately Gaussian distribution. Here, in a large sample, no single event significantly alters the whole. The greatest observation of the sample is impressive, but it is insignificant for the total sample. The extreme event can be excluded without too many consequences and considered an “outlier”.

In contrast, scalable professions are characterized by intense randomness or type 2, in which there is no typical member (the most typical is giant or dwarf). This type of activity produces the winner-takes-almost-all effect. Taleb (2010) suggests that scalable activities are associated with Mandelbrotian geometry and power laws, to the detriment of the “bell-shaped curve”. In this case, a single observation can disproportionately affect the whole. As the tails of this distribution are thicker than those of the Gaussian bell-shaped distribution, no extreme events need be excluded and treated as “outlier”.

3 METHODOLOGY

3.1 MATERIALS

The article proposes an investigation on the distribution of Brazilian workers' remuneration, duly tabulated by occupation groups, based on microdata from the Annual Social Information Report (RAIS) for 2017.

To group the professions, the Brazilian Classification of Occupations (CBO) was used. The CBO is the standardizing document for the classification, naming and coding of titles and contents of occupations in the Brazilian labor market. The enumerative function of the CBO is used in administrative records such as the Annual Social Information Report – RAIS.

Since its publication at the beginning of the last quarter of the last century, the CBO has undergone punctual updates, without structural and methodological changes. The international classification made public in 1988 under the acronym CIUO 88 in Spanish – ISCO 88 and CITP 88, in English and French, respectively – introduced new criteria for aggregation of occupations.

Therefore, starting from the RAIS base for the year 2017, we considered the monthly average nominal income microdata for 12 different occupations. From the base of the state of Rio de Janeiro, we collected data for motorcycle messenger, bank clerk, dentist, architect, actor and soccer player. From the base of São Paulo, income data were collected for lawyers, elementary school Portuguese teachers, fashion designers, radio and television broadcasters, visual artists and music performers.

It should be noted that the data for occupations that are candidates for scalables are indisputably conservative. As RAIS is an administrative register with a national scope, its information reveals the characteristics of the formal labor market based on statements sent by Brazilian employers. Thus, football superstars who play in international clubs, for example, are not considered in the records. In the same way, for soccer players who work in the Brazilian market, income from other sources, such as advertising contracts, is not computed.

In addition, the choice of professions mainly met the criterion of availability, given that for many occupations, especially those candidates for scalable, there was not enough data. One possible explanation for the scarcity of earnings data for many scalable occupations is that many of these individuals are part of the informal market, work as freelancers or are registered with an adverse occupation code.

Based on these conservative data, the aim is to derive the probability distributions of the Brazilian worker's income from the scalability characteristics of these professions and compare them. For that, a set of statistical techniques that allow to diagnose and characterize the power law will be presented, as well as methods to calculate its parameters.

3.2 METHODS

3.2.1 Descriptive statistics and visual inspection

In the present work, descriptive statistics are presented, such as percentiles, mean, standard deviation, asymmetry and kurtosis of earnings tabulated by occupation. The histograms of income distributions are also presented. A histogram is a frequency graph intended to illustrate how a particular sample or population of data is distributed.

Afterwards, the normal quantile plots of remuneration are displayed. The QQ plot, or quantile-quantile plot, is a graphical tool to help us assess whether a data set comes from some theoretical distribution, such as normal or exponential. For example, if we run a statistical analysis that assumes our dependent variable is normally distributed, we can use a normal QQ plot to verify this assumption. It's just a visual check.

3.2.2 Power law and estimation of Pareto exponent

The probability density function (PDF) for a Pareto Type I random variable is:

$$p(x) = \frac{\alpha x_{min}^{\alpha}}{x^{\alpha+1}} \qquad (1)$$

It is very important to also consider the cumulative distribution function or CDF of a power-law distributed variable, which we denote P(x) and which is defined as P(x) = Pr(X ≤ x). For example, in the continuous case:

$$P(x) = \int_{x_{min}}^{x} p(z)dz = 1 - \left(\frac{x_{min}}{x}\right)^{\alpha} \qquad (2)$$

The complementary cumulative distribution function (CCDF), or Survivor function, in a Pareto type I model shows the fraction of a given population with incomes greater than x - that is, S(x) = Pr(X > x) - and is given by:

$$S(x) = \int_{x}^{\infty} p(z)dz = 1 - P(x) = \left(\frac{x}{x_{min}}\right)^{-\alpha} \qquad (3)$$

where $x \geq x_{min} > 0$, and $x_{min} > 0$ is the lower income limit. The α parameter is the shape parameter ("tail index") that describes the weight of the right tail of the distribution, with smaller values corresponding to heavier tails. The k-th moment exists only if $k < \alpha$.

An important power law graphical representation mechanism derived from equation (3) was known as Zipf plots. These are plots of the logarithm of the Survivor function against the logarithms of income (for incomes in ascending order and greater than x_{min}). In the event of a power law, the format of these plots generates straight line intervals with an inclination equal to -α.

There are several methods for calculating the scaling parameter of a power law. For example, the exponent of a type I Pareto distribution can be estimated by an ordinary least squares regression of the logarithm of the complementary cumulative distribution function (or Survivor function) over the logarithm of income and a constant term.

An ordinary least squares estimate is consistent, but the standard error is imprecise because the positive autocorrelation in the residuals from income rankings is neglected (JENKINS, 2017). An alternative is maximum likelihood (ML) estimation. For the continuous case of a Pareto type I model, the maximum likelihood estimator (HILL, 1975) for the shape coefficient is:

$$\hat{\alpha} = n\left[\sum_{i=1}^{n} \ln \frac{x_i}{x_{min}}\right]^{-1} \qquad (4)$$

where x_i, i = 1. . . n are the observed values of x such that $x_i \geq x_{min}$.

This estimator produces a consistent standard error, but is susceptible to bias in the presence of extremely high incomes (JENKINS, 2017). This is a serious limitation given the nature of our data. However, as our dataset is conservative - as explained above, superstars operating in markets beyond national borders are absent - we also provide ML estimates.

In addition, both the ordinary least squares (OLS) and the maximum likelihood (ML) estimators are biased in small samples (JENKINS, 2017). As the sample size in our datasets for each occupation is not large enough, we consider here the OLS estimator proposed by Gabaix and Ibragimov (2011) that solves the small sample problem and produces a consistent standard error, given by:

$$\log\left(Rank - \frac{1}{2}\right) = a - b\log(x) \qquad (5)$$

where x in this case is income. The change through the subtraction of 1/2 is optimal, cancels the bias and the standard error of α is not the standard error of the OLS estimator, but is, asymptotically, equal to $\left(\frac{2}{n}\right)^{\frac{1}{2}}\alpha$ (GABAIX; IBRAGIMOV, 2011).

Additionally, the graph of the log of rank-1/2 versus the log of income in straight line form suggests that the decay of the upper tail of the income distribution follows a power law.

As can be inferred, there is also a type II Pareto model that adds an extra parameter and sometimes produces a better fit to the data. However, in most cases, the improvement in the quality of the fit is negligible and this must be balanced against the greater simplicity of the Pareto type I model. Furthermore, the Pareto type II model collapses to the type I after proper parameterization. Here, we do not consider the Pareto type II model.

4 ESSAY 1: TWO-CLASS EARNINGS DISTRIBUTION MODEL

4.1 INTRODUCTION

In this first essay, we tested the effectiveness of a two-class income distribution model, which reconciles Pareto's law for individuals with higher incomes and the log-normal distribution for the rest of the sample, considering the attribute of scalability of occupations.

4.2 RELATED LITERATURE

In the literature related to econophysics, one can define income distribution as the probability *P(x)dx* that, in the "equilibrium" or "steady state" of the system, a randomly chosen person has income between x and $x + dx$.

In his seminal work, Pareto (1897) found that the upper portion of the income distribution follows a power law:

$$P(x) \sim x^{-\alpha} \qquad (6)$$

Later, Gibrat (1931) found that despite Pareto's law being valid only for the upper tail, the average income interval is well represented by the log-normal probability density:

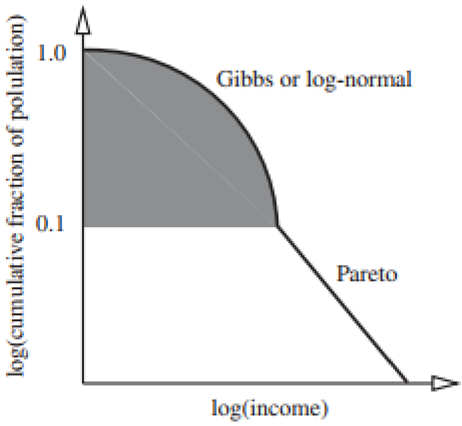
$$P(x) \sim \frac{1}{x\sqrt{2\pi\sigma^2}} exp\left\{-\frac{log^2\left(\frac{x}{x_0}\right)}{2\sigma^2}\right\} \qquad (7)$$

where x_0 is the mean and σ^2 is the variance.

Since then, several studies have shown that income distributions maintain some stable and robust properties (YAKOVENKO; ROSSER, 2009). Thus, this type of regularity observed in income distribution models suggests a kind of “natural” law of economics (CHAKRABARTI ET AL, 2013).

Empirically, a series of works reveals that in the graphical representation of the logarithm of the accumulated distribution of income against the logarithm of income, at least 90% of the sample fits well into the lognormal format (or Gibbs distribution). The remainder of the sample, which denotes the highest earnings, fits well with a power law. The following figure illustrates this pattern.

Figure 1. Two-Class Income Distribution



Fonte: Chakrabarti et al., 2013

However, the attribute of scalability of professions has not been explored in this research agenda, although it has enormous potential to explain salary differences.

Based on this finding, an attempt was made to fully derive the probability distributions of the Brazilian worker's income in a two-class model based on the scalability characteristics of the professions.

4.3 METHODOLOGY

We estimate the following model to express the probability of an individual having an individual income greater than or equal to *x*:

$$P(x' \geq x) \cong \begin{cases} \Phi\left(\frac{(\ln x) - \mu}{\sigma}\right), & x \leq x_{min} \\ \left(\frac{x}{x_{min}}\right)^{-\alpha}, & x \geq x_{min} \end{cases} \quad (8)$$

where Φ is the cumulative distribution function of a standard normal distribution N(0,1).

As the complementary cumulative distribution function (CCDF) plotted on a log-log scale presented a decreasing linear slope, the presence of a power law in the upper tail was considered.

The Kolmogorov-Smirnov statistic was used for the task of drawing an optimal cut-off point between the middle stratum income and the upper tail (CLAUSET; SHALIZI; NEWMAN, 2009). This statistic provides the maximum distance D between the empirical and adjusted cumulative distribution functions:

$$D = \max_{x \geq x_{min}} |S(x) - F(x)| \qquad (9)$$

where *S(x)* is the empirical complementary cumulative distribution function (CCDF) and *F(x)* is the theoretical CCDF of the power law model that best fits the sample data for $x \geq x_{min}$. For each possible choice of x_{min}, MATLAB's plfit.m function (available at http://www.santafe.edu/~aaronc/powerlaws) estimated "alpha" using the maximum likelihood method and calculated the Kolmogorov-Smirnov D statistic.

Subsequently, estimates of x_{min} whose value gives the smallest D statistic were selected. The optimal cut-off point is therefore the value of x_{min} that minimizes D - to a sufficiently large p-value, say, $p > 0.05$.

In the next section, the graphic results of the best fit procedure of the CCDF's functions plotted in log-log scale are displayed.

In addition, the two parameters characteristic of the log-normal distribution were estimated. Generally, the mean μ and the variance σ^2 of log(x) are used to specify it. However, there are advantages to using "back-transformed" values, that is, values in terms of x (LIMPERT; STAHEL; ABBT, 2001):

$$\mu^* = e^{\mu} \qquad (10)$$

$$\sigma^* = e^{\sigma} \qquad (11)$$

Thus, μ^* represents the median of the log-normal distribution and also the geometric mean of the distribution in terms of the original data. The parameter σ^*, the so-called geometric standard deviation, determines the shape of the distribution. Therefore, as μ^* and σ^* are in the original measurement units, they are easier to interpret.

4.4 RESULTS

4.4.1 Visual inspections

A QQ chart is a scatter plot created by plotting two sets of quantiles against each other. If the two sets of quantiles come from the same distribution, we should see the points forming a line that is approximately straight. QQ charts take sample data, sort it in ascending order, and then plot it against calculated quantiles from a theoretical distribution.

From the normal quantile plots, it can be seen that the earnings from non-scalable professions such as bank clerk, motorcycle courier, Portuguese teacher, lawyer and dentist had quantiles compatible with the quantiles of the normal distribution, with the exception of the values of the tails lower and upper. The histograms for these occupations display a format compatible with the lognormal distribution.

Likewise, it is clear that the normal quantile plots of the remuneration of scalable professions such as actor, football player, visual artist and musician performing instrumentalist showed incompatibility with the quantiles of the normal distribution at all levels, even generating a certain distance in values around the median. Histograms for these occupations show thicker tails than the Gaussian distribution.

Because of these initial results based on visual inspections, we will test the empirical adherence of the two-class income distribution model to the data.

Figure 2. Normal quantile plots

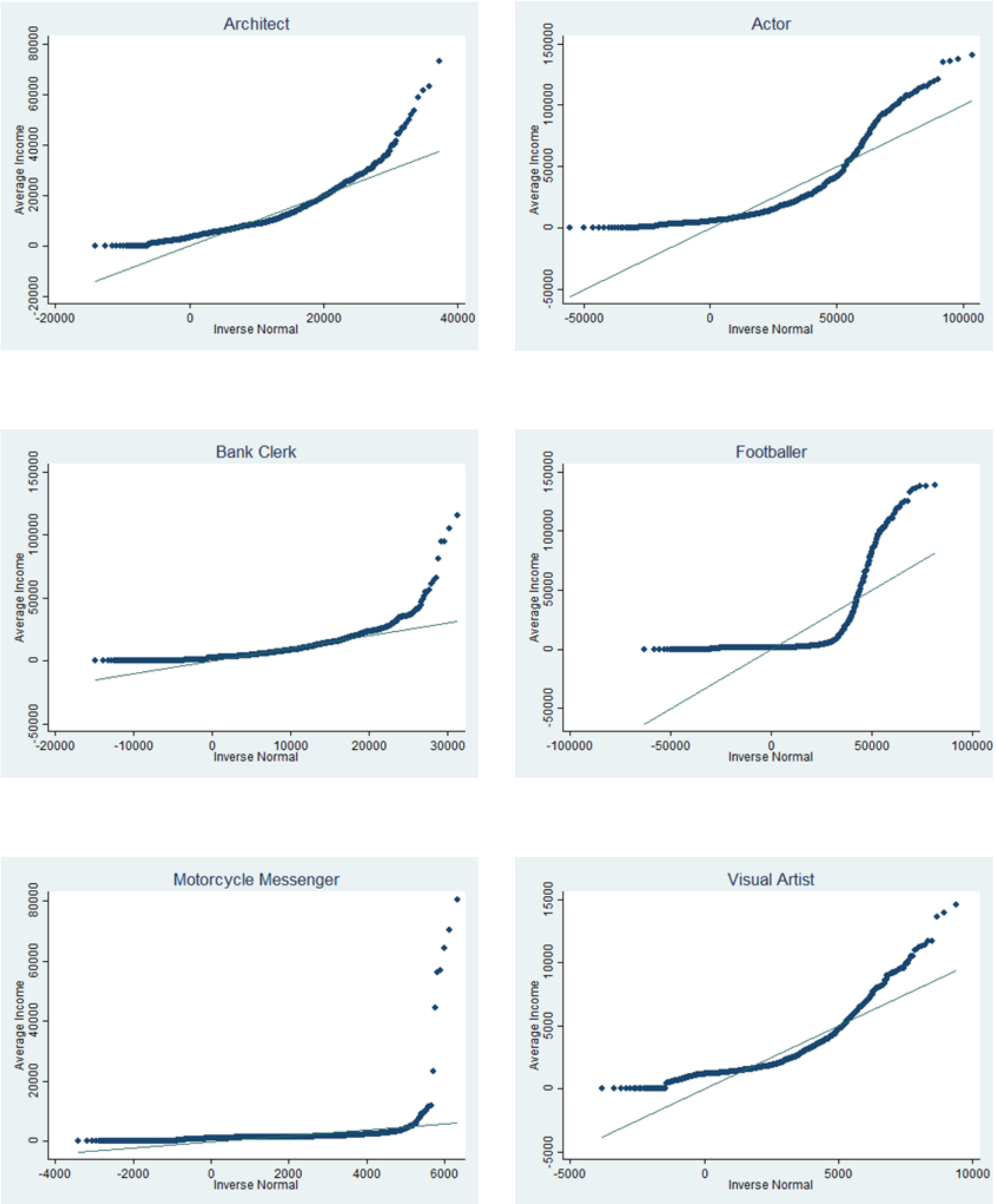


Figure 2. Normal quantile plots – Continue

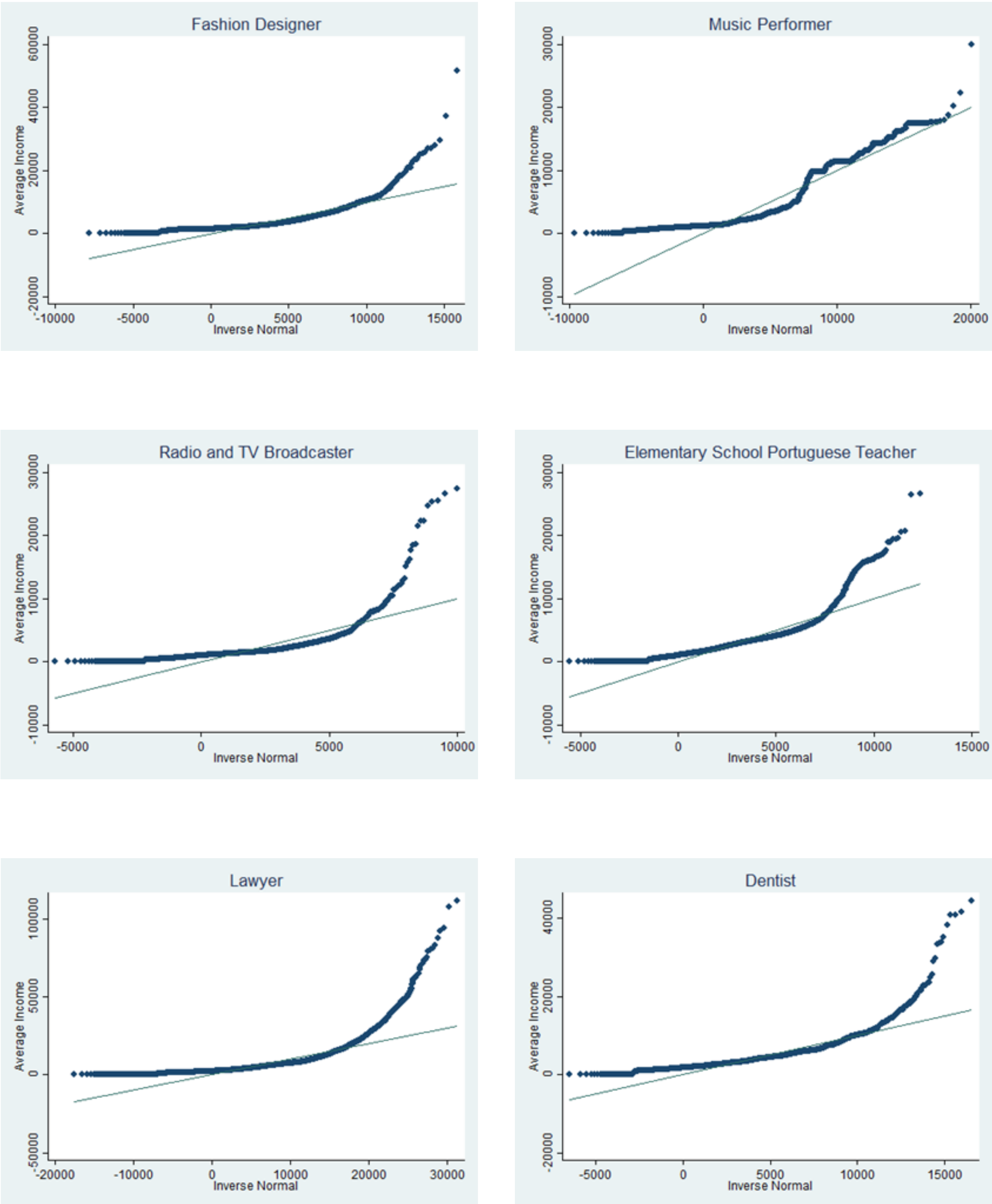


Figure 3. Histograms

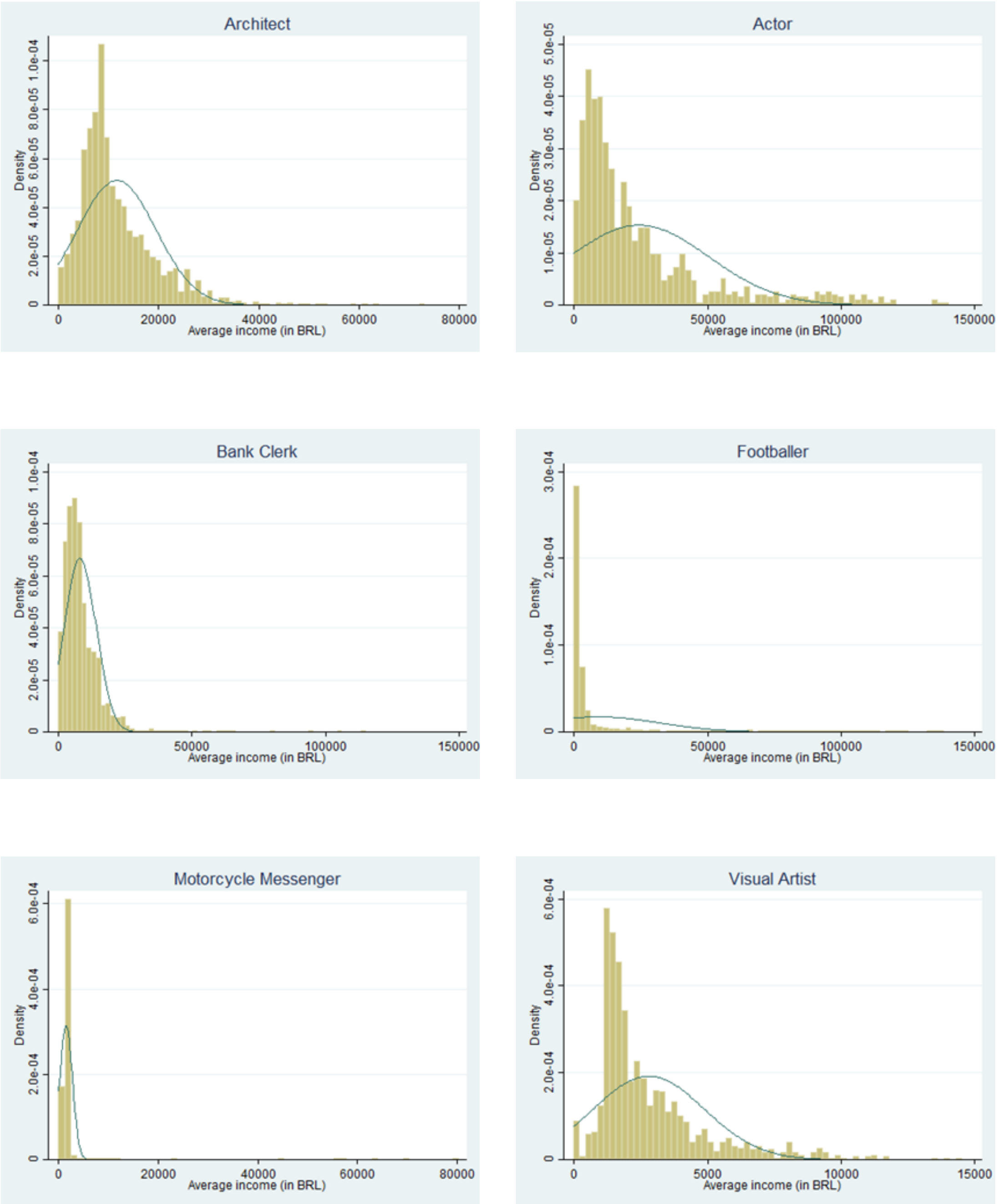
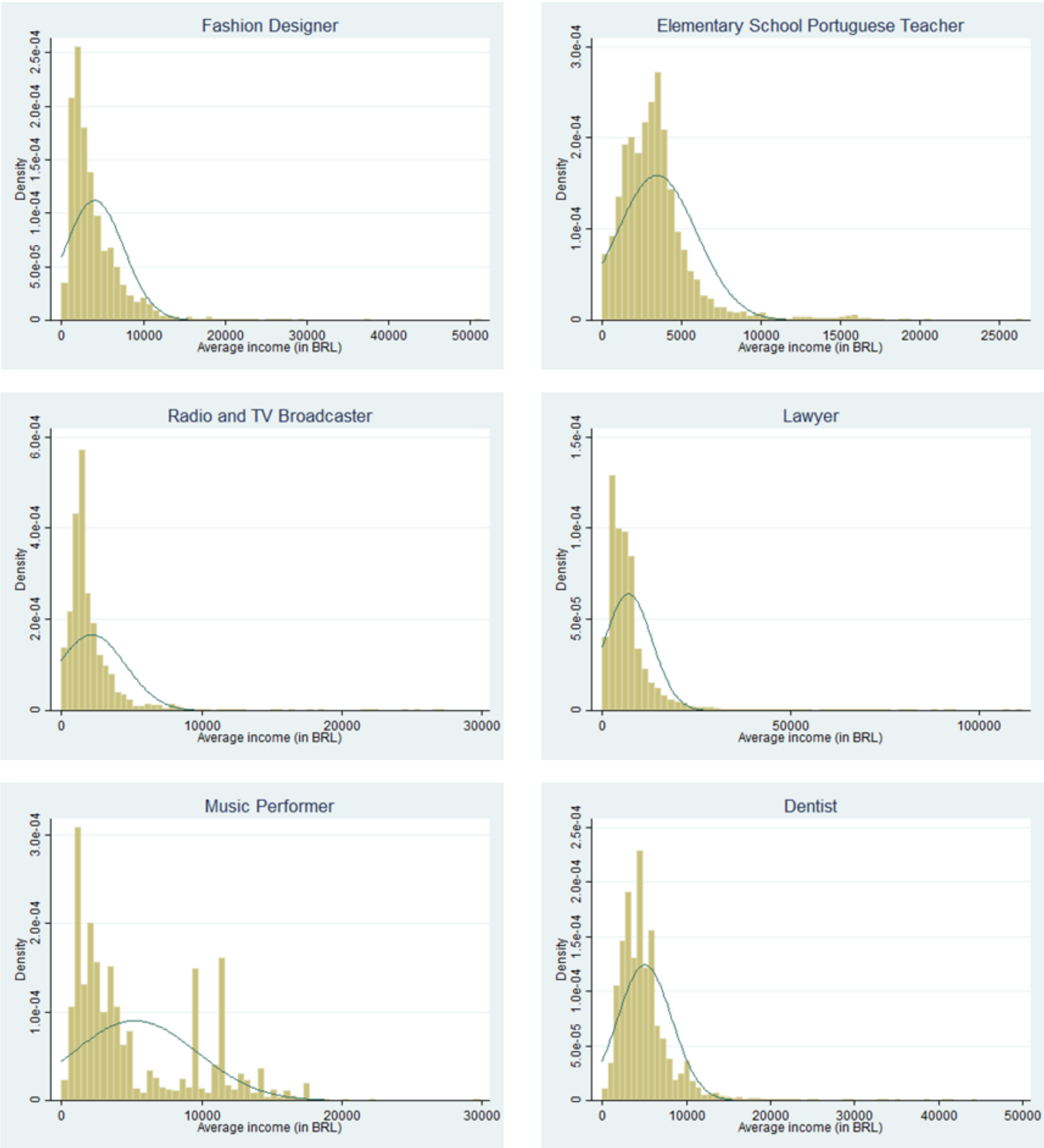


Figure 3. Histograms – Continue



4.4.2 Two-class income distribution model estimation

The results of the two-class income distribution model specified on section methodology can be seen in the table below.

Table 1. Results of the two-class income distribution model

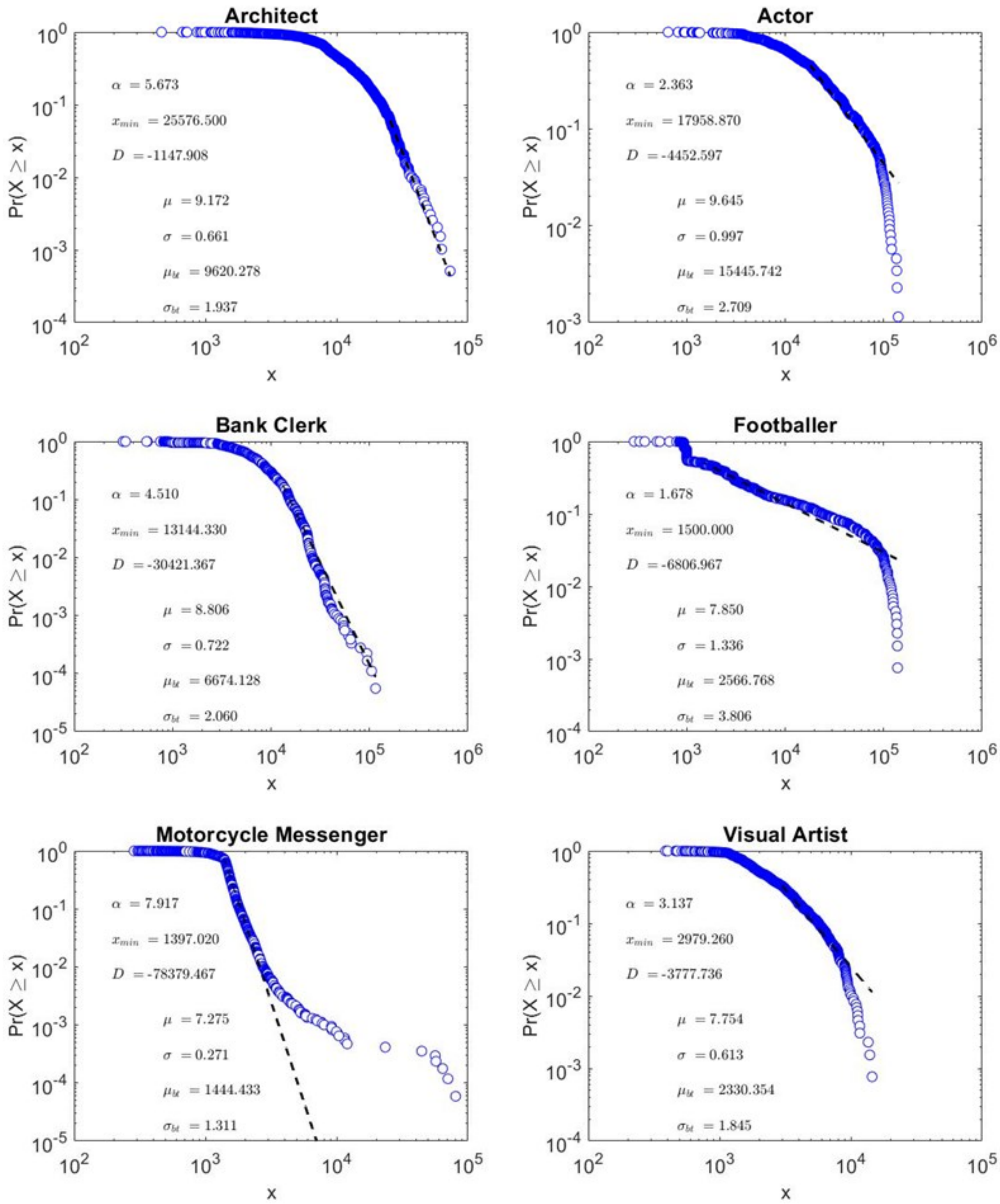
	Power Law			Log Normal PDF	
Occupation	x_{min}	α	D	μ^*	σ^*
Footballer	1500	1.6775	-6807	2566.8	3.8057
Actor	17959	2.3629	-4452.6	15446	2.7093
Radio Tv Broadcaster	2650.7	3.0046	-3620.6	1707.3	1.9894
Visual Artist	2979.3	3.1372	-3777.7	2330.4	1.8452
Fashion Designer	5376.9	3.4256	-4821	3188.1	1.9678
Portuguese Teacher	3426.1	3.7114	-19292	2873.5	1.9078
Lawyer	26722	4.3583	-4120.8	5345.8	2.0428
Bank Clerk	13144	4.5100	-30421	6674.1	2.0595
Dentist	13317	4.5731	-1196.1	4369.5	1.7259
Architect	25577	5.6733	-1147.9	9620.3	1.9369
Music Performer	17777	6.3849	-54.5407	3595.1	2.4713
Motorcycle Messenger	1397	7.9171	-78379	1444.4	1.3115

Source: Own elaboration based on RAIS 2017 microdata

Thus, the two-class income distribution model, which reconciles Pareto's law for higher-income individuals and the log-normal distribution for the rest of the sample, was compatible with the interoccupational categorization proposed by Taleb (2010). In it, non-scalable occupations are expected to have low income variance, while scalable ones are characterized by large wage inequalities.

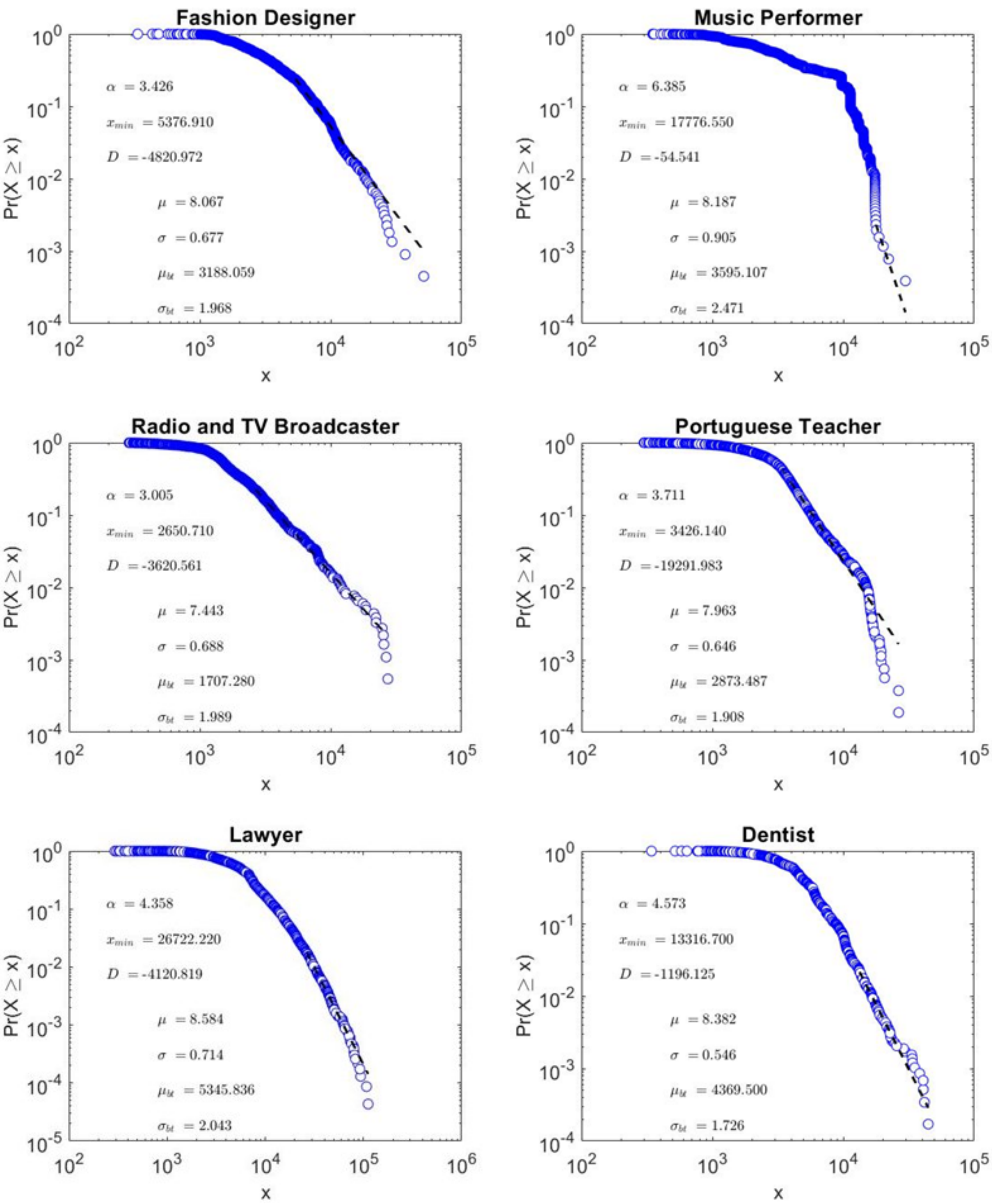
Insofar as the Pareto coefficient α is typically interpreted as an inverse measure of concentration at the top of incomes, the results expressed in Table 1 show less concentrated upper incomes for non-scalable professions such as motorcycle messenger, architect, dentist and bank clerk. In contrast, we can observe heavier upper tails for scalable occupations such as football player, actor, radio and tv broadcaster and visual artist.

Figure 4. Complementary Cumulative Density Function (y axis) and earnings (x axis) in log scale[1]



[1] The dashed line is the power law distribution tail computed.

Figure 4. Complementary Cumulative Density Function (y axis) and earnings (x axis) in log scale - Continue



4.5 CONCLUSIONS

The results found allow us to conclude that, considering the two-class income distribution model, we cannot reject the hypothesis that non-scalable occupations are more egalitarian than those scalable for the upper portion of the data from the optimal cutoff point.

The income distributions of scalable occupations decay as a power law and have heavier tails - smaller Pareto exponents - than non-scalable occupations.

5 ESSAY 2: TOP EARNINGS OF SCALABLE VS. NON-SCALABLE OCCUPATIONS

5.1 INTRODUCTION

Usually, the most used empirical reference to adjust the top earnings is the Pareto type I model. As will be shown, arbitrary choice of cutoff points for income, such as the highest 5% or 10%, can lead to wrong inferences. For this reason, we address the optimal cutoff problem.

5.2 RELATED LITERATURE

In recent decades, interest in research on the distribution of higher incomes has grown. The choice of the upper income stratum as the object of study is not limited to the purpose of explaining only the distributive dynamics among members of the wealthier social strata.

In fact, the different parts of the distribution are interdependent, insofar as the outcome of one group is affected by the outcome of the others. In this sense, the choice of the group with the highest income is based on some specific aspects of this set, such as its command over resources, its command over people and its global significance (ATKINSON; PIKETTY, 2007).

However, the different interoccupational patterns of distribution of the highest incomes have not yet been explored in a specific study. In this context, the attribute of scalability of professions proposed by Taleb (2010) constitutes an alternative approach to address the issue of wage distribution. Thus, this essay will support the concept of scalability as a determining element to explain the highest incomes from work among some selected professions.

5.3 METHODOLOGY

There are some commonly used methods to estimate the shape parameter α of the Pareto model type I. In this essay we considered the OLS estimator proposed by Gabaix and Ibragimov (2011) – equation 5 – and the ML estimator proposed by Hill (1975) – equation 4.

An important question related to implementation is: if we assume that rents are described by a Pareto model above some threshold, what should that threshold be? What is the cutoff point to use to distinguish between higher incomes and not-so-high incomes? Is the highest income group the top 10% (RUIZ; WOLOSZKO, 2015), the top 5% (ATKINSON, 2016) or the top 1% (ALVAREDO, 2011)? This discussion will be empirically presented in the following section.

5.4 RESULTS

5.4.1 Estimation

To begin our analysis, let's consider the traditional 10% and 5% cutoff points. It is important to note that smaller values of the Pareto exponent mean heavier tails.

Table 2. Pareto coefficients for pre-established thresholds: OLS estimation

Occupation	Top 10 percent incomes	Top 5 percent incomes
Motorcycle messenger	2.94	2.39
Bank clerk	4.17	4.49
Dentist	3.66	3.47
Lawyer	2.85	3.13
Elementary school Portuguese teacher	2.77	3.28
Architect	4.50	4.52
Fashion designer	3.01	3.00
Radio and TV broadcaster	2.11	2.39
Visual artist	4.50	6.19
Music performer	6.64	8.44
Actor	4.19	8.10
Footballer	1.64	4.14

Source: Own elaboration based on RAIS 2017 microdata

The results in Table 2, which reveal the Pareto coefficients estimated using the method proposed by Gabaix and Ibragimov (2011), suggest a diffuse pattern among the selected professions.

Of the entire sample, the heaviest tail was the soccer player for the highest-paid 10%. Although this individual result is in line with expectations, we can see that if we consider the top 5% of earnings, the coefficient increases a lot, showing itself even higher than that of some notably non-scalable occupations.

The high values of the α coefficient found for actor, musician and visual artist indicate that the income distribution between the top 10% and 5% is less unequal than the upper strata of occupations such as motorcycle messenger, lawyer and Portuguese teacher.

As a comparison, we present in table 3 the estimates of the Pareto α coefficient obtained through the maximum likelihood method.

Table 3. Pareto coefficients for pre-established thresholds: ML estimation

Occupation	Top 10 percent incomes	Top 5 percent incomes
Motorcycle messenger	4.94	4.33
Bank clerk	3.62	4.55
Dentist	3.59	3.63
Lawyer	2.49	2.81
Elementary school Portuguese teacher	2.62	2.56
Architect	4.25	4.60
Fashion designer	2.89	3.12
Radio and TV broadcaster	1.92	2.12
Visual artist	3.47	4.87
Music performer	7.64	6.72
Actor	2.69	6.79
Footballer	1.03	2.70

Source: Own elaboration based on RAIS 2017 microdata

The results in table 3 confirm the diffuse pattern found for the pre-established thresholds of 5% and 10%. Although heavier tails were observed for soccer player and radio and television broadcaster, low values for the Pareto exponent were also found for remarkably non-scalable occupations, such as lawyer and Portuguese teacher.

In addition, it appears that the Pareto α coefficient estimates for non-scalable professions were less sensitive to the values used as threshold, namely, the 90th and 95th percentiles.

Setting the cutoff point for the highest incomes in the Pareto Type I model as the highest 5% or 10% can lead to wrong inferences. A higher cutoff point lowers the Pareto exponent estimate, artificially leading to more inequality among top incomes (JENKINS, 2017). For this reason, we address the problem of the optimal cut-off point here.

5.4.2 Optimal cut-off point: the Kolmogorov-Smirnov statistics

There is some evidence that adopting a higher cutoff point decreases the estimate of the scale coefficient α, which characterizes the Pareto type I distribution. Thus, assigning a higher threshold would increase inequality between higher incomes, keeping others constant variables (BURKHAUSER, 2012).

Some criteria have been proposed for the computation of the ideal cutoff point of the Pareto distribution (see, for example, CLAUSET; SHALIZI; NEWMAN, 2009).

A common and pragmatic method of choosing the cut-off point is to visualize the point beyond which the accumulated distribution function becomes more or less linear in a graph constructed on a logarithmic scale. Another visual inspection tool for this purpose is to plot the estimated α exponent as a function of the threshold and identify a point beyond which the value appears relatively stable. These approaches, however, can be considered subjective and present sensitivity to noise or fluctuations in the tail of the distribution (CLAUSET; SHALIZI; NEWMAN, 2009).

Clauset, Shalizi and Newman (2009) advocate a "more objective and principle-based approach, based on minimizing the "distance" between the power-law model and the empirical data". In this sense, they suggest measuring the distance between fitted and empirical distributions using the Kolmogorov-Smirnov (KS) statistic.

Figure 5 shows the Zipf plots derived from the empirical estimates of ln *S(x)*, where *S(x)* ≤ 0.5. From a first visual inspection of the tails, we cannot rule out the presence of power laws after some threshold x_{min}. It is important to note that few distributions follow a power law across the entire range of values. Normally, the power law appears after a lower limit and usually disappears after an upper limit; this is precisely the reason why a distribution is characterized by a power-law tail (NEWMAN, 2005).

Figure 5. Zipf plots for $S(x) \leq 0.5$

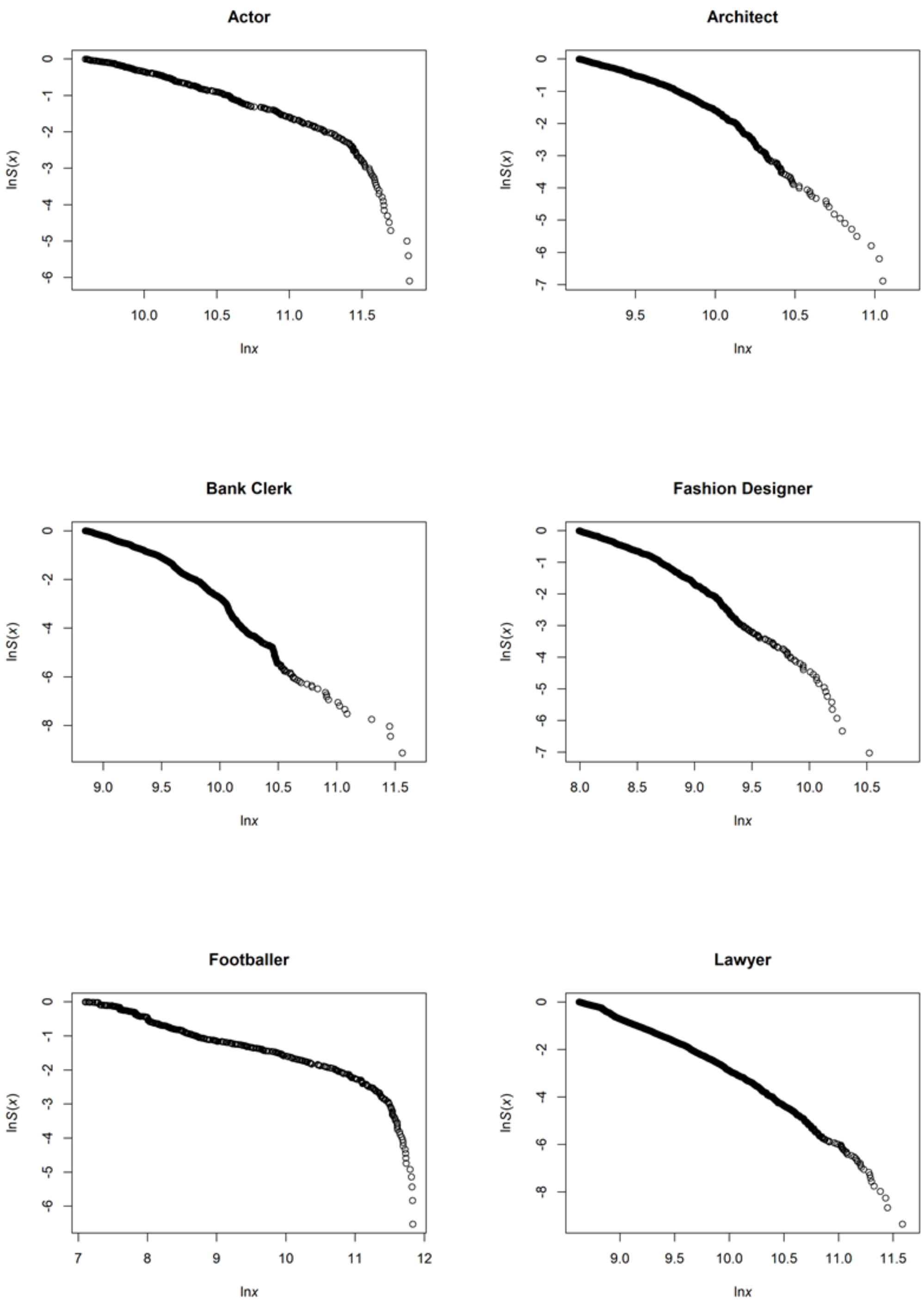
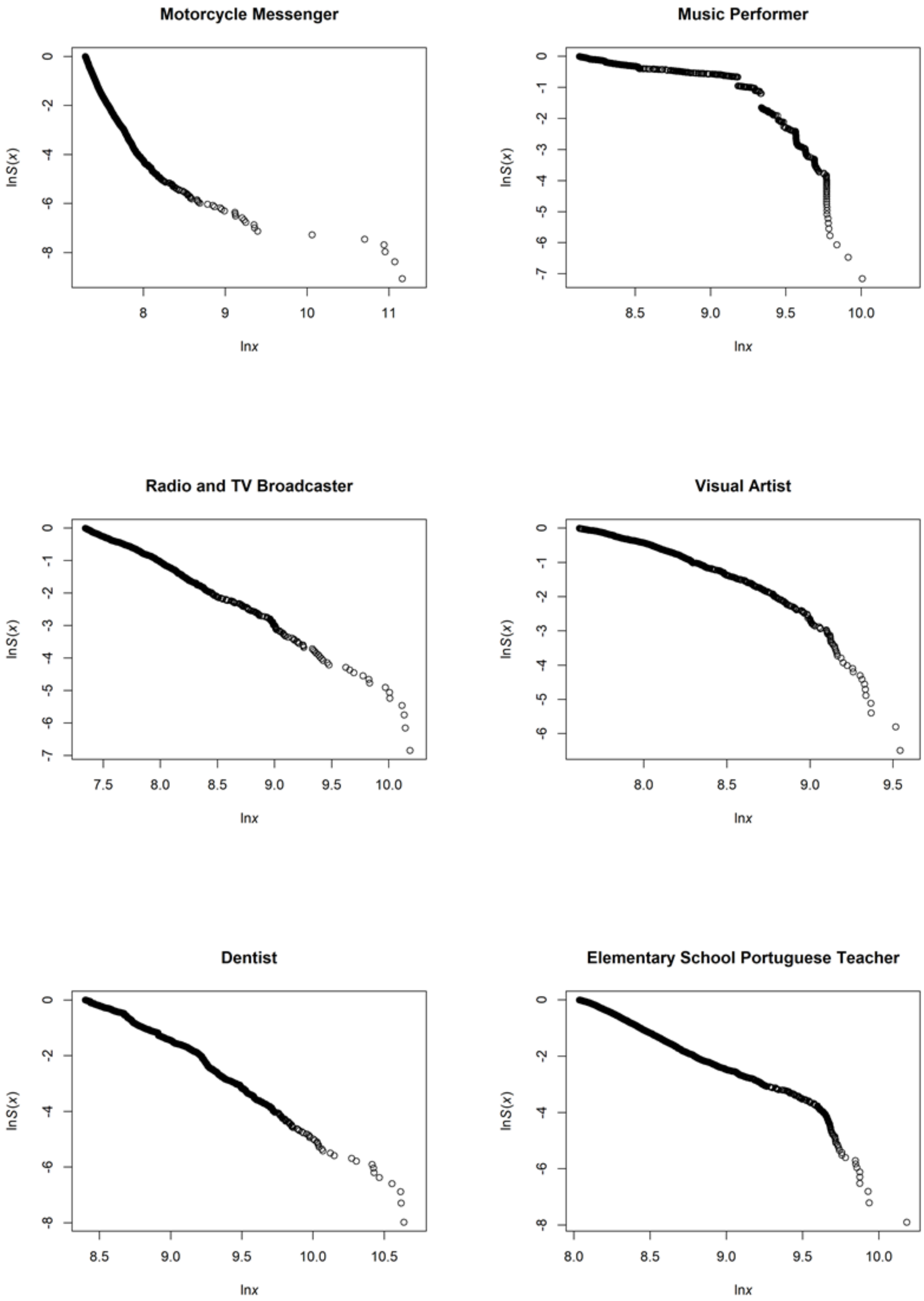


Figure 5. Zipf plots for $S(x) \leq 0.5$ - Continue



As the significance of the KS statistic depends on the sample size, rather than minimizing D in Equation (9) to find the optimal one, we determined the ideal cutoff point in percentage terms associated with the smallest KS statistic with a p value > 0.05 (Fig. 6).

Figure 6. D statistics and their p-values vs. thresholds: OLS(left) and ML(right)

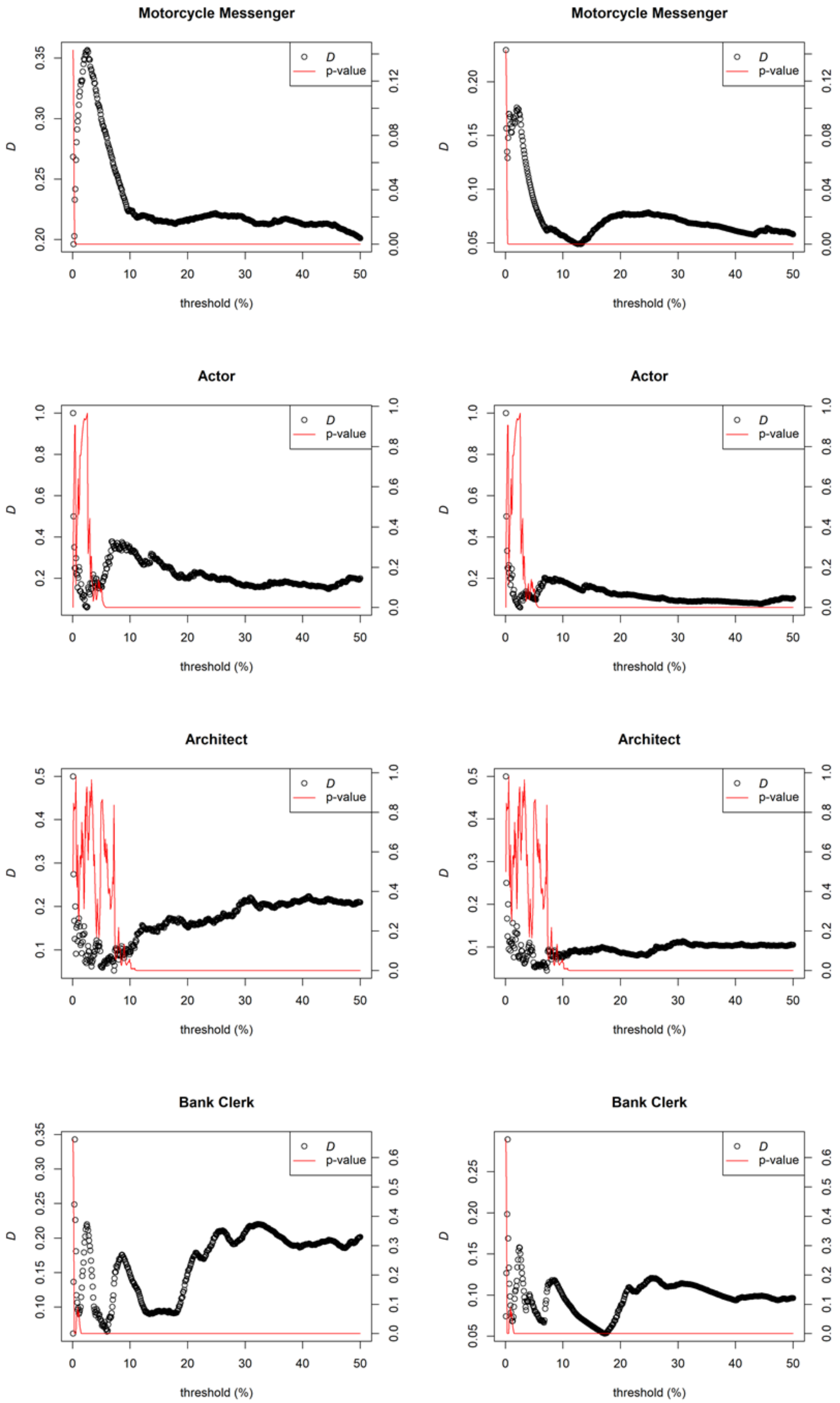


Figure 6. D statistics and their p-values vs. thresholds: OLS(left) and ML(right) - Continue

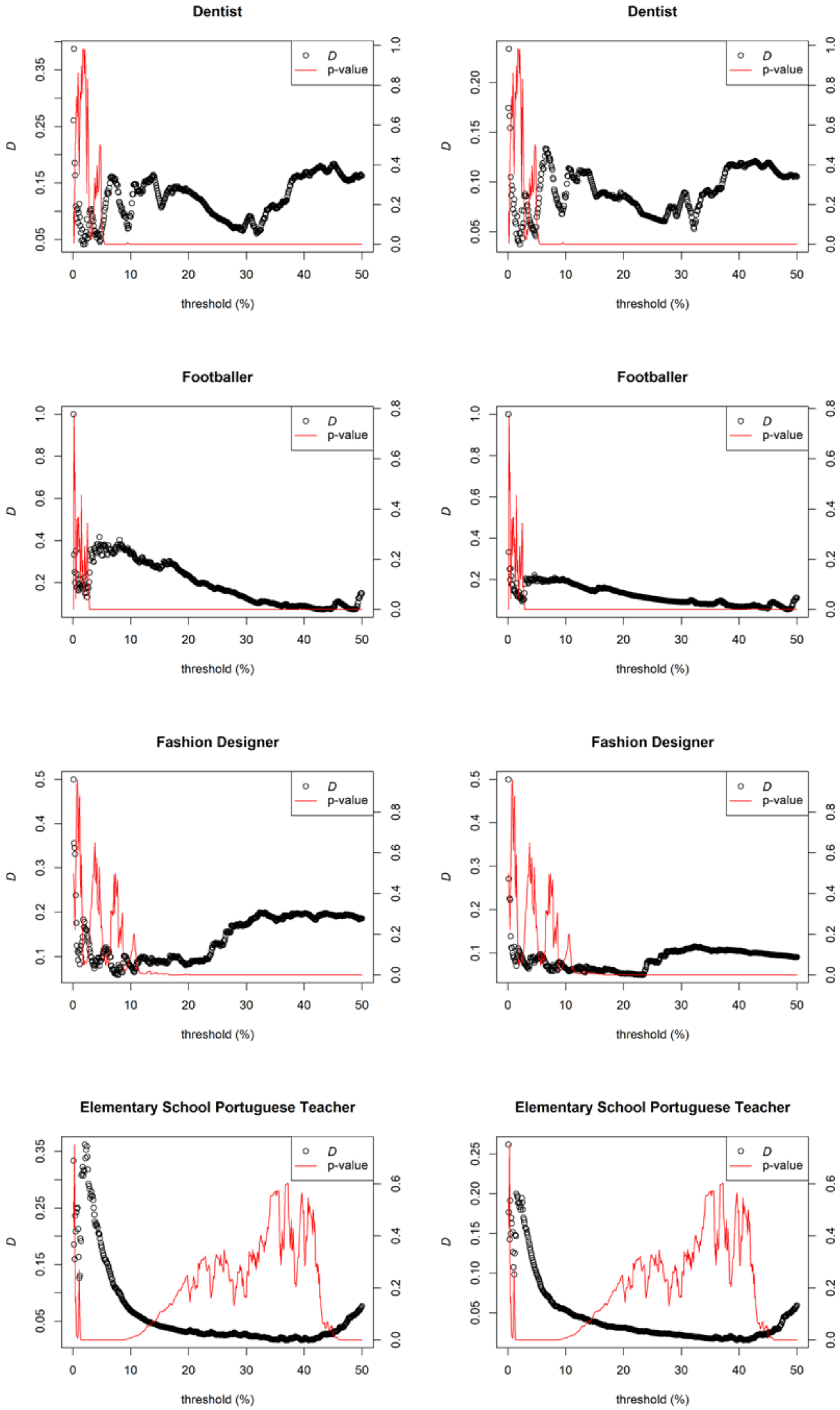
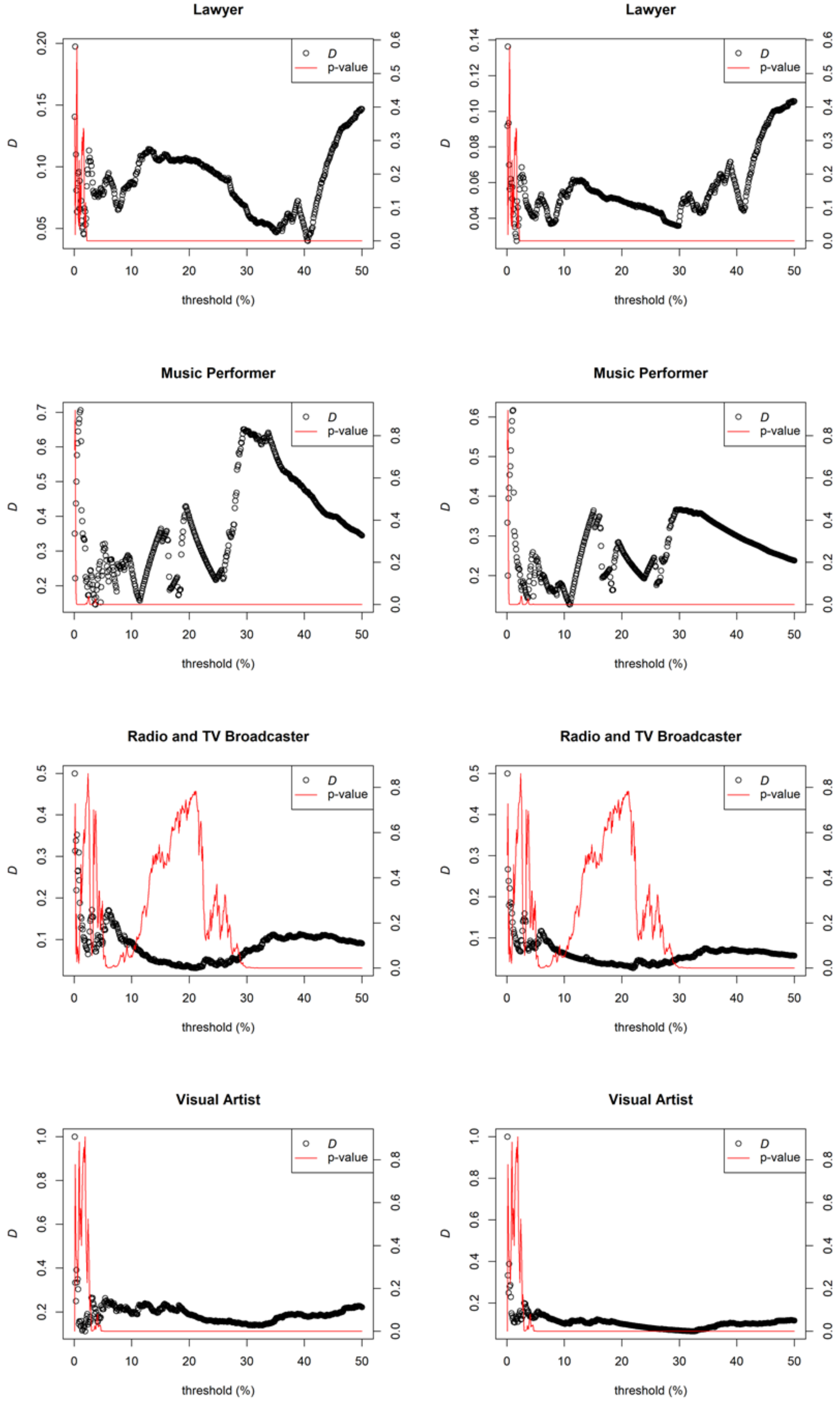


Figure 6. D statistics and their p-values vs. thresholds: OLS(left) and ML(right) - Continue



Tables 4 and 5 provide a summary of the estimates for OLS and ML, respectively. It is important to highlight that the OLS estimates employ regressions against the log of earnings, as proposed by Gabaix and Ibragimov (2011).

Table 4. OLS estimates for the scaling regime above the KS thresholds

Occupation	τ	*n*	α	*G*	*E/ln n*	*D*	*p-value*
Motorcycle Messenger	0.22	38	1.06	0.89	0.5304	0.19	0.128
Radio and TV Broadcaster	21.00	397	2.06	0.32	0.0300	0.03	0.766
Elementary School Portuguese Teacher	39.60	2155	2.73	0.22	0.0118	0.02	0.566
Fashion Designer	7.80	176	3.05	0.20	0.0134	0.06	0.469
Bank Clerk	0.10	18	3.19	0.19	0.0217	0.06	0.663
Dentist	1.80	106	3.50	0.17	0.0109	0.04	0.984
Lawyer	1.70	396	3.58	0.16	0.0081	0.05	0.325
Music Performer	0.20	5	4.27	0.13	0.0203	0.22	0.921
Architect	7.20	142	4.58	0.12	0.0057	0.05	0.837
Visual Artist	1.90	25	7.05	0.08	0.0035	0.11	0.908
Footballer	2.50	34	8.34	0.06	0.0022	0.13	0.343
Actor	2.50	22	9.87	0.05	0.0018	0.06	0.955

Source: Own elaboration based on RAIS 2017 microdata

Table 5. ML estimates for the scaling regime above the KS thresholds

Occupation	τ	*n*	α	*G*	*E/ln n*	*D*	*p-value*
Motorcycle Messenger	0.43	75	1.67	0.43	0.0731	0.12	0.194
Radio and TV Broadcaster	21.90	414	2.00	0.33	0.0319	0.03	0.813
Fashion Designer	23.40	529	2.43	0.26	0.0189	0.05	0.126
Elementary School Portuguese Teacher	40.70	2215	2.71	0.23	0.0118	0.02	0.620
Bank Clerk	0.12	22	2.80	0.22	0.0275	0.06	0.688
Lawyer	1.70	396	3.40	0.17	0.0091	0.03	0.876
Visual Artist	10.10	134	3.46	0.17	0.0107	0.10	0.090
Dentist	2.10	123	3.59	0.16	0.0100	0.04	0.996
Architect	7.20	142	4.68	0.12	0.0054	0.04	0.896
Music Performer	0.20	5	5.60	0.10	0.0113	0.20	0.962
Footballer	2.50	34	7.51	0.07	0.0028	0.09	0.676
Actor	2.60	23	9.99	0.05	0.0017	0.06	0.958

Source: Own elaboration based on RAIS 2017 microdata

Tables 4 and 5 also display Gini and entropy indices. The Gini index (G) ranges from 0 (perfect equality) to 1 (perfect inequality). In the Pareto type I model, it is given by:

$$G = \frac{1}{2\alpha - 1}, \qquad (12)$$

to $\alpha > 1$ (JENKINS, 2017). We report the Gini index, but caution that it may be unreliable under heavy tails; at least this is true for non-parametric methods (FONTANARI; TALEB, 2018).

The generalized entropy index (E) in the Pareto type I model is given by:

$$E = \log\frac{\alpha}{\alpha - 1} - \frac{1}{\alpha}, \qquad (13)$$

for $\alpha > 1$ (Jenkins 2017), and can also be interpreted as a measure of inequality. To meaningfully compare occupations, we divide E by ln n to obtain a "normalized" entropy index in the range of 0 to 1.

Fig. 7 shows the presence of two scaling regimes in Zipf plots. Tables 4 and 5 present the results for the portion of data above the optimal cut-off points in percentage terms τ (vertical lines in Fig. 7). Here, motorcycle messengers have income distributions with heavier tails than those of actors and soccer players. It is important to highlight once again that smaller Pareto exponent values mean heavier tails. However, the reverse is true for the part of the data between the medians and ideal cutoff points τ (Table 6).

Table 6. Results for the scaling regime from medians to optimal cut-offs

Occupation	*OLS*			*ML*		
	α	G	E/ln n	α	G	E/ln n
Motorcycle Messenger	6.67	0.08	0.0013	7.50	0.07	0.0011
Broadcaster	5.31	0.10	0.0032	4.14	0.14	0.0055
Fashion Designer	2.97	0.20	0.0107	2.15	0.30	0.0235
Portuguese Teacher	23.92	0.02	0.0001	15.19	0.03	0.0004
Bank Clerk	2.67	0.23	0.1036	1.99	0.34	0.0215
Lawyer	2.64	0.23	0.0104	2.12	0.31	0.0177
Visual Artist	2.37	0.27	0.0195	1.72	0.41	0.0450
Dentist	3.46	0.17	0.0065	2.74	0.22	0.0112
Architect	3.36	0.17	0.008	2.41	0.26	0.0180
Music Performer	1.66	0.43	0.0449	1.23	0.69	0.1223
Footballer	0.81	--	--	0.67	--	--
Actor	1.74	0.40	0.0462	1.29	0.64	0.1198

Source: Own elaboration based on RAIS 2017 microdata
Note: For $\alpha < 1$, G and E cannot be computed

This differentiated result arises because the ideal cutoff points τ have high variance - they range from 0.1 to 39.6% in Tables 4 and 5. Therefore, there are very different concentration dynamics in the extreme gains between occupations. In this case, the Gini and entropy indices can add relevant explanations about this dynamic. For the power law regime above (Tables 4 and 5), the extreme income distributions of actors and soccer players are more equal than those of motorcycle messengers, but they are more unequal for the power law regime from median to τ (Table 6).

In short, the ranking of occupations from non-scalable to scalable depends not only on the chosen higher income threshold, but also on the scaling regime. From the median to the ideal cutoff point, soccer players are a more scalable occupation than motorcycle messengers. However, above τ, this is no longer true.

Figure 7. Two scaling regimes in Zipf plots: vertical lines are optimal cutoff points in percentage terms

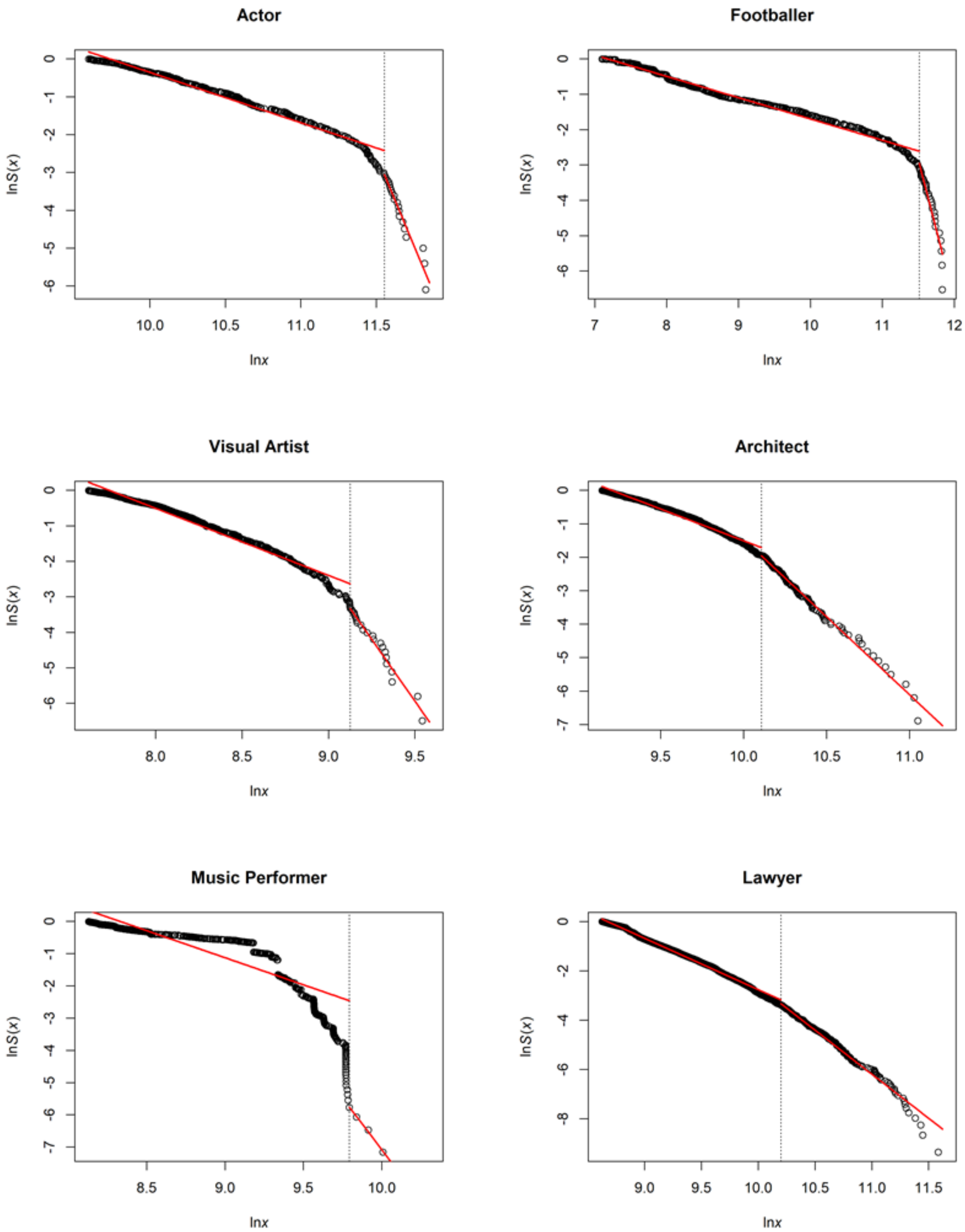
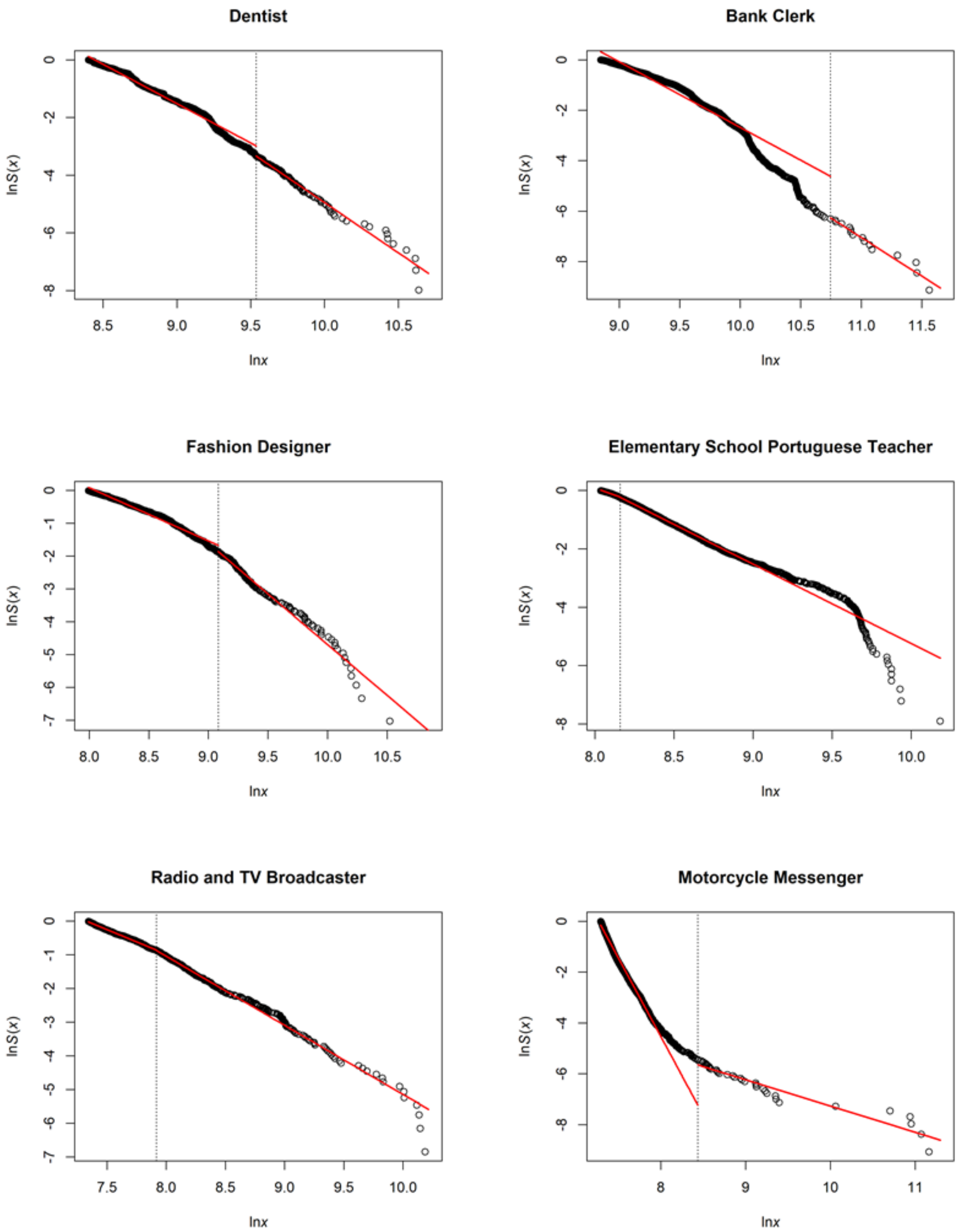


Figure 7. Two scaling regimes in Zipf plots: vertical lines are optimal cutoff points in percentage terms - Continue



5.4.3 The R^2 coefficient

In this section, we develop the analysis based on the values of the R^2 coefficient. Table 7 provides a summary of the OLS and ML estimates for the part of the data above the ideal cut-off points. Here, instead of $\log\left(\text{rank}-\frac{1}{2}\right)$ we will use the logarithm of the complementary cumulative distribution function S(x) as the explained variable in the OLS regression.

Table 7. Pareto coefficients above optimal cutoffs R^2

Occupation	τ	n	R^2	OLS			ML		
				α	G	E/ln n	α	G	E/ln n
Motorcycle Messenger	0.35	61	0.93	1.07	0.87	0.4285	1.51	0.49	0.1026
Broadcaster	3.23	61	0.95	2.28	0.28	0.0338	2.65	0.23	0.0235
Fashion Designer	2.21	50	0.97	3	0.2	0.0185	2.74	0.22	0.0228
Dentist	0.85	50	0.96	3.16	0.19	0.0163	3.73	0.15	0.0112
Bank Clerk	0.66	121	0.95	3.45	0.17	0.0109	4.13	0.14	0.0073
Lawyer	0.21	50	0.96	4.02	0.14	0.0095	3.63	0.16	0.0119
Architect	2.53	50	0.99	4.06	0.14	0.0093	4.43	0.13	0.0077
Visual Artist	9.13	121	0.95	4.46	0.13	0.0062	3.51	0.17	0.0105
Actor	5.95	53	0.95	6.75	0.08	0.0031	5.27	0.1	0.0052
Musician	8.8	227	0.91	6.95	0.08	0.0021	5.26	0.11	0.0038
Portuguese Teacher	1.4	76	0.95	7.3	0.07	0.0024	6.77	0.08	0.0028
Footballer	2.77	38	0.91	7.32	0.07	0.0028	6.09	0.09	0.0042

Source: Own elaboration based on RAIS 2017 microdata

The values of R^2 show a behavior of the upper tail estimated via OLS similar to the asymptotic behavior of the ML estimates (DOREA et al., 2016). The ranking of Pareto exponents based on estimates via OLS is close to that found for estimates via ML. The ranking of professions, however, is far from the expected pattern.

Motorcycle messengers, for example, have an income distribution with a heavier upper tail than soccer players (smaller values of the Pareto exponent mean heavier tails.). The median earnings of motorcycle messengers and soccer players are similar – R$ 1,468.72 and R$ 1,213.71, respectively. But their corresponding extreme superior earnings cannot be adequately captured above the ideal cut-off points.

The Gini and entropy indices also reveal an unexpected pattern. The extreme income distributions of football players are more equal than those of motorcycle messengers. The ideal cut-off points exhibit high variance - they range from 0.21% to 9.13% in Table 7 - which reveals very different dynamics of extreme earnings concentration between professions. Fig. 8 shows the R^2 values and the corresponding tail indices of the OLS and ML estimates.

Figure 8. R^2 statistics (left) and the corresponding tail index (right) vs. cut-offs

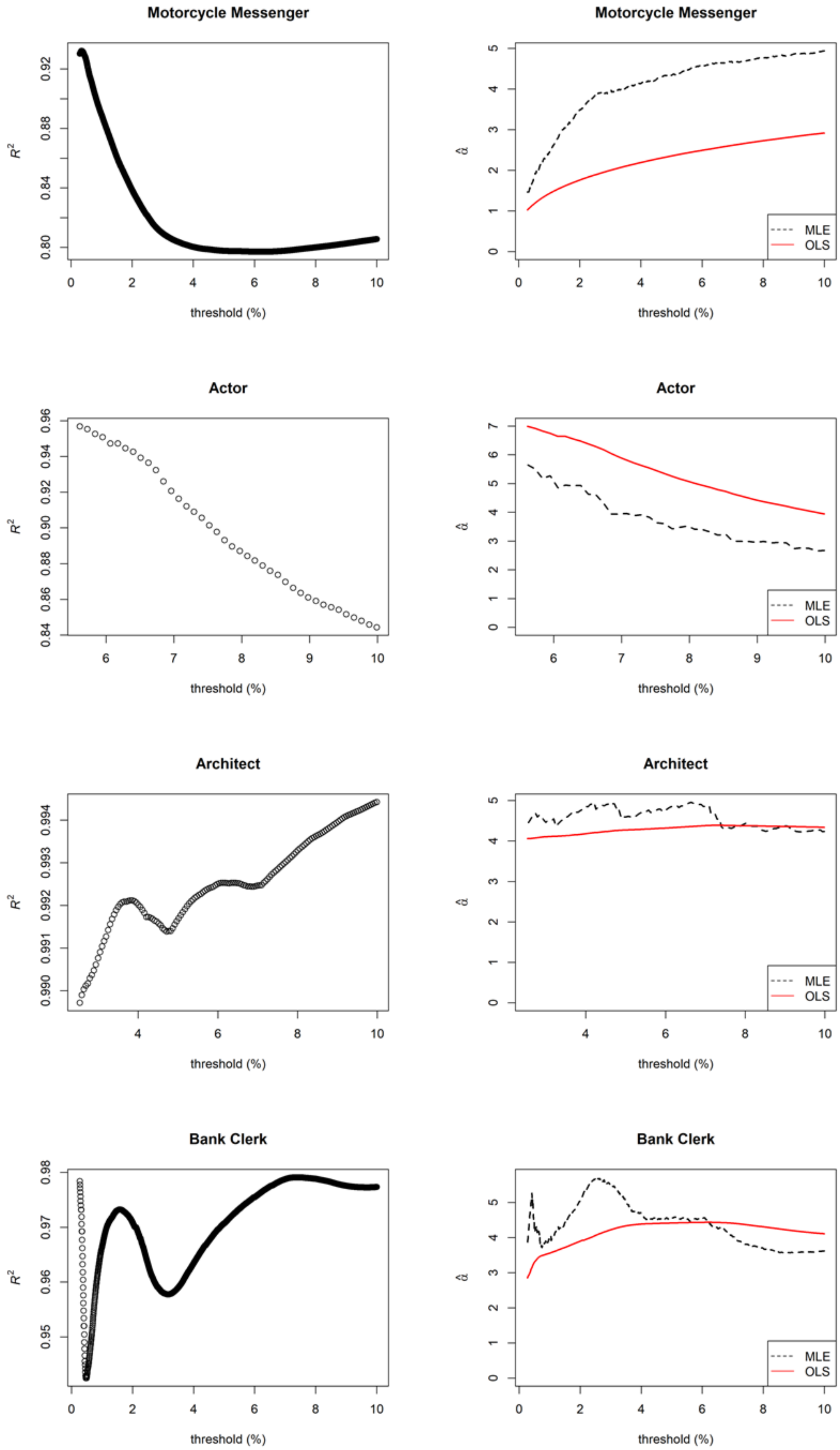


Figure 8. R^2 statistics (left) and the corresponding tail index (right) vs. cut-offs - Continue

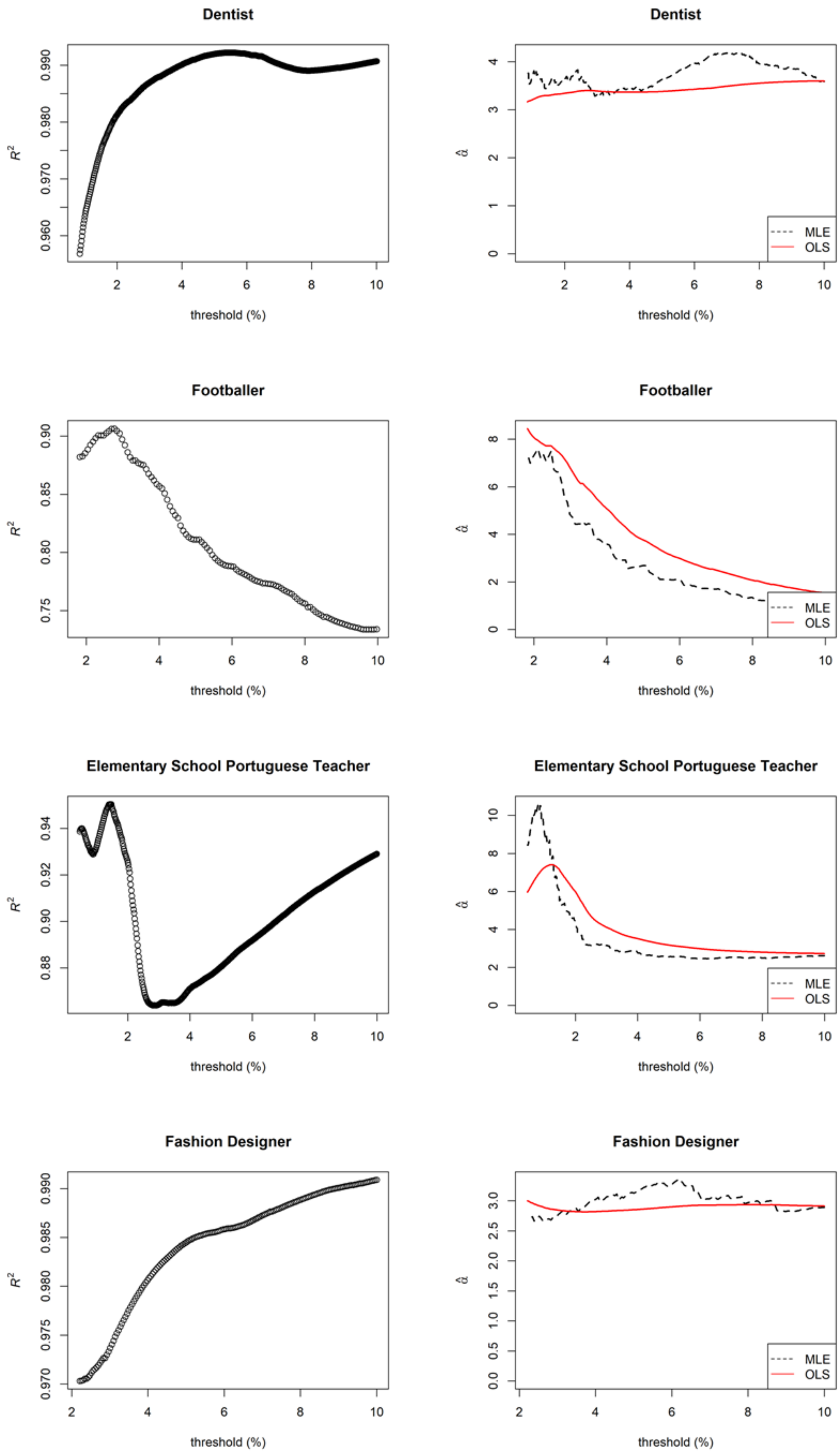


Figure 8. R^2 statistics (left) and the corresponding tail index (right) vs. cut-offs - Continue

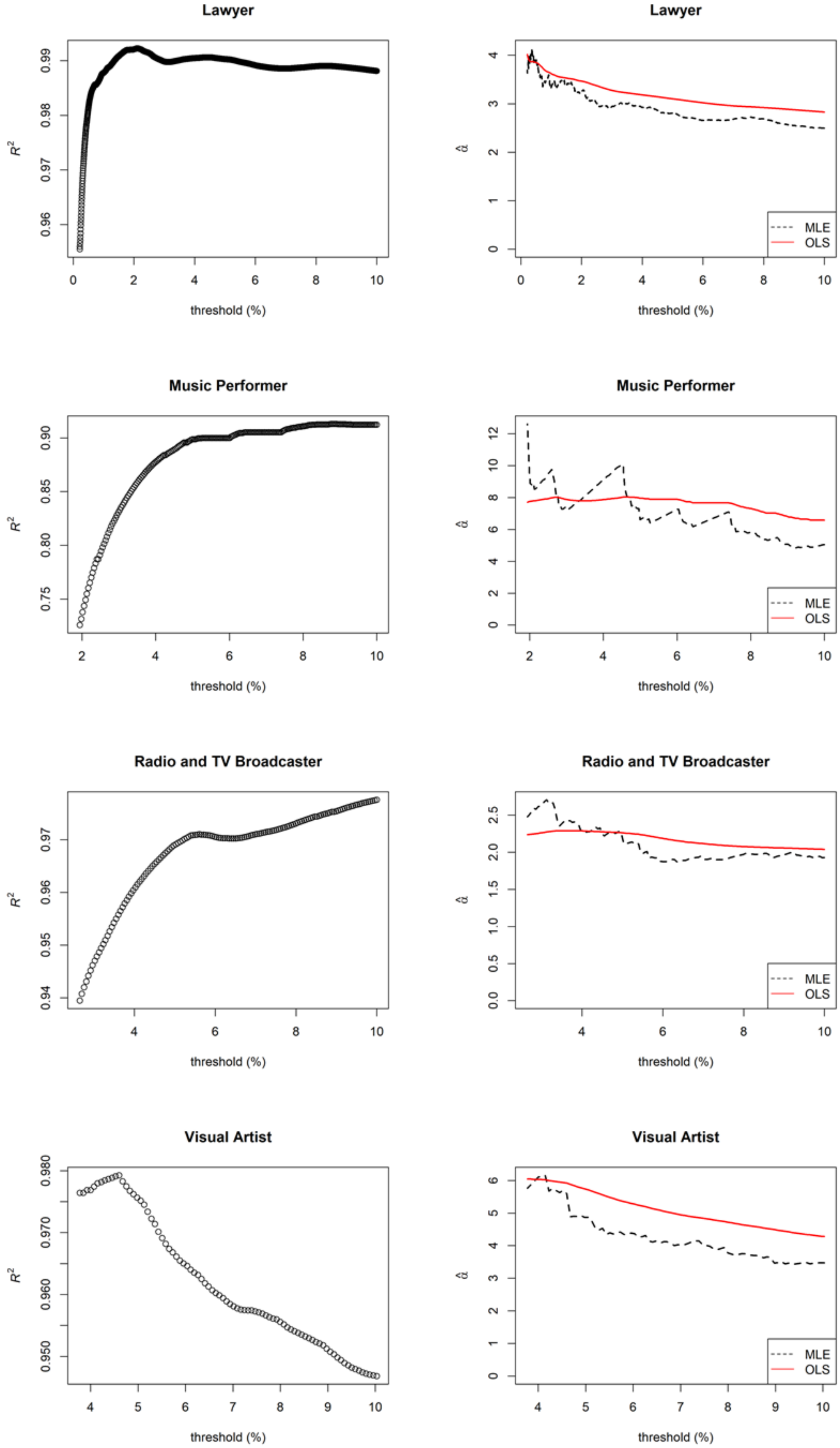


Figure 9 shows the scaling on the Zipf plot for the part of the data above the ideal cut-off points.

Figure 9. Scaling in Zipf plots: vertical lines are optimal cutoff points in percentage terms

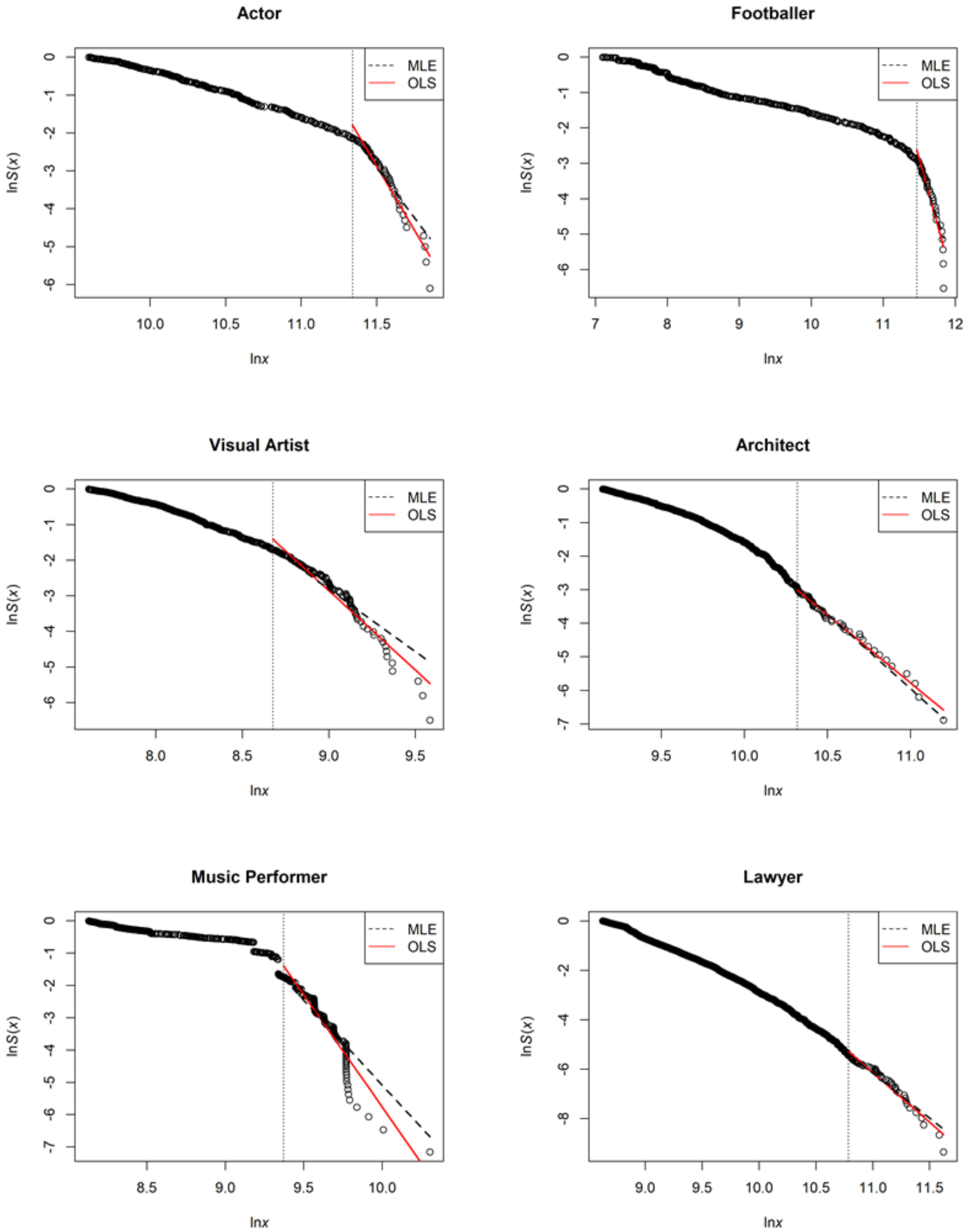
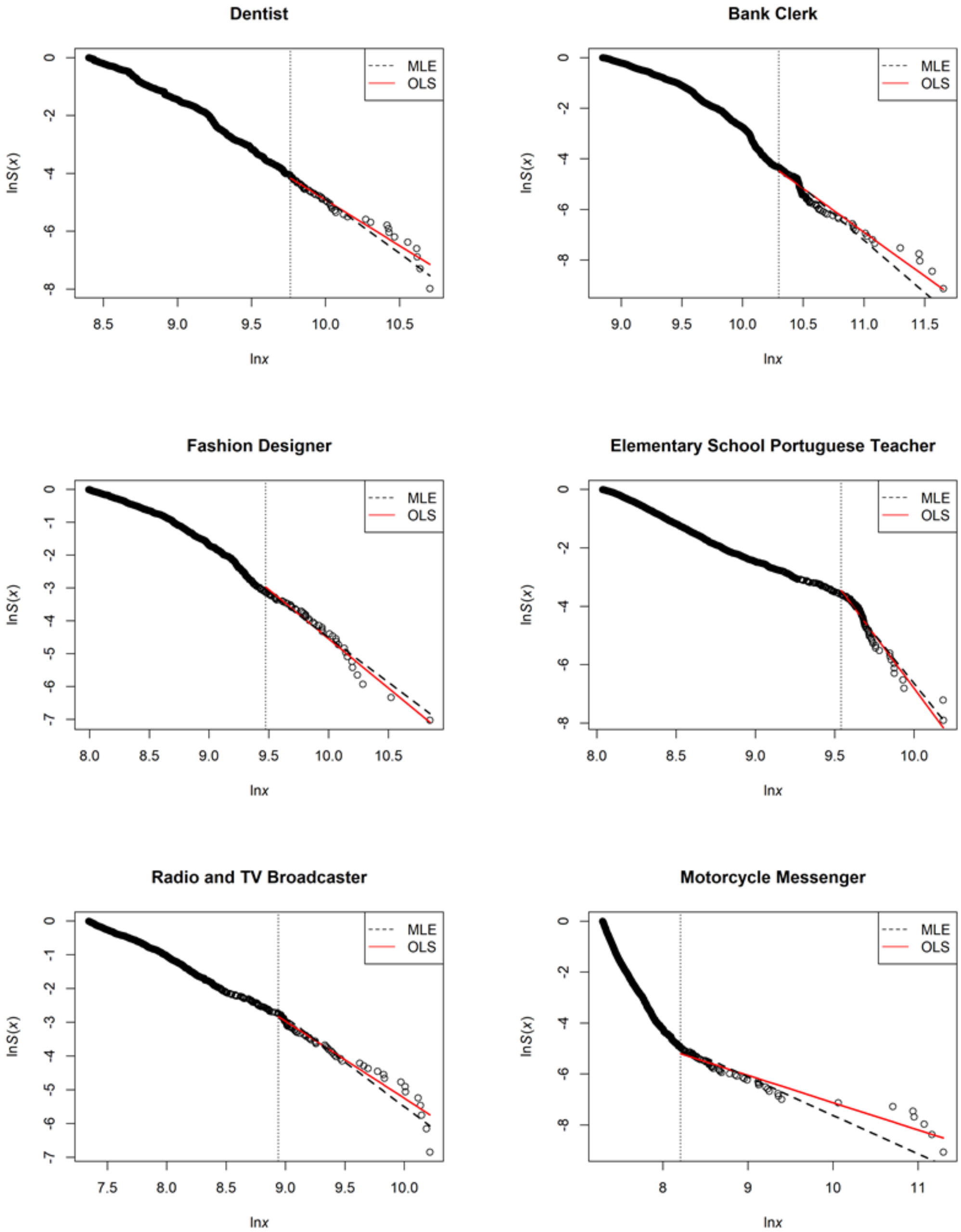


Figure 9. Scaling in Zipf plots: vertical lines are optimal cutoff points in percentage terms - Continue



There are drastic changes in the scaling regime (Fig. 9). For this reason, we also calculate OLS estimates for the portion of data above the medians. Table 8 shows that the expected ranking can be roughly replicated.

Table 8. OLS estimates for the scaling above medians

Occupation	τ	n	α	G	E/ln n	R^2
Footballer	18.85	231	0.70	2.46	-	0.9800
Musician	31.98	330	0.83	1.50	-	0.9800
Actor	34.79	125	0.87	1.33	-	0.9799
Architect	48.40	30	0.97	1.06	-	0.9799
Visual Artist	36.43	176	1.13	0.80	0.1979	0.9800
Portuguese Teacher	49.32	36	1.41	0.55	0.0670	0.9797
Dentist	47.66	92	1.45	0.53	0.0606	0.9800
Broadcaster	49.05	18	1.53	0.49	0.0595	0.9786
Fashion Designer	7.74	932	1.62	0.45	0.0490	0.9800
Bank Clerk	8.59	7454	1.93	0.35	0.0234	0.9800
Lawyer	19.54	5384	2.02	0.33	0.0201	0.9800
Motorcycle Messenger	49.75	35	6.26	0.09	0.0016	0.9795

Source: Own elaboration based on RAIS 2017 microdata
Note: For $\alpha < 1$, G and E cannot be computed

Now, the distribution of income of soccer players has heavier tails (a smaller exponent of Pareto) than that of motorcycle messengers. Therefore, for earnings above the medians, we cannot reject the hypothesis that income distributions from scalable occupations exhibit heavier tails.

5.5 CONCLUSIONS

We calculated the Pareto exponents of the income distributions of twelve selected professions using OLS and ML. So, we test the hypothesis that these non-scalable occupations are more egalitarian than these scalable ones. We conclude that we cannot reject such a hypothesis for the portion of income data between the median and the optimal cut-off point.

It was found, therefore, that the income distributions of these scalable occupations decay as a power law and have heavier tails - smaller Pareto exponents - than these non-scalable occupations. However, there is another power law regime above the ideal cutoff points, where these results are reversed.

This differentiated general result arises because the ideal cutoff points for extreme earnings have high variance and, therefore, a very different concentration dynamics between occupations. Thus, the ranking of occupations from non-scalable to scalable depends not only on the chosen higher income threshold, but also on the scaling regime in the tails.

6 ESSAY 3: THE "LOCAL" APPROACH OF THE INVERTED COEFFICIENT AND THE GENERALIZED PARETO CURVES

6.1 INTRODUCTION

A more comprehensive analysis of income inequality must address distributions in their entirety, from the upper tail of incomes, where power laws are a good description, to the lower portions, where they are not. For this purpose, a good analysis tool is a non-parametric approach called generalized Pareto curve, which is the graphical representation of the inverted Pareto coefficients $b(p)$. So, instead of trying to make a preconceived functional shape fit the data by tuning a set of parameters, we start with the observed Pareto curve. In this way, we intend to compare the behavior of the inverted Pareto coefficient over the entire range of income distributions between selected scalable and non-scalable occupations.

6.2 RELATED LITERATURE

The specialized economic literature has been finding out for some time that the upper portion of the income distribution presents a good fit with the Pareto model, which means that the arrangement of the highest incomes follows a power law.

However, the empirical analysis of the data reveals that, although the approximation with the Pareto distribution is acceptable for some purposes, it does not hold even on top of the observations. In practice, assuming that a certain upper portion of the data is distributed according to the Pareto model imposes the strict restriction that inequality is configured in the same way in all higher-income groups within this interval.

This is equivalent to saying that the complete distribution between the 10%, 1% or 0.1% of higher incomes is the same, which is not necessarily the case. If this property manifests itself, we enter into the discussion of the "fractal" nature of inequality. However, this result is not so common in practice.

In this sense, the real distributions of income and wealth usually have Pareto coefficients that depend on the rank $p \in [0,1]$ in the distribution. By allowing the variation of these coefficients, we allow more flexibility and precision, keeping the Pareto model as a baseline (BLANCHET; FOURNIER; PIKETTY, 2017).

Thus, the "local" approach of the inverted Pareto coefficient, calculated for the entire sample from the observation rank, reveals more faithfully the nature of inequality in a given income distribution.

Formally, the inverted Pareto coefficient $b(p)$ is defined as the ratio between the average income above rank p and the p-th quantile, that is:

$$b(p) = \frac{\mathbb{E}[X|X > Q(p)]}{Q(p)} \qquad (14)$$

Ultimately, $b(p)$ = 1 defines a situation in which all individuals above rank p have the same income or wealth, so that there is no inequality above p (BLANCHET et al. 2018). Therefore, the larger the coefficient $b(p)$, the higher the corresponding level of inequality.

In practice, inverted Pareto coefficients range from 1.5 to 3.5 (PIKETTY, 2014). An inverted coefficient of 1.5 means that the average income beyond a certain threshold is equal to one and a half times the value of that threshold (people who have more than one million euros have, on average, 1.5 million euros, and so on for any threshold), which corresponds to a relatively smooth inequality. An inverted coefficient of 3.5, on the contrary, corresponds to a very strong inequality.

6.3 METHODOLOGY

For any income level x > 0, the inverted Pareto coefficient is $b^*(x) = \mathbb{E}[X|X > x]$ or:

$$b^*(x) = \frac{1}{(1 - F(x))x} \int_x^{+\infty} zf(z)dz \qquad (15)$$

where F(x) is the cumulative distribution function of x.

The inverted Pareto coefficient can also be expressed as a function of the fractile p with p = F(x) and b(p) = $b^*(x)$:

$$b(p) = \frac{1}{(1 - p)Q(p)} \int_p^1 Q(u)du \qquad (16)$$

If X follows a Pareto distribution with coefficient α and lower bound $\bar{x}$, so that $F(x) = 1 - \left(\frac{\bar{x}}{x}\right)^{\alpha}$, then $b(p) = \frac{\alpha}{\alpha - 1}$ is constant and the participation of the top $100 \times (1 - p)\%$ is an increasing function of b and is equal to $(1 - p)^{\frac{1}{b}}$. Otherwise, b(p) will vary.

Thus, the function $b: p \mapsto b(p)$ defined over $[\bar{p}, 1[$ with $\bar{p} = F(\bar{x})$ is called the generalized Pareto curve (BLANCHET; FOURNIER; PIKETTY, 2017).

For a strict power law (a Pareto distribution), the Pareto curve is constant. But strict power laws rarely exist in practice, so we can characterize the Pareto curve when the power-law behavior is only approximated.

In practice, Blanchet, Fournier, and Piketty (2017) found that b(p) varies within the upper tail of observed income and wealth distributions (including between the top 10% or the top 1%), but the b(p) curves are relatively similar (usually U-shaped).

Probability distributions can be divided into three categories, based on the behavior of their generalized Pareto curve (BLANCHET; FOURNIER; PIKETTY, 2017).

First, power laws for which b(p) converges to a constant strictly greater than one. Second, "thin-tailed" distributions for which b(p) converges to one. The third category includes distributions with erratic tail behavior for which b(p) can oscillate at an ever faster rate without converging to anything. This last category does not include any standard parametric family of distributions and its members can essentially be considered pathological. If we exclude it, we are left with a straightforward dichotomy between power laws and thin tails.

When $\lim_{p \to 1} b(p) > 1$, such that X is an asymptotic power law, the generalized Pareto curve can still be used to observe how the distribution converges. If b(p) increases close to p = 1, the tail will get fatter with higher levels of income. But if b(p) decreases, it's getting thinner.

With a strict power law such that b(p) is constant, the level of inequality remains the same as we move forward in the distribution. The share of the richest 10% among the entire population is the same as the share of the richest 1% among the richest 10% or the share of the richest 0.1% among the richest 1%. This property is often called the "fractal" nature of inequality.

6.4 RESULTS

The results found reveal diversified patterns of inverted coefficients and, consequently, of the shape of the curves for the different types of occupation. In fact, this represents an alternative way of illustrating the different characteristics concerning the distribution of income for professions that are distinguished, above all, by the attribute of scalability.

With the exception of the soccer player occupation, all curves were decreasing in the initial portion of their respective distributions, that is, the inverted Pareto coefficient decreases as we advance through the first percentiles. This means that initially, as a rule, income inequality decreases with p.

For the first fractiles, the coefficient b(p) found is very high, which suggests that the income of the extremely poorest is very small in relation to the average income of people who are richer than them. However, as we move forward in distribution, the income of individuals begins to represent a larger fraction of the average income above them, thus decreasing the value of b(p).

The main differences found lie at the top of the distributions, that is, at the last percentiles. For the non-scalable occupations, especially architect, bank clerk and courier, the curve presents an inflection in the last levels of p.

Thus, the slopes of b(p) change around the most advanced percentiles and then increase to the top of the distribution, which shows that within the group of better paid individuals there is also an important intragroup inequality, since the extreme portion of the best pay constitutes a highly concentrated subgroup. Thus, the layout of the curves for these occupations approximates the U-shape found by Blanchet, Fournier and Piketty (2017) for aggregated data.

For the scalable professions, especially fashion designer, music performer, radio and television broadcaster and visual artist, there was a certain equality at the top of the distributions, as the coefficient b(p) remained more or less constant or even decreased in the last fractiles.

The results, therefore, do not reject the conclusions of essay 2 that above the ideal cutoff points to characterize the top of incomes, scalable occupations are more egalitarian than non-scalable ones.

Figure 10. Generalized Pareto curves

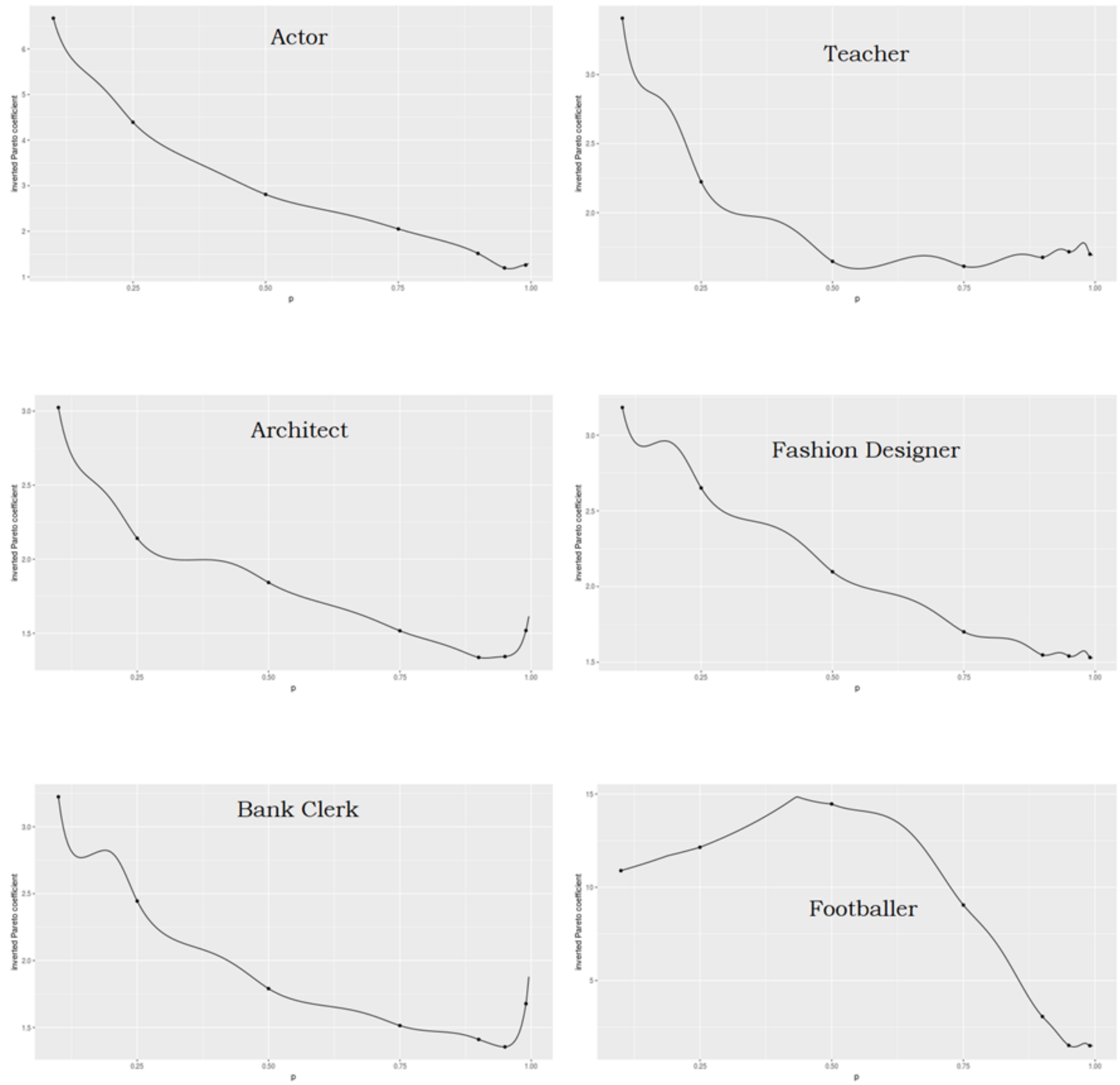
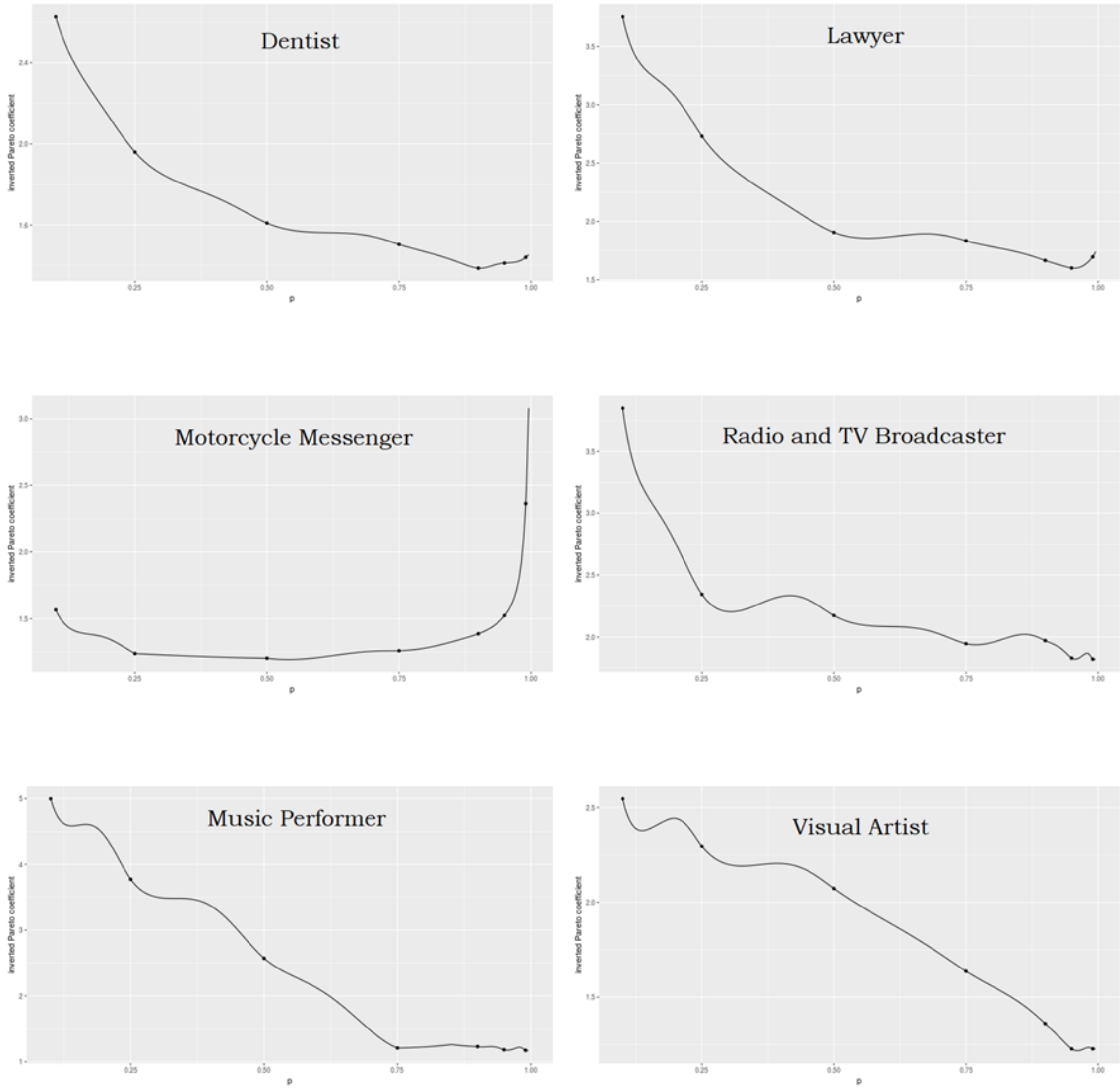


Figure 10. Generalized Pareto curves - Continue



6.5 CONCLUSIONS

For the purpose of addressing the income distributions in their entirety, from the upper tail, where power laws are a good description, to the lower portions, where they are not, we use generalized Pareto curves.

In general, income distributions from our scalable occupations decay as a power law and have heavier tails than our non-scalable occupations. Above the ideal cutoff points, however, there is another power law regime, where these results are reversed.

The results of essay 3, therefore, are in line with the conclusions of essay 2 that in the region that characterizes the top incomes, our selected scalable occupations are more egalitarian than the non-scalable ones.

In view of this, there is evidence of a very different concentration dynamics between the types of professions. Thus, the categorization of occupations from non-scalable to scalable depends not only on the highest income threshold chosen, but also on the scaling regime in the tails.

7 CONCLUSION

The present thesis aimed to verify if there is a statistically significant difference between the earnings of selected scalable and non-scalable occupations. For this purpose, the work was divided into three essays, each one containing a different theoretical model.

The results of essay 1 allow us to conclude that, considering the two-class income distribution model, we cannot reject the hypothesis that our non-scalable occupations are more egalitarian than those scalable for the upper portion of the data from the optimal cutoff point. The higher incomes of both types of occupations fit well as a power law and the earnings distribution of scalable occupations have heavier tails - smaller Pareto exponents - than non-scalable occupations.

In essay 2, we focused on the top earnings of the selected professions. Our aim was to test the hypothesis that the higher earnings of non-scalable occupations are more egalitarian than scalable ones. We conclude that we cannot reject such a hypothesis for the portion of income data between the median and the optimal cut-off point. However, there is another power law regime above the ideal cutoff points, where these results are reversed.

In the 3rd essay, we compared the behavior of the inverted Pareto coefficient over the entire range of earnings distributions between scalable and non-scalable occupations using generalized Pareto curves. The results of essay 3 are in line with the conclusions of essay 2 that in the region that characterizes the top earnings scalable occupations are more egalitarian than non-scalable ones.

This differentiated general result arises because the ideal cutoff points for extreme earnings have high variance and, therefore, a very different concentration dynamics between occupations. Thus, the ranking of occupations from non-scalable to scalable depends not only on the chosen higher income threshold, but also on the scaling regime in the tails.

In addition to revealing a distinct earnings concentration mechanism between the types of professions, the present work also sheds light on the macroeconomic dimension. Based on the intra and inter-occupational characterizations of earnings distribution, we provided new elements and perspectives to address income inequality at the aggregate level.

BIBLIOGRAPHIC REFERENCES

Alvaredo, F. A note on the relationship between top income shares and the Gini coefficient. Economic Letters, 110 (3), p. 274-277, 2011

Atkinson, A. Desigualdade: o que pode ser feito? LeYa, São Paulo, 2015

Atkinson, A.; Bourguignon, F. Handbook of Income Distribution. Elsevier, First edition, Volume 2, 2015

Atkinson, A. Pareto and the upper tail of the income distribution in the UK: 1799 to the present. Economica, 84 (334), p. 129-156, 2016

Atkinson, A.; Piketty, T. Top incomes over the twentieth century. Oxford: Oxford University Press, 2007

Atkinson, A.; Piketty, T.; Saez, E. Top incomes in the long run of history, J. Econ. Lit. 49 (1), p. 3-71, 2011

Becker, G.S., 1964. Human Capital. Columbia University Press, New York, second ed., 1975

Becker, G; Chiswick, B. Education and the distribution of earnings. American Economic Review 56. 1966

Benhabib, J.; Bisin, A.; Zhu, S. The distribution of wealth and fiscal policy in economies with finitely lived agents, Econometrica 79 (1), p. 123-157, 2011

Blanchet, T.; Fournier, J.; Piketty, T. Generalized Pareto curves? Theory and applications, World Inequality Lab Working Paper, No. 2017-3, 2017

Blanchet, T.; Garbinti, B.; Goupille-Lebret, J.; Martínez-Toledano, C. Applying Generalized Pareto Curves to Inequality Analysis. WID.world Working Paper No. 2018/3, 2018

Clauset, A.; Shalizi, C.; Newman, M. Power-law distributions in empirical data. SIAM Rev., 51, 2009

Chakrabarti, B.; Chakraborti, A.; Chakravarty, S.; Chatterjee, A. Econophysics of Income and Wealth Distributions. Cambridge University Press, 2013

Clementi, F.; Gallegati, M. Power law tails in the Italian personal income distribution. Physica A 350 (2-4), 2005

David, H. Order Statistics, John Wiley & Sons, New York, 1970

Dorea, C.; Otiniano, C.; Matsushita, R.; Rathie, P. Lévy flight approximations for scaled transformations of random walks, Comput. Stat. Data An. 51 (12), p. 6343-6354, 2007

Dragulescu, A.; Yakovenko, V. Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States, Physica A 299 (1-2), p. 213-221, 2001

Fontanari, A.; Taleb, N.; Cirillo, P. Gini estimation under infinite variance, Physica A 502, p. 256-269, 2018

Friedman, M. Choice, chance, and the personal distribution of income. Journal of Political Economy, Vol. 61, No. 4, 1953

Friedman, M; Kuznets, S. Income from Independent Professional Practice. National Bureau of Economic Research, New York, 1945

Gabaix, X. Power laws in economics: An introduction. Journal of Economic Perspectives 30 (1), p. 185-206, 2016

Gabaix, X.; Ibragimov, R. Rank-1/2: A simple way to improve the ols estimation of tail exponents. Journal of Business and Economic Statistics 29, 2011

Gabaix, X.; Lasry, J.; Lions, P.; Moll, B. The dynamics of inequality, Econometrica 84 (6), p. 2071-2111, 2016

Gibrat, R. Les Inégalités Économiques, Sirey, Paris, 1931

Hill, B. A simple general approach to inference about the tail of a distribution. Annals of Statistics, 3 (5), p. 1163–1174, 1975

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

Kalecki, M. On the Gibrat distribution, Econometrica 13 (2), p. 161-170, 1945

Limpert, E.; Stahel, W.A.; Abbt, M. Log-normal Distributions across the Sciences: Keys and Clues. BioScience, Volume 51, Issue 5, p. 341-352, 2001

Lucas, R.; Moll, B. Knowledge growth and the allocation of time, J. Polit. Econ. 122 (1), p. 1-51, 2014

Mandelbrot, B. The Pareto-Lévy law and the distribution of income, Int. Econ. Rev. 1 (2), p. 79-106, 1960

Mandelbrot, B. The stable Paretian income distribution when the apparent exponent is near two, Int. Econ. Rev. 4 (1), p. 111-115, 1963

Mandelbrot, B. The variation of certain speculative prices. Journal of Business 36 (4), 1963

Marshall, A. Principles of Economics. 8th edition, 1920, Macmillan, London, 1890

Mill, J. Principles of Political Economy. Collected Works of John Stuart Mill, volumes 2-3, 1965, University of Toronto Press, 1848

Mood, A.; Graybill, F.; Boes, D. Introduction to the Theory of Statistics, McGraw-Hill, New York, 1974

Newman, M. Power laws, Pareto distributions and Zipf's law, Contemp. Phys. 46 (5), p. 323-351, 2005

Pareto, V. Cours d'Economie Politique, Rouge, Lausanne, 1897

Pareto, V. Manual of Political Economy. 1971, Macmillan, London. 1909

Piketty, T. Capital in the Twenty-First Century. Cambridge: Harvard University Press, 2014

Piketty, T.; Zucman, G. Capital is back: Wealth-income ratios in rich countries, 1700-2010, Q. J. Econ. 129 (3), p. 1255-1310, 2014

Roy, A. The distribution of earnings and of individual output. The Economic Journal, Vol. 60, No. 239, 1950

Roy, A. Some thoughts on the distribution of earnings. Oxford Economic Papers, Vol. 3, No. 2, 1951

Ruiz, N.; Woloszko, N. What do household surveys suggest about the top 1% incomes and inequality in OECD countries? Economics Department Working Paper No. 1265. Paris: OECD, 2015

Sen, A. On Economic Inequality. Clarendon Press, Oxford. 1973

Smith, A. An Inquiry into the Nature and Causes of the Wealth of Nations. 1976, Oxford University Press. 1776

Taleb, N. The Black Swan: The Impact of the Highly Improbable. Random House: New York. 2010

Taleb, N.; Douady, R. On the super-additivity and estimation biases of quantile contributions, Physica A 429, p. 252-260, 2015

Tao, Y. Swarm intelligence in humans: A perspective of emergent evolution, Physica A 502, p. 436-446, 2018

Tao, Y.; Wu, X.; Zhou, T.; Yan, W.; Huang, Y.; Yu, H.; Mondal, B.; Yakovenko, V. Exponential structure of income inequality: Evidence from 67 countries, J. Econ. Interact. Coord. 14 (2), p. 345-376, 2019

Toda, A.; Walsh, K. The double power law in consumption and implications for testing Euler equations, J. Polit. Econ. 123 (5), p. 1177-1200, 2015

von Neumann, J.; Morgenstern, O. Theory of Games and Economic Behavior, second ed. Princeton University Press, Princeton, 1947

Walras, L. Elements of Pure Economics. Irwin, Homewood, IL. 1874-1877, 1954

APPENDIX

Elementary School Portuguese Teacher

	Percentiles	Smallest		
1%	0	0		
5%	672.74	0		
10%	1097.45	0	Obs	5,442
25%	1888.06	0	Sum of Wgt.	5,442
50%	3089.465		Mean	3422.304
		Largest	Std. Dev.	2512.233
75%	4141.02	20531.01		
90%	5737.94	20671.4	Variance	6311317
95%	7482	26463.54	Skewness	2.64903
99%	15039.9	26514.69	Kurtosis	14.50088

Fashion Designer

	Percentiles	Smallest		
1%	0	0		
5%	1205.86	0		
10%	1369.355	0	Obs	2,260
25%	1852.72	0	Sum of Wgt.	2,260
50%	2958.155		Mean	4009.788
		Largest	Std. Dev.	3563.587
75%	5059.505	27999.99		
90%	7926.775	29374.07	Variance	1.27e+07
95%	10194.49	37107.15	Skewness	3.590398
99%	18212.65	51419.04	Kurtosis	28.05365

Lawyer

	Percentiles	Smallest		
1%	0	0		
5%	1516.25	0		
10%	2000	0	Obs	23,291
25%	3113.82	0	Sum of Wgt.	23,291
50%	5570.05		Mean	6874.834
		Largest	Std. Dev.	6231.704
75%	7994.3	92093.98		
90%	13008.87	93778.86	Variance	3.88e+07
95%	17679.01	107441.9	Skewness	3.959314
99%	31260.92	111634.4	Kurtosis	32.94417

Music Performer

	Percentiles	Smallest		
1%	403.05	0		
5%	903.82	0		
10%	1145.015	0	Obs	2,580
25%	1746.215	0	Sum of Wgt.	2,580
50%	3390.99		Mean	5225.587
		Largest	Std. Dev.	4420.112
75%	9716.49	18750		
90%	11351.04	20187.5	Variance	1.95e+07
95%	13346.29	22194.65	Skewness	.9916074
99%	17493.25	29864.94	Kurtosis	3.027481

Radio and TV Broadcaster

	Percentiles	Smallest		
1%	0	0		
5%	369.67	0		
10%	617.04	0	Obs	1,890
25%	1141.86	0	Sum of Wgt.	1,890
50%	1544.095		Mean	2169.442
		Largest	Std. Dev.	2395.544
75%	2462.06	25248.24		
90%	3830.575	25482.07	Variance	5738632
95%	5732.77	26544.62	Skewness	5.109324
99%	12039.74	27330.54	Kurtosis	40.87181

Visual Artist

	Percentiles	Smallest		
1%	0	0		
5%	910.8	0		
10%	1187.27	0	Obs	1,326
25%	1464.74	0	Sum of Wgt.	1,326
50%	2020.655		Mean	2791.406
		Largest	Std. Dev.	2076.794
75%	3499.99	11713.94		
90%	5685.47	13570	Variance	4313074
95%	7346.03	13954.72	Skewness	1.885234
99%	9900	14537.5	Kurtosis	7.223483

Actor

	Percentiles	Smallest		
1%	0	0		
5%	2662.5	0		
10%	4000.5	0	Obs	891
25%	7046.65	0	Sum of Wgt.	891
50%	14674.25		Mean	24250.49
		Largest	Std. Dev.	26045.69
75%	30000	134801.7		
90%	60283.47	135975.6	Variance	6.78e+08
95%	90430.57	137061.4	Skewness	1.952031
99%	114773.7	140524.4	Kurtosis	6.589535

Architect

	Percentiles	Smallest		
1%	0	0		
5%	2546.91	0		
10%	4191.77	0	Obs	1,973
25%	6600	0	Sum of Wgt.	1,973
50%	9370		Mean	11629.86
		Largest	Std. Dev.	7817.682
75%	14983.63	58596.77		
90%	22227.78	61635.6	Variance	6.11e+07
95%	26366.87	63048.43	Skewness	1.852732
99%	37254.57	73134	Kurtosis	9.284978

Bank Clerk

	Percentiles	Smallest		
1%	0	0		
5%	937	0		
10%	2790.63	0	Obs	18,442
25%	4134.55	0	Sum of Wgt.	18,442
50%	6948.905		Mean	8209.523
		Largest	Std. Dev.	5984.19
75%	10747.83	94214.1		
90%	15396.2	94851.6	Variance	3.58e+07
95%	19356.68	105038.7	Skewness	2.477931
99%	26274.55	115233	Kurtosis	22.81583

Dentist

	Percentiles	Smallest		
1%	1029.58	0		
5%	1623.76	0		
10%	2070.18	0	Obs	5,875
25%	3068.73	0	Sum of Wgt.	5,875
50%	4446.98		Mean	5047.19
		Largest	Std. Dev.	3216.743
75%	6092.17	40800.17		
90%	8728.8	40892.88	Variance	1.03e+07
95%	10452.57	41742.33	Skewness	3.198155
99%	16539.52	44584.25	Kurtosis	25.40443

Footballer

	Percentiles	Smallest		
1%	0	0		
5%	855.55	0		
10%	937	0	Obs	1,374
25%	992.75	0	Sum of Wgt.	1,374
50%	1213.71		Mean	9245
		Largest	Std. Dev.	22720.97
75%	3666.66	136000		
90%	22657.14	137492	Variance	5.16e+08
95%	65601.85	138000	Skewness	3.571644
99%	116912	138431.4	Kurtosis	15.68053

Motorcycle Messenger

	Percentiles	Smallest		
1%	0	0		
5%	712.39	0		
10%	1000.42	0	Obs	17,429
25%	1328	0	Sum of Wgt.	17,429
50%	1468.72		Mean	1467.839
		Largest	Std. Dev.	1263.12
75%	1598.67	57038.4		
90%	1810.81	64170.3	Variance	1595472
95%	2042.31	70356	Skewness	42.79039
99%	2725.65	80561.4	Kurtosis	2241.216