ARTIGO **01 • 2020**

Indicadores de Desigualdades Fiscais e de Rendimento

Alexandre Mergulhão

Abstract

This paper presents a comprehensive review of inequality, redistribution and progressivity measures, developed by the literature over the years. Providing an analytical, conceptual and graphical framework, it explains each of these multiple indicators and discusses differences among them. It is intended to serve as a small manual for people interested in these topics, hopefully enhancing public administration's capacity. It is also the methodological support of work presented to the Superior Council of Statistics, where results based on Portugal's administrative personal income tax data are analysed.

Resumo

Este artigo apresenta uma revisão abrangente das medidas de desigualdade, redistribuição e progressividade que têm vindo a ser desenvolvidas pela literatura académica. Com base num enquadramento analítico, conceptual e gráfico, explica cada indicador e discute as diferenças entre os múltiplos existentes. Este trabalho tenciona constituir um pequeno manual para pessoas interessadas em estudar estes temas, bem como servir de suporte metodológico da apresentação no GT-IDS do Conselho Superior de Estatística, onde são analisados os resultados, calculados com base em dados administrativos do Imposto sobre o Rendimento das pessoas Singulares (IRS).

Palavras-chave: metodologia, indicadores, desigualdade, progressividade, redistribuição.

"[t]he fiscal history of a people is above all an essential part of its general history. […] The spirit of a people, its cultural level, its social structure, the deeds its policy may prepare — all this and more is written in its fiscal history, stripped of all phrases. He who knows how to listen to its message here discerns the thunder of world history more clearly than anywhere else." Joseph Schumpeter (1918)

Gabinete de Planeamento, Estratégia, Avaliação e Relações Internacionais
Ministério das Finanças
Rua da Alfândega n.º 5ª • 1100 – 016 Lisboa
www.gpeari.gov.pt

1. Introdução

Este trabalho tem como objetivo servir de manual metodológico para análises de distribuição, com o foco em rendimentos e impostos diretos. Assim, ambiciona dotar a administração pública de uma súmula sobre indicadores de desigualdades, bases de dados disponíveis e ferramentas de *softwares* estatísticos, para estudos de distribuição mais aprofundados.

A investigação sobre questões de distribuição, na ciência económica, foi relativamente marginal e tardia, contando com os trabalhos pioneiros de Hugh Dalton (1920's), Simon Kuznets (1950's) e Anthony Atkinson (1970's). No entanto, o interesse público e académico em relação ao nível, à evolução, causas e consequências das múltiplas dimensões das desigualdades sociais tem vindo a aumentar substancialmente, desde os anos de 1990. Em Portugal, destacam-se os trabalhos de Carlos Farinha Rodrigues (*e.g.* 2016 ou 1993).

Os avanços significativos na qualidade e abrangência de dados provenientes de inquéritos às famílias, internacionalmente estandardizados, tornaram possível a constatação do aumento das desigualdades de rendimento e riqueza, na maioria dos países (ver Branko Milanovic, 2002). Mais recentemente, alguns investigadores dedicaram-se ao estudo das desigualdades concentrando-se nos extremos mais elevados das distribuições – *e.g.* percentagem do rendimento total detida pelos *top* 5%, 1% e 0,1% – recorrendo a dados administrativos, nomeadamente das autoridades tributárias nacionais (ver Piketty, Zucman, Saez, Alvaredo *etc.*).[1] Estes trabalhos culminaram na criação da *World Inequality Database* (2011 e 2015) que pretende criar séries históricas de Contas Nacionais de Distribuição, combinando dados de inquéritos, dados fiscais e de Contabilidade Nacional, duma forma sistemática, rigorosa e transparente.

De acordo com os dados oficiais do Eurostat, oriundos do Inquérito às Condições de Vida e Rendimento (EU-

[1] Estas metodologias pretendem colmatar limitações, não só do facto das medidas convencionais de desigualdade – como o coeficiente de Gini – serem pouco sensíveis a dispersões nos extremos da distribuição, assim como aquelas, relacionadas com a natureza dos inquéritos – *e.g.* subestimação dos rendimentos mais elevados e baixa representatividade em relação à Contabilidade Nacional.

SILC), Portugal tem pertencido sistematicamente ao *top* 3 de desigualdade na União Europeia.[2] Contudo, as desigualdades sociais acarretam um carácter multidimensional, impossibilitando uma análise plena recorrendo apenas a indicadores de rendimento.

Posto isto, o Conselho Superior de Estatística constituiu, em setembro de 2017, o Grupo de Trabalho sobre Indicadores das Desigualdades Sociais (GT-IDS), com vista à produção de indicadores variados. Este grupo conta com a participação de diversas entidades públicas, assim como parceiros sociais, de entre as quais o GPEARI, que coordena o subgrupo de Fiscalidade e Rendimento. Neste contexto, a Autoridade Tributária e Aduaneira (AT) disponibilizou as bases de dados anonimizadas das Notas de Liquidação do IRS de 2016 e 2017. Estas bases de dados foram utilizadas para construir os indicadores referidos neste artigo, tendo sido os respetivos resultados, e propostas de novos indicadores, discutidos no GT-IDS.

2. Enquadramento Conceptual, Analítico e Gráfico

2.1. Enquadramento conceptual

A desigualdade monetária diz respeito às diferenças relativas ou absolutas na distribuição entre unidades. Para efeitos de análise, estas unidades podem representar: indivíduos (ou sujeitos passivos), agregados fiscais, agregados familiares ou médias de países, etc. Por sua vez, as variáveis monetárias, como os rendimentos, assumem diferentes categorias. Como podemos ver no Quadro 1, o enquadramento adaptado da OCDE define as seguintes componentes de rendimento.

Os **rendimentos de mercado** constituem salários e retribuições do trabalho, rendimento do trabalho por conta própria, rendas, juros, mais-valias e dividendos. O **rendimento disponível** é o resultado após todos os compromissos e benefícios do Estado – compostos respetivamente por impostos[3] (IRS) e contribuições sociais (CSS) a pagar, e por transferências sociais, com destaque para as pensões públicas, a receber.

Quadro 1 - Enquadramento Conceptual do Rendimento
Componentes do Rendimento
Ordenados e salários
+ Rendimentos do trabalho por conta-própria
+ Rendimentos de propriedade
+ Outros rend. privados (bónus, dividendos, etc.)
Rendimentos de Mercado/Primário
+ Pensões
Rendimentos de Mercado + Pensões
+ Outras transferências monetárias da SS
Rendimento Bruto/Original
- Contribuições do trab. para a SS
- Impostos sobre o rendimento
- Outras transf. correntes pagas
Rendimento Disponível

Quadro 1 – Enquadramento conceptual das componentes do rendimento de pessoas singulares (indivíduo, agregado fiscal ou familiar). Fonte: Adaptado da OCDE (2008) e de Carlos Farinha Rodrigues (2019).

Define ainda as instâncias entre estes dois extremos, como é o caso do **rendimento bruto** (ou original) que representa o rendimento disponível antes do pagamento de IRS e CSS, mas após receção de transferências sociais. Este é o enquadramento conceptual presente nos inquéritos às famílias que são a fonte das estatísticas oficiais de desigualdade.

Contudo, os conceitos presentes nas bases de dados da AT têm diferenças que merecem ser clarificadas. Primeiro, o conceito de agregado fiscal agrupa as pessoas que declaram rendimentos (conjunta ou separadamente) e seus dependentes associados, ao passo que um agregado familiar agrupa pessoas que residem na mesma habitação (Decoster, 1998). Logo, o agregado familiar pode ser composto por vários agregados fiscais. Em segundo lugar, os inquéritos baseiam-se em amostras representativas da população (no mínimo 10,5 mil pessoas), enquanto a AT inclui todas as declarações de rendimentos. Terceiro, em Portugal, muitas prestações sociais não são tributáveis e, por isso, não são declaradas no IRS, enquanto os inquéritos contabilizam todos os rendimentos do Quadro 1. Para além disso, o código do IRS contempla a opção de não englobamento para rendimentos que não sejam do trabalho ou pensões. Isto significa que, estes dados da AT não incluem

[2] Entre 2004 e 2012, Portugal pertenceu ao *top* 3 dos países da UE em termos de desigualdade de rendimento disponível – após impostos, contribuições e todas as transferências sociais, incluindo as pensões – com a exceção de 2008 onde registou o quarto Gini mais elevado. Entre 2012 e 2017, foi quase sempre o segundo país mais desigual em termos de rendimento de mercado – antes de impostos e contribuições, e também de qualquer transferência social.

[3] Doravante impostos referem-se a impostos diretos (IRS).

rendimentos de capitais (e.g. juros e dividendos), prediais (rendas) nem incrementos patrimoniais (mais-valias) elevados visto que é racional não englobar e pagar apenas 28% de taxas liberatórias sobre cada categoria individual. Assim, o rendimento bruto da AT, na verdade, aproxima-se dos rendimentos de mercado (do trabalho) mais pensões. Acresce o facto de a declaração de rendimentos ser condição para a elegibilidade de algumas transferências socias, o que explica a presença de vários agregados fiscais com rendimentos brutos extremamente baixos.[4]

2.2. Enquadramento analítico e gráfico

Para qualquer variável contínua (*e.g.* rendimento bruto) temos uma distribuição definida por um vetor ordenado $x = (x_1, x_2, \dots, x_n, \dots, x_{max})$ onde, por exemplo, x_5 representa o rendimento da 5.ª unidade (*e.g.* agregado fiscal) mais pobre. Este vetor tem uma **função densidade** $f(x)$ que, basicamente, estima a proporção da população com rendimento x.[5]

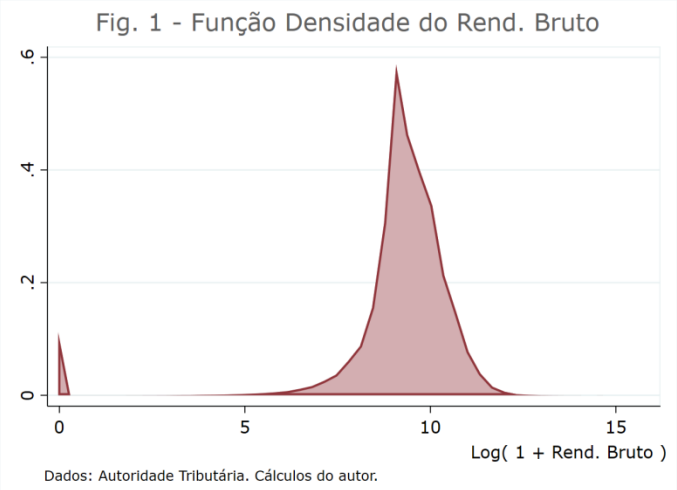

Figura 1 – Função densidade (kernel) do rendimento bruto declarado para fins tributários, em sede de IRS, com base nos dados da Nota de Liquidação de IRS de 2017. Efetuou-se uma transformação logarítmica (natural) à variável para fins de visualização. Por exemplo, o valor 10 representa rendimentos de 22 mil euros anuais, aproximadamente.

O integral, ou a área vermelha de zero até cada ponto, da anterior função designa-se por **função de distribuição cumulativa** $F(x) = \int_{x_1}^{x_{max}} f(x)dx$ que atribui a proporção $(0 \leq F(x) \leq 1)$ da população com rendimentos iguais ou inferiores a x.

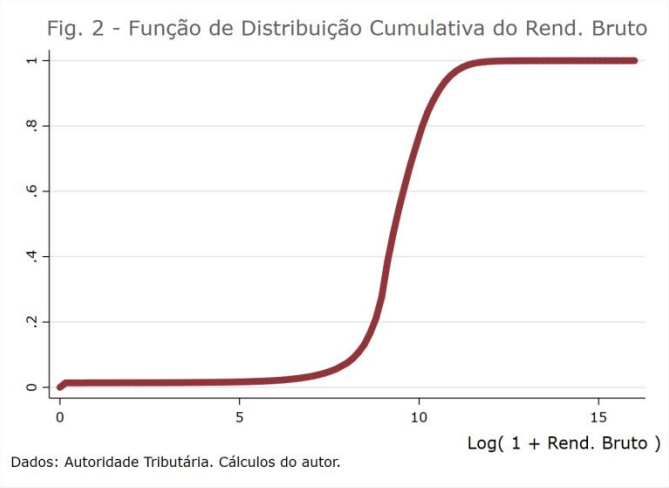

Figura 2 – Função de distribuição cumulativa é igual ao integral (inverso de derivada) da densidade. Assim, representa a progressão da área da figura 1. Por exemplo, a imagem do valor 10 é aproximadamente 0,7, significando que 70% dos agregados fiscais têm rendimentos iguais ou inferiores a cerca de 22 mil euros anuais.

Finalmente, uma distribuição pode ainda ser expressa em percentagem do total, em vez de valores absolutos, constituindo uma **função da proporção do rendimento total cumulativa**:

$$\Phi(x) = \frac{1}{\mu}\int_{x_1}^{x_n} x.f(x)dx = \frac{"média\ de\ x_1\ a\ x_n"}{"média\ total"}$$

$$= \frac{\int_{x_1}^{x_n} x.f(x)dx}{\int_0^{x_{max}} x.f(x)dx} = \frac{\int_{x_1}^{x_n} \phi(x)dx}{\int_0^{x_{max}} \phi(x)dx} = \frac{"\%\ x_1\ a\ x_n"}{100\%}$$

Onde o denominador é a média global (μ rendimento médio) e $\phi(x)$ é a função que atribui a percentagem do rendimento total detida pelo indivíduo com o rendimento x.

Esta representação da distribuição pode ser melhor compreendida quando interpretada por **quantis**. Qualquer distribuição pode ser dividida em partes, com igual frequência/probabilidade p, através de quantis. Genericamente, quantis são a função $q(p) = x$ que atribui o valor de x (*e.g.* rendimento) tal que $F(x) = p$.

[4] Por fim, o total de IRS efetivamente pago designa-se por coleta líquida.

[5] Em bom rigor, uma função densidade probabilística é estimada através de um método. Estatisticamente, recorre-se a uma estimação Kernel que atribui a probabilidade não do ponto em si, mas sim dos pontos vizinhos.

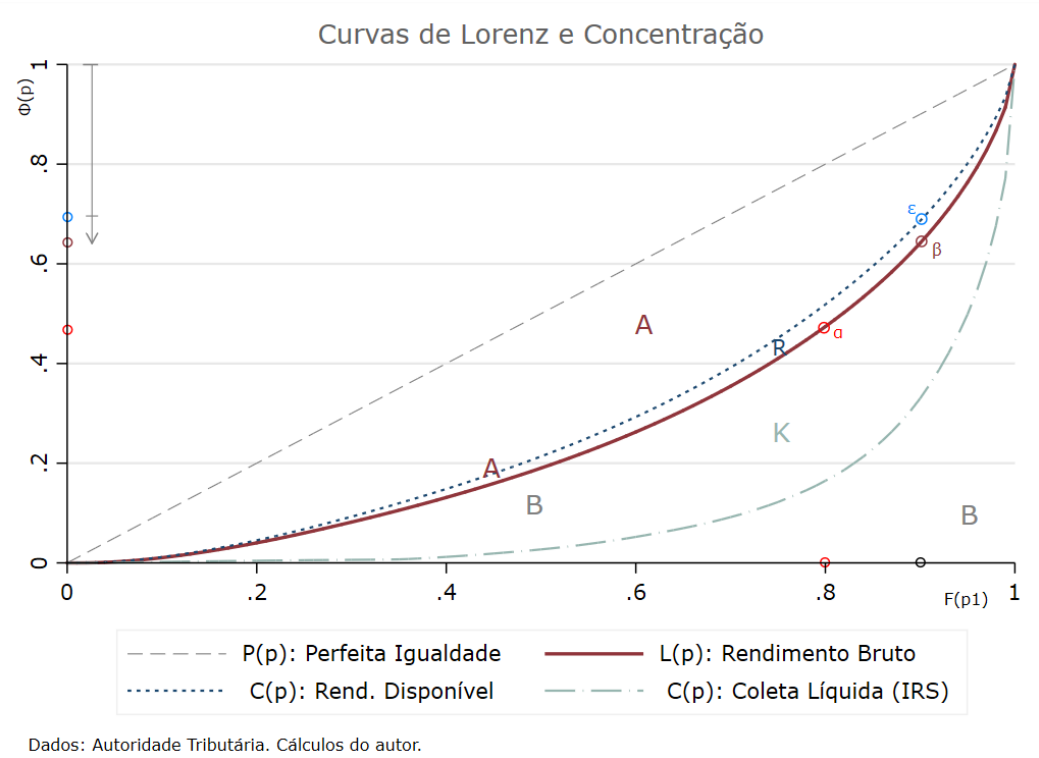

Figura 3 – Curvas de Lorenz do rendimento bruto e perfeita igualdade, e curvas de concentração (ordenadas pelo rendimento bruto) para o rendimento disponível (i.e. bruto após IRS) e para a Coleta Líquida (IRS efetivamente pago). A área A está compreendida entre a diagonal P(p) e a curva de Lorenz do rendimento bruto.

Portanto, encontramos a proporção p partindo do menor valor (x_1) e terminando no quantil $q(p)$, que é um nível de rendimento (x). Assim, decis são um exemplo de quantis onde p toma o valor de 0,1, dividindo a distribuição em 10 intervalos. Quartis, quintis, vintis e percentis dividem em quatro (p =0,25), cinco (p =0,2), vinte (p =0,05) e cem (p =0,01) partes iguais, respetivamente. Por fim, a mediana é o quantil que delimita o x a partir do qual atingimos metade (ou o "meio") da distribuição $(F(x) = 0{,}5)$, e pode ser expresso em percentis (50.º), vintis (10.º) e quartis (2.º), mas não em quintis, por exemplo.

2.3. Curvas de Lorenz e de Concentração

Chegamos finalmente, ao conceito de **curva de Lorenz** (Max O. Lorenz, 1905) que constitui a base da grande maioria dos indicadores de desigualdade. Esta curva é a representação gráfica entre a distribuição cumulativa da população $F(x)$ e a função cumulativa da proporção do rendimento total $\Phi(x)$:

$$\mathrm{L}(p) = \Phi(F(x))$$

A curva diz-nos qual é a fração do rendimento total $\Phi(x) = p_2$ detida pelos $F(x) = p_1$ por cento mais *pobres*. Logo, a reta $p_2 = p_1$ representa a situação de perfeita igualdade, uma vez que, não só 50% da população detém 50% do rendimento total mas o mesmo sucede para qualquer percentagem, implicando que toda a população recebe o rendimento médio.

Resumindo, a curva de Lorenz deve ser lida da esquerda para a direita, onde para cada ponto p_1 (percentagem da população mais *pobre*) temos a imagem em percentagem do rendimento total que aufere.[6] Logo, quanto mais convexa for a curva (mais afastada da reta diagonal) maior é a desigualdade porque uma percentagem maior da população mais *pobre* tem menos do rendimento total.

[6] Doravante, por facilidade de linguagem, entenda-se Lorenz por curva de Lorenz e por mais *pobres* as unidades com os rendimentos mais baixos (*i.e.* de zero até certo ponto – valor ou percentagem). Pela mesma razão, coeficiente/índice de Gini é abreviado para Gini.

Os indicadores analisados de seguida devem ser lidos tendo sempre como referência visual a Figura 3 apresentada acima.

3. Indicadores sumários de desigualdade

Uma vez compreendido o enquadramento conceptual, analítico e gráfico é possível apreender a intuição subjacente à grande maioria dos indicadores de desigualdade.

3.1. Coeficientes de Gini e de Concentração

A divergência da curva de Lorenz $\mathrm{L}(p)$ em relação a reta de perfeita equidade pode ser medida por índices de desigualdade, dos quais o **coeficiente de Gini** (Corrado Gini, 1912) é o mais popular. Seguindo o raciocínio anterior, na reta $P(p_1) = p_2$, os 20% mais *pobres* recebem 20% do rendimento total, no entanto, esta situação hipotética nunca se verifica. Empiricamente, os x% mais *pobres* recebem $(x-a)$% do rendimento total, e quanto maior for o a maior será o desvio entre a reta e a curva de Lorenz $\int_0^1 [P(p_1) - \mathrm{L}(p_1)]dp_1$, e portanto, o índice de Gini:[7]

$$G(x) = \frac{\int_0^1 [P(p_1) - \mathrm{L}(p_1)]dp_1}{\int_0^1 P(p_1)dp_1} =$$

$$= \frac{A}{triângulo(A+B)} = 2A$$

Logo, o índice Gini é igual ao dobro da área A. Isto porque $\Phi(x)$ e $F(x)$ representam frações – *i.e.* variam entre 0 e 1 –, o que implica que a área do $triângulo(A+B) = 0{,}5$ e que $0 \leq \mathrm{G}(x) \leq 1$ – onde o máximo representa o cenário hipotético em que apenas uma unidade detém todo o rendimento.

De forma a não confundir curvas de Lorenz com índice de Gini repare-se que a equação da área de $A + B = 0{,}5$ implica que $2A + 2B = 1 \Leftrightarrow 2A = 1 - 2B = 1 - 2\int_0^1 L(p_1) = G(x)$. Ou seja o Gini é igual a um menos o dobro da área abaixo da curva de Lorenz (i.e. área B).

Imaginemos agora que queríamos saber quanto IRS, em percentagem do total, é que a metade mais *pobre* paga efetivamente. A função que nos dá a resposta para os 50% mais *pobres*, ou qualquer outra percentagem, é a **curva de concentração** da coleta líquida $\mathrm{C}(p_1)$, ordenada pela $F(x)$ do rendimento bruto. Uma curva de concentração é equivalente à de Lorenz, com a diferença de que a variável de interesse não é a mesma que a variável da população: devemos explicitar a variável de ordenação.[8] Caso contrário, assume-se que a ordem é a mesma e, seguindo o exemplo anterior, saberíamos quanto IRS é que a metade que paga menos imposto efetivamente o faz. Portanto, o **índice de concentração** de uma dada variável é o indicador sumário de desigualdade "Gini", mas com a ordem cumulativa de outra variável. Conclui-se, então, que a Lorenz é o caso específico das curvas de concentração onde apenas se usa uma variável.

3.2. Indicadores de Redistribuição e Progressividade

O conceito de redistribuição prende-se com as diferenças relativas ao longo da distribuição, antes e depois de um número de medidas orçamentais. Este tem subentendido o conceito de progressividade (positiva) na medida em que é expectável que a desigualdade após Estado seja menor do que a original: os percentis mais altos passem a ter menos rendimento em relação aos mais baixos. Desta feita, alguma literatura refere-se a indicadores de redistribuição e progressividade como sinónimos. Contudo, é mais correto referir o primeiro aquando da comparação entre rendimento bruto e disponível, e empregar o segundo aquando da comparação entre rendimento bruto e as responsabilidades tributárias.

Para analisar a redistribuição, o indicador mais utilizado é o **Reynolds-Smolensky** (1977), que compara o Gini antes do Estado com o Gini após transferências e impostos, da forma mais simples: $RS = G_X - G_N$.[9] Este também é apelidado de

[7] Existem outras fórmulas do Gini equivalentes como a decomposição entre intra e inter-grupos (Milanovic, 2010) ou a "convinence covariance" (Kakwani, 1980). A formulação escolhida é a mais intuitiva em termos gráficos.

[8] Doravante a variável de ordem será o rendimento bruto (*i.e* ordenada por $F(p_1)$), exceto indicação contrária.

[9] Terminada a explicação genérica do enquadramento analítico, passarei a denominar a distribuição do rendimento bruto (*gross)* por X e a do rendimento disponível (*net*-IRS) por N.

redistribuição absoluta (*e.g.* Frederick Solt, 2019), tomando a forma *relativa* quando se divide o anterior resultado pelo G_X.[10] Outras medidas são, basicamente, variantes: *progressão efetiva* de Musgrave-Thin (1948) $MT = (1 - G_N)/(1 - G_X)$; e Pechman-Ocker (1974) $PO = RS/G_X = 1 - G_N/G_X$.

Uma medida menos usual é o **coeficiente de Pietra**[11] que estima a proporção do rendimento total que teria de ser redistribuída das famílias acima da média para aquelas abaixo da média, de forma a atingir o estado de perfeita igualdade. Por isso, também é conhecido como **índice Robin Hood**. Na Figura 3, podemos visualizá-lo como a maior distância vertical entre a diagonal e a Lorenz – que ocorre no ponto em que a derivada da curva é igual a um. Alternativamente, é igual ao dobro da área do maior triângulo inscrito na área A, onde o vértice é o ponto anterior $L'(p_1) = 1$: logo, é sempre menor ou igual ao Gini (Fellman, 2018), que como vimos é o dobro da área A. Em termos práticos existem várias maneiras de o estimar:

$$Robin\ Hood = \frac{\frac{1}{n}\sum_{i=1}^{n}|x_i - \mu|}{2\mu} \quad \textit{onde n é \# observações}$$

$$\frac{Desvio\ absoluto\ médio}{2 \times média\ total} = \frac{\sum_{\{i:\, x_i \geq \mu\}}(x_i - \mu)}{\sum_{i=1}^{n} x_i}$$

$$= \left[\sum_{1}^{u} shares\ dos\ \#\frac{1}{p}\ intervalos\ de\ q(p) > p\%\right] - [u \times p]\%$$

Quadro 2 - Distrib. das fatias do rendimento por decis					
Decil	1	2	3	4	5
share/ fatia	1,08	2,48	4,13	5,74	7,33
Decil	6	7	8	9	10
share/ fatia	8,98	10,83	13,09	16,41	29,93

Quadro 2 – Exemplo duma distribuição de rendimento, expressa em percentagem, do total, detida por cada intervalo de decis. Fonte: Kucharski, "Gini index and Robin Hood índex" slides.

Tomando decis como exemplo de intervalos de quantis $p = 0{,}1$, a forma mais direta e prática de o calcular seria somar as *shares*/fatias (%) dos decis superiores a 10% (*e.g.* $u = 3$ casos) e subtrair $3 \times 10\%$. Assim, aplicando a última fórmula ao exemplo do Quadro 2, teríamos (10,83% + 13,09% + 16,41% + 29,93%) – (4 x 10%) = 70,26% – 40% = 30,26%.

Voltando às medidas usuais de redistribuição, importa destacar uma dimensão oculta nestes indicadores. A **curva de concentração do rendimento disponível** (com a população ordenada pelo rendimento bruto) diz-nos quanto é que os p_1% mais *pobres* recebem após transferências e impostos. Por outro lado, a **curva de Lorenz do rendimento disponível** diz-nos quanto é que os p_3% mais *pobres* recebem, também, após transferências e impostos. O ponto é o seguinte: mesmo que o valor de p_1 e p_3 seja igual, por exemplo 20, a primeira refere-se aos 20% mais *pobres* em termos de rendimento bruto e a segunda em termos de rendimento disponível. Estas frações não representam a mesma população, visto que a ordem muda (*re-ranking*) (*e.g.* a pessoa mais rica antes de se pagar IRS pode não ser a mais rica após o pagamento do imposto).[12]

Desta feita, temos a decomposição do índice RS em **redistribuição/equidade vertical** (VE) e reordenação (re):

$$RS = G_X - G_N = VE - re =$$
$$= [G_X - C_N] - [G_N - C_N]$$

Em termos de áreas da Figura 3:

$$VE = 2A - [2A - 2R] = 2R$$

O efeito de **reordenação capta a dimensão da desigualdade horizontal** (Atkinson, 1980) que diz respeito ao impacto desigual que uma política orçamental tem em agregados que não são diferentes. Esta dimensão seria visualizada no gráfico pelo dobro da área entre a curva de concentração e a de Lorenz do rendimento disponível – que estaria entre a primeira e a Lorenz do rendimento bruto, diminuindo o efeito redistributivo da política. Caso não haja reordenação a curva de concentração seria a mesma que a Lorenz do rendimento disponível. Contudo, não é possível apresentar esta última porque teria um eixo horizontal diferente.

[10] Assim, o Reynolds-Smolensky relativo estima a percentagem da desigualdade inicial (e.g. de mercado) que é eliminada pelas políticas fiscais e sociais do Estado – ou por apenas uma medida.
[11] Também designado como índice de: Hoover (1936), Schutz (1951), Ricci (1879-1946) e Lindahl (1891-1960).
[12] Isto pode dever-se a vários motivos como diferenças no montante de deduções à coleta consoante o número de filhos ou o tipo de despesas realizadas, mas também por benesses menos óbvias.

Caixa - Taxas e Cálculo do IRS

Um equívoco muito disseminado na opinião pública prende-se com o mito de o rendimento disponível diminuir por via da subida de escalão, após ter ganho mais rendimento bruto num ano. Na verdade, o IRS é calculado de forma escalonada ("em escadinha") aplicando-se as taxas marginais apenas ao excedente em relação ao limite do escalão anterior.

Ao aplicarmos a taxa marginal ao total do rendimento, como se tende a fazer, teremos que subtrair a parcela a abater – que muitas vezes não é apresentada.

Escalões 2017	Rendimento Coletável	Taxa marginal	Taxa média	Parcela a abater
1	<= 7.091	14,5%	14,5%	-
2	> 7.091	28,50%	23,6%	992,74
3	> 20.261	37%	30,3%	2.714,93
4	> 40.522	45%	37,61%	5.956,69
5	> 80.640	48%	-	8.375,89

Tomemos como exemplo uma pessoa que auferiu um rendimento coletável anual de 25.000€. A coleta bruta (antes das deduções à coleta) seria 7.091 x 14,5% + (20.261–7.091) x 28,5% + (25.000–20.261) x 37%= 1.028,195 + 3.753,45 + 1.753,43 = 6.535,075€. A taxa marginal do escalão em que a pessoa se situa (3.º escalão) apenas se aplica ao excedente de 4.739 e não ao total – e a anterior (28,5%) a 13.170. Equivalentemente, aplicando a parcela a abater, teríamos 25.000 x 37% – 2.714,93 = 6.535,075€.

Logo **o rendimento disponível nunca diminuirá pela subida de escalão**.

O terceiro método aplica a taxa média do escalão anterior (2.º) ao seu limite superior e a taxa marginal do escalão da pessoa (3.º) ao excedente. A coleta bruta seria 20.261 x 23% + (25.000-20.261) x 37% = 6.535,075€. Originando uma taxa média final de 6.535,075 / 25.000 = 26,1403%.

A taxa média é a proporção do rendimento total que é paga em imposto, repartida pelos máximos de cada escalão. A taxa marginal é o acréscimo de imposto por cada € de rendimento extra no respetivo escalão. Um imposto é progressivo se as taxas marginais forem maiores que as taxas médias – *i.e.* se as taxas médias efetivas aumentarem com o nível de rendimento.

Como podemos ver na Figura 3, o IRS efetivamente pago também pode ser visualizado por uma curva de concentração que evidencia a progressividade do imposto, se estiver abaixo da Lorenz. O principal **indicador de progressividade**, construído por Kakwani (1977), capta o desvio do sistema de imposto em relação ao imposto proporcional (*flat tax*), onde a mesma taxa se aplica a toda gente, o que implica que as duas curvas são iguais. Assim, mede a diferença entre o índice de concentração do imposto e o Gini do rendimento bruto; ou o dobro da área entre a curva de Lorenz e a de concentração:

$$Kak = C_t - G_X = 2K$$
$$= 2\int_0^1 L(p_1) - Ctax(p_1)\, dp_1$$

Kakwani (1977) demonstra que o seu índice pode ser decomposto como a média ponderada da progressividade das diferentes componentes i do sistema fiscal (entre IRS, CSS e outros impostos, por exemplo):

$$Kak_{Total} = \sum_i \frac{t_i}{t} \times Kak_i$$

Onde t_i é a taxa média de cada componente e t é a taxa média global que é, por sua vez, determinada pela receita fiscal total (T) a dividir pelo rendimento bruto nacional (X). Outra das razões para que o *Kakwani* seja a medida de progressividade mais utilizada prende-se com a relação direta com o conceito de redistribuição vertical:

$$VE = \frac{t}{(1-t)} \times Kak$$
$$onde\ \frac{t}{(1-t)} = \frac{t.X}{X/_X - T/_X} = \frac{t.X}{X-T} = \frac{T}{N}$$
$$VE = \frac{T}{N} \times Kak$$

Desta feita, a **redistribuição é uma função da dimensão do imposto ($T/_N$) e do nível de progressividade.** Esta formulação realça a importante diferença entre o conceito de progressividade e o de redistribuição.

Isto quer dizer que **é possível que o nível de redistribuição aumente num ano em que a progressividade do imposto diminui, ou vice-versa.** Este é o caso onde, não havendo alterações

do número de escalões, haja uma atualização dos limites dos escalões abaixo da evolução dos rendimentos médios, provocando um aumento dos agregados que se encontram pouco acima do mínimo de existência maior do que nos escalões mais altos. A redistribuição aumenta porque aqueles que não pagam imposto estão relativamente mais iguais àqueles que pagam, mas a progressividade diminui porque o imposto passa a estar relativamente menos concentrado no topo da distribuição. Naturalmente, se a progressividade for negativa – *i.e.* concentração do imposto menor que a concentração do rendimento bruto – o efeito redistributivo do imposto também será negativo.

$$-1 \leq \textit{índices de progressividade} \leq 1$$

Outros indicadores de progressividade[13] compararam a distribuição inicial com a concentração do imposto de outras formas. Simultaneamente ao Kakwani, Suits (1977) desenvolveu outro índice de progressividade. A diferença é que *Suits* integra em relação ao rendimento e não sobre a população: eixo horizontal da figura 3 é a Lorenz do rendimento bruto $L(p_1)$ e não a função cumulativa da população $F(p_1)$. Note-se que a proporção cumulativa do imposto pago pela população p é a mesma independentemente de p ser expressa como proporções cumulativas do número de declarações (*Kakwani*) ou em proporções cumulativas do rendimento total (*Suits*).[14] Deste modo, *Suits* traduz-se, essencialmente, no índice de Kakwani normalizado pelo declive da curva de Lorenz em cada ponto. Este declive $L'(p_1)$ é menor que um para p_1 menores, igual a um num único percentil, e maior que um para percentis mais elevados. Atribui, por isso, pesos maiores a rendimentos mais elevados (Seaks and Smith, 1981):

$$Suits = 2\int_0^1 [L(p_1) - Ctax(p_1)]L'(p_1)\, dp_1$$

Portanto, para as mesmas curvas de Lorenz e concentração, *Suits* é maior ou igual a *Kakwani* desde que as curvas sejam convexas (*i.e.* haja desigualdade). Seaks and Smith (1981) demonstram, também, que se a distribuição de rendimento se alterar de tal forma que $L'(p_1)$ varia significativamente, estes indicadores poderão variar com sinais opostos.

Mais recentemente, Stroup (2005) desenvolveu um rácio das convexidades relativas da Lorenz e da curva de concentração do imposto. A diferença entre as convexidades é normalizada em termos da convexidade da curva de Lorenz. Isto é, a área entre a Lorenz e a curva de concentração do imposto (K) a dividir pela área abaixo da curva de Lorenz (B).

$$Stroup = \frac{\int_0^1 L(p_1) - Ctax(p_1)\, dp_1}{\int_0^1 L(p_1)\, dp_1} = \frac{K}{B}$$

$$= 1 - \frac{\int_0^1 Ctax(p_1)\, dp_1}{\int_0^1 L(p_1)\, dp_1} = 1 - \frac{(B-K)}{B}$$

Stroup (2013) colmata uma limitação importante dos outros dois índices de progressividade. Segundo o autor, e seguindo a lógica do Gini, **uma medida de progressividade deve atingir o seu valor máximo na situação em que apenas a unidade mais *rica* x_{max} paga imposto.** Uma vez que os outros índices de progressividade se baseiam na diferença entre as curvas de Lorenz e concentração,[15] é possível que tomem o mesmo valor em duas situações muito distintas:

1- Lorenz é praticamente linear (perfeita igualdade) e a $\mathrm{Ctax}(\mathrm{p}_1)$ é relativamente convexa, medindo uma área cinzenta $\mathrm{K} = \mathrm{a}$;

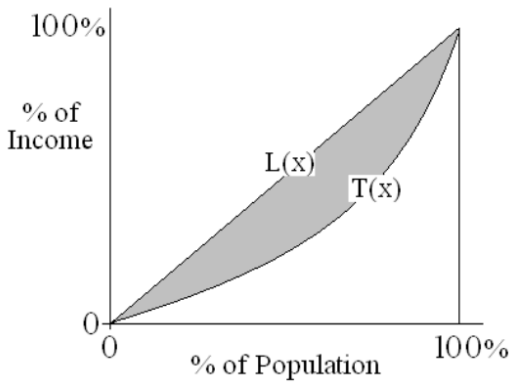

Figura 4 – Cenário hipotético 1 – curva de Lorenz do rendimento bruto e curva de concentração do imposto pouco concentradas.
Fonte: Stroup (2013)

[13] A única exceção é Khetan-Poddar (1976) que mede o resíduo do Gini original em relação ao resíduo da concentração do imposto $KP = (1 - G_X)/(1 - C_T)$, sendo o imposto progressivo se o indicador for maior que um.

[14] Por exemplo, a figura 3 mostra-nos que 60% do número de declarações mais baixas paga cerca de 5% do IRS total. Estes 60% do número de declarações detêm 19% do rendimento total. Logo a população que representa 60% das declarações mais baixas é a mesma população que representa os 19% de rendimentos mais baixos, que paga 5% do IRS.

[15] *Kakwani* na diferença (vertical) absoluta e *Suits* na diferença (vertical) multiplicada pela inclinação da Lorenz.

2- Lorenz é extremamente convexa (próxima do canto inferior direito) enquanto a $\mathrm{Ctax}(\mathrm{p}_1)$ é uma reta horizontal até ao último p_1 onde se torna uma reta vertical. Esta representa a situação de progressividade máxima, porém, de novo, $\mathrm{K} = \mathrm{a}$.

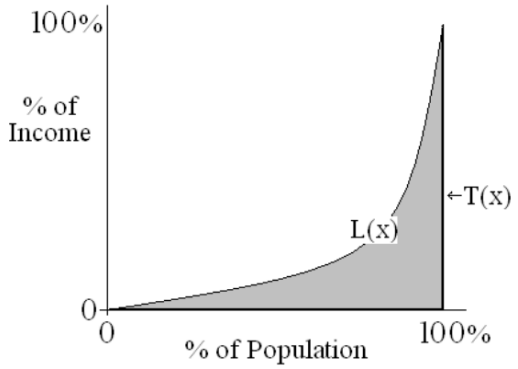

Figura 5 – Cenário hipotético 2 – curva de Lorenz do rendimento bruto muito desigual e curva de concentração do imposto o mais concentrada possível.
Fonte: Stroup (2013)

Este indicador resolve o problema ao dividir a área cinzenta K pela área B, fazendo com que a progressividade no segundo cenário seja aproximadamente um, visto que K e B são aproximadamente iguais. Em suma, o índice de *Stroup* faz com que, para a mesma concentração de imposto, a medida de progressividade não varie em função da desigualdade original (curva de Lorenz).

Sabendo que é praticamente impossível que a distribuição de rendimentos passe rapidamente da primeira para a segunda situação, esta limitação ganha relevância quando estudamos a evolução da progressividade ao longo de vários anos.

4. *Shares* ou Fatias

A desigualdade também pode ser medida através da **percentagem do rendimento total detida por diferentes grupos da população: *shares* ou fatias.** À semelhança das limitações explicadas anteriormente, podemos estar perante níveis diferentes de desigualdade para o mesmo valor do Gini. Isto porque o Gini é mais sensível a variações em torno da moda do que nos extremos da distribuição.

Assim, Piketty, Saez, Zucman e outros autores têm-se concentrado na evolução da fatia dos grupos mais abastados. Estes indicadores são estimados através da curva de Lorenz, função cumulativa da proporção do rendimento total, explicada anteriormente. Genericamente, a fatia dum *top* global:

$$Top\ (1-p)\% = [1-\Phi(p)]\ \%$$
$$= \left[\frac{rend.\,médio\ [q(p);q(1)]}{rend.\,médio\ total} \times (1-p)\right]\%$$

Por exemplo, para calcularmos a fatia do *top* 10%, $(1-p) = 0{,}1$, logo $p = 0{,}9$ é a proporção da população a partir da qual todas as unidades pertencem ao *top* 10%. Assim, será o rácio do rendimento médio daqueles com rendimentos entre o percentil 90 e o rendimento máximo a dividir pelo rendimento médio global, multiplicado pela dimensão do intervalo. Por sua vez, expresso em percentagem, é igual a cem porcento menos a proporção cumulativa do rendimento dos 90% mais *pobres* (a imagem do ponto β da figura 3): 100% – 64,15% = 35,85%.

Adicionalmente é possível calcular fatias de partes intermédias da distribuição. Por exemplo, a fatia da metade mais *pobre* do *top* 20% (*i.e.* 9º decil ou D9) é igual à diferença entre a proporção cumulativa (Lorenz) do percentil 90 e aquela do percentil 80 – imagem do ponto β menos a imagem do ponto α. Assim como é igual ao rendimento médio do 9º decil sobre a média de todos os rendimentos, multiplicado por 0,1 (e ainda por 100 se expresso em percentagem).

Neste contexto, a redistribuição é medida, em pontos percentuais, pela variação da fatia do *top* 10% antes e depois do imposto. Este valor é calculado pela diferença absoluta entre a imagem do ponto β (64,15%) e aquela do ponto ε (68,47%) – como evidenciam os segmentos de reta verticais na Figura 3 (=4,32p.p.).

5. Rácios

Rácios são os indicadores de desigualdades mais simples, fáceis de calcular e interpretar. A desigualdade é frequentemente dividida em disparidade da metade inferior (P50/P10 = mediana/D1) e disparidade da metade superior (P90/P50 = D9/mediana) – geralmente mais elevada. Outro rácio de percentis comum é o intervalo interquartil (P75/P25 = Q3/Q1). Os rácios de fatias mais usados são o S90/S10 e S80/S20.

Figura 6 – Reta real que demonstra a diferença entre decis D (valor monetário) e as fatias S (%) que cada decil (intervalo com 10% da população) detém.

É importante perceber a distinção entre rácios de percentis e rácios de fatias. Enquanto os primeiros dizem respeito a valores monetários de rendimento correspondentes a uma posição na distribuição, os segundos referem-se à proporção do rendimento total que um intervalo da população detém. Portanto, **o rácio de percentis P90/P10 será menor que o rácio de fatias S90/S10.** Por exemplo, o rácio P90/P50 diz-nos quantos rendimentos medianos é que a unidade mais *pobre* do *top* 10% recebe. Por sua vez, o rácio S90/S50 diz-nos quantos rendimentos médios da metade mais *pobre* (valor abaixo do P50) recebem, em média, aqueles que pertencem ao *top* 10% (acima do P90).

Imaginemos um cenário em que, de um ano para o outro, apenas os rendimentos do *top* 1% aumentam. Nesta situação, o rendimento da primeira unidade (mais *pobre*) do *top* 10% não se altera, mas o rendimento médio do *top* 10% aumenta. Logo, o rácio P90/P10 não varia – visto que a ordem e valores até, pelo menos, o 99º percentil são exatamente os mesmos – enquanto o rácio S90/S10 pode aumentar substancialmente.

Os rácios de percentis são indicadores fáceis de calcular e de interpretar, não são sensíveis a *outliers* nem à inflação, mas ignoram grande parte da distribuição utilizando apenas dois pontos.

5.1. Rácio de Palma

O rácio de Palma é igual à fatia do *top* 10% dividida pela fatia dos *bottom* 40% (S90/S40) e tende a captar a desigualdade impercetível ao Gini.

Num artigo seminal de 2011, Palma demonstra que a fatia recebida pelas "classes média e média-alta" (entre D5 e D9, inclusive) apresenta uma homogeneidade surpreendente entre países de todo o mundo e ao longo dos anos. Esta parte da distribuição tende a deter pouco mais de 50% do rendimento total, o que implica que a desigualdade resulta da repartição da outra metade do rendimento entre os 40% mais *pobres* (D1-D4) e o *top* 10% (D10).

Apesar de ser um indicador pouco usual é de grande e crescente importância. Veja-se, por exemplo, a redação da primeira meta do 10º Objetivo de Desenvolvimento Sustentável ("Reduzir as desigualdades"): *Até 2030, progressivamente alcançar, e manter de forma sustentável, o crescimento do rendimento dos 40% da população mais pobre a um ritmo maior do que o da média nacional.* Esta poderá ter sido influenciada pelos trabalhos de José Gabriel Palma.

5.2. Coeficiente de variação

O coeficiente de variação é dos indicadores de desigualdade mais antigos e amplamente utilizados. Este é simplesmente o rácio entre o desvio padrão e a média. Por um lado, duas distribuições podem ter a mesma média, mas desvios padrão muito diferentes. Por outro, é importante estandardizar a variância (quadrado do desvio padrão) pela média uma vez que a possibilidade de disparidade aumenta com a média.

Este indicador é fácil de calcular e estimar, não é afetado pela inflação e incorpora todas as observações. Para além disso, opostamente ao Gini, é mais sensível a transferências entre os extremos.

6. Outros indicadores

Como vimos, todos os indicadores acima são compreendidos através do mesmo enquadramento gráfico das curvas de Lorenz e de concentração – espaço da Figura 3. Em baixo são apresentados e analisados indicadores de desigualdade com outro enquadramento.

6.1. Índice de Atkinson

Num artigo seminal, o economista britânico Anthony Atkinson (1970) prova a intuição expressa por Dalton 50 anos antes: subjacente a qualquer medida sumária de desigualdade está um conceito/julgamento de bem-estar social e é sobre esse com que nos devemos preocupar**. Qualquer medida de desigualdade deve diminuir após uma transferência duma pessoa mais rica para uma pessoa mais pobre**.[16] Acontece que o Gini, por

[16] Assumindo que a transferência não origina uma reordenação substancial – condição Pigou-Dalton. Tanto o desvio relativo médio como os rácios de percentis (*e.g.* intervalo interquartil) não verificam esta propriedade. As outras duas condições para

exemplo, atribui mais importância a transferências que ocorrem perto do centro de distribuição (a moda) do que nos extremos.

Alicerçado nas ideias de Dalton, Atkinson propõe o conceito de **rendimento equivalente igualmente distribuído –** $Yede$ **–** nível de rendimento *per capita* que, se fosse atribuído igualmente por todos, não alteraria o bem-estar social atual (dada uma certa aversão à desigualdade). Segundo a função de bem-estar social proposta pelo autor, se o rendimento de uma pessoa aumentar 1%, o seu peso para o bem-estar social diminuirá ε% relativamente ao seu valor inicial. Assim, ε é o grau de aversão à desigualdade duma sociedade. Este parâmetro varia entre zero (indiferença total) e infinito. Valores mais altos representam maior propensão, por parte das pessoas, para aceitar rendimentos menores em nome de maior igualdade. Isto porque quanto maior for ε mais importância será dada pelo índice à desigualdade na cauda da distribuição (entre os mais pobres). Os valores convencionais para ε são: 0,5, 1, 1,5 ou 2.[17]

$$At(\varepsilon, X) = 1 - \frac{Yede}{\mu}$$
$$= 1 - \left[\sum_i \left(\frac{x_i}{\mu}\right)^{1-\varepsilon} \times f(x_i)\right]^{1/1-\varepsilon} \quad \varepsilon \neq 1$$

Para cada grau de aversão à desigualdade, o índice de Atkinson $At(\varepsilon, X)$ representa a proporção do rendimento total que seria necessária para manter o mesmo nível de utilidade, se o rendimento fosse igualmente distribuído. Posto de outra forma, se o rendimento fosse igualmente distribuído, o mesmo bem-estar social poderia ser alcançado com apenas $1 - At(\varepsilon, X)$ % do atual rendimento nacional.

Sabendo das limitações do Gini, Kiefer (1984) propõe um novo indicador de redistribuição definido como a diferença entre os índices de Atkinson antes e depois do imposto.

6.2. Índices de Entropia Generalizada

Inspirado no conceito de entropia de sistemas termodinâmicos, Claude Shannon (1948) desenvolveu uma teoria para medir o valor da informação, onde acontecimentos menos prováveis acarretam mais informação. Por sua vez, o nível de informação esperado é estabelecido pela entropia (Cowell, 2002).

$$Entropia = \sum_{i=1}^{N} prob_i \times \ln\left(\frac{1}{prob_i}\right)$$
$$\rightarrow Ig(fatia) = \sum_{i=1}^{N} fatia_i \times \ln\left(\frac{1}{fatia_i}\right)$$

Baseado nesta teoria, Henri Theil (1967) desenvolveu uma metodologia para medir igualdade usando as fatias do rendimento de cada grupo em lugar das probabilidades dos acontecimentos. Note-se que a perfeita igualdade ocorre quando todas as unidades têm as mesma $fatia_i = 1/N$ – onde N é o número de unidades. Nessa situação a fórmula da igualdade torna-se $\ln N$. A desigualdade é, então, medida pela diferença entre a igualdade perfeita e a igualdade verificada: $\ln N - Ig(fatia)$. Assim, o índice de desigualdade varia entre o mínimo de zero e o máximo de $\ln N$. Este índice de Theil T pode ser reescrito em função do número de unidades e a relação entre o rendimento individual e o rendimento médio:

$$Theil\ T = \sum_{i=1}^{N} \frac{1}{N} \times \frac{x_i}{\mu} \times \ln\left(\frac{x_i}{\mu}\right)$$

Essencialmente, é a média ponderada (pela população N) dos desvios proporcionais do rendimento (x_i) relativamente ao rendimento médio (μ). O último termo determina se o contributo dessa unidade i é positivo ou negativo.[18]

James K. Galbraith (*e.g.* 2008), e o projeto UTIP, têm vindo a defender o uso deste indicador em vez do Gini. Ao contrário deste último, que geralmente requer micro-dados, é possível estimá-lo apenas com dados agregados. O índice de Theil T tem a vantagem de ser aditivamente decomponível,[19] distinguindo a contribuição da desigualdade *entre* grupos daquela

uma medida de desigualdade são: não variarem com a escala ou média, e não variarem com a dimensão da população (Heshmati, 2004).

[17] Se $\varepsilon = 1$ então $A(\varepsilon, X) = 1 - \frac{\prod x_i^{(1/N)}}{\mu}$.

[18] Note-se que ln(x)>0 se x>1 e ln(x)<0 se x<1.

[19] Desigualdade total é a soma ponderada das desigualdades de cada subgrupo mais a contribuição que advém das desigualdades entre as médias desses subgrupos.

que advém de diferenças *dentro* dos grupos: $Theil\ total = Theil_{within} + Theil_{between}$. Por sua vez, a componente total *entre* grupos é igual à soma dos *Theil* dos grupos que a compõem – fórmula acima de $Theil\ T$. Conceição et al. (2001) demostram esta propriedade fractal do índice e a consistência do $Theil_{between}$ em espelhar a evolução da desigualdade global apesar de o $Theil_{within}$ não ser observável. Isto quer dizer que é possível estimar a evolução da desigualdade dum país, e as contribuições de cada grupo, tendo apenas dados agregados do número de observações e valor médio – *e.g.* sectores, profissões, municípios ou distritos etc.[20]

Henri Theil desenvolveu ainda outro índice conhecido por *Theil L*, invertendo o papel das fatias:

$$Theil\ L = \sum_{i=1}^{N} \frac{1}{N} \times \ln(\frac{\mu}{x_i})$$

Ambos medem a divergência entre diferentes fatias de rendimento e proporções da população, mas com funções de distância distintas. O *Theil T* dá mais peso a dispersões acima da média enquanto o *Theil L* é mais sensível a variações entre rendimentos mais baixos.

Mais tarde, Cowell (1980) demonstra que os dois índices de Theil são casos particulares do índice generalizado de entropia:

$$GE(\alpha) = \frac{1}{\propto(\propto -1)}\left[\frac{1}{N}\sum_{i=1}^{N} \times \left(\frac{x_i}{\mu}\right)^{\propto} - 1\right]$$

Onde α representa o peso atribuído a distâncias entre rendimentos em diferentes partes da distribuição: valores maiores dão mais importância a desigualdades no topo e vice-versa. É este parâmetro que define vários índices de entropia:

- GE(-1)
- GE(0) = Desvio Médio Logarítmico (*Theil L*)
- GE(1) = índice de Theil (*Theil T*)
- GE(2) = Metade do quadrado do coeficiente de variação

O índice de *Theil L* é mais conhecido por **desvio médio logarítmico**, atribui mais peso a desigualdades entre unidades com rendimentos mais baixos e também permite a desagregação por grupos (*e.g.* socioeconómicos).

Shorrocks (1980) prova a relação direta entre estes índices e os índices de Atkinson, onde $\propto = 1 - \varepsilon$. Todos estes índices são aditivamente decomponíveis ao contrário dos índices de Atkinson que são decomponíveis, mas não aditivamente: $A(\varepsilon) = A(\varepsilon)_{within} + A(\varepsilon)_{between} - A(\varepsilon)_{within} \times A(\varepsilon)_{between}$.

Esta é a grande vantagem dos índices de entropia. As desvantagens são a sua difícil interpretação (Sen, 1973: 36) e não possibilitar a comparação direta entre populações com dimensões ou estrutura de grupos diferentes.

Na prática, podemos encontrar no *website* do projeto UTIP manuais com exemplos e ficheiros Excel para o cálculo do *Theil T*. Para além disso existem vários comandos de *Stata*[21] que calculam todos estes indicadores automaticamente (ver Quadro 3).

7. Conclusão

Genericamente, a desigualdade mede a dispersão entre a percentagem da população e a proporção dos recursos que essa população detém. As desigualdades sociais são multidimensionais, abrangendo igualdade de oportunidades, educação, saúde, trabalho, etc., não se restringindo a variáveis monetárias.

Este trabalho procurou apresentar e explicar cada um dos indicadores de desigualdades desenvolvidos pela literatura académica. Com enfase no rendimento e no Imposto sobre o Rendimento das pessoas Singulares (IRS), pretende ser um manual que contribuía para a capacitação da administração pública em análises de distribuição (desigualdade, redistribuição e progressividade).

Existem vários indicadores de desigualdades, todos com vantagens e desvantagens. A medida mais utilizada é o índice de Gini, mas este é pouco sensível

[20] O $Theil_{between}$ é o mínimo do $Theil\ T$ e $Theil_{população} > Theil_{concelho} > Theil_{distrito} > Theil\ T$. Note-se que o máximo destes indicadores varia com o número de grupos. Logo, se dois países tiverem quantidades de grupos diferentes ou se estes variarem ao longo do tempo, este indicador não é apropriado. O mesmo se aplica se apenas tivermos um ano na amostra.

[21] *Stata* é um *software* de análise estatística e de tratamento de dados, com uma *syntax* simplificada, que contém inúmeros programas desenvolvidos para análises econométricas e gráficas, amplamente utilizado em economia, sociologia, biomedicina e ciência política.

a diferenças que ocorrem nos extremos da distribuição e não é aditivamente decomponível. A grande maioria pode ser compreendida e visualizada através do enquadramento analítico e gráfico das curvas de Lorenz e de concentração (Figura 3). No entanto, existem indicadores com outro enquadramento que permitem examinar efeitos em partes diferentes da distribuição e que são aditivamente decomponíveis em grupos.

No anexo são apresentadas as bases de dados internacionais que contêm diversas medidas de desigualdades de rendimento. A *Luxembourg Income Study* (LIS) e as *Statistics on Income and Living Conditions* (EU-SILC) disponibilizam os microdados a entidades de investigação. Deve-se realçar o facto de Portugal ser o único país da área do euro, para além de Malta e Chipre, que não consta na LIS. Esta é a maior base de dados harmonizados sobre o rendimento, abrangendo 5 décadas e 50 países de todo o mundo. Por isso, está na base de muitas outras bases de dados, designadamente a *Leiden LIS Budget Incidence Fiscal Redistribution Dataset* que contém muitas variáveis, nomeadamente sobre os efeitos fiscais, não disponíveis em qualquer outra fonte.[22]

Tendo o objetivo de ser um manual prático, apresentam-se ainda todos os pacotes de *Stata* que permitem o cálculo de todos estes indicadores (Quadro 3, anexo). Sabendo que o *Stata* não é gratuito, são também apresentados os pacotes do *software* livre *R*. Este é mais flexível e tem maiores potencialidades gráficas, de tratamento de dados, e de computação que o *Stata*. Apesar do *R* apresentar um número reduzido de programas predefinidos para estudos de distribuição, o conteúdo metodológico do presente artigo possibilita a construção manual dos indicadores.

A criação, monitorização e avaliação de políticas eficazes de combate às desigualdades sociais requerem a adequada estimação e análise de indicadores de desigualdades de rendimento e, em parte, dos efeitos redistributivos do sistema fiscal. Por sua vez, uma análise aprofundada das desigualdades fiscais e de rendimento exige o conhecimento sobre as vantagens e complementaridades dos diversos indicadores existentes, de modo a identificar a combinação mais apropriada para cada problemática.

"[…] in different stages of society, the proportions of the whole produce of the earth which will be allotted to each of these classes [rentiers, capitalists and labour], under the names of rent, profit, and wages, will be essentially different; To determine the laws which regulate this distribution, is the principal problem in Political Economy."

David Ricardo, 1817

Agradecimentos

O autor agradece ao GPEARI a oportunidade de trabalhar no GT-IDS do Conselho Superior de Estatística e a disponibilização dos microdados, devidamente anonimizados, por parte da Autoridade Tributária e Aduaneira.
Um enorme agradecimento ao Pedro Duarte Silva, diretor do Departamento de Análise, Estudos e Previsão do GPEARI, pelas várias orientações, críticas e sugestões feitas ao longo destes meses.
Todos os erros são da inteira responsabilidade do autor.

[22] https://www.lisdatacenter.org/resources/other-databases/

8. Bibliografia

Alvaredo, F., Atkinson, A. B., Piketty, T., & Saez, E. (2013). The top 1 percent in international and historical perspective. Journal of Economic perspectives, 27(3), 3-20.

Atkinson, A. (1980), Horizontal Inequity and the Distribution of the Tax Burden in H. Aaron & M. Boskin (eds.), The Economics of Taxation, pp.3-18.

Atkinson, A. B. (1970). On the measurement of inequality. Journal of economic theory, 2(3), 244-263.

B. Milanovic, (2002). "True World Income Distribution, 1988 and 1993: First Calculations Based on Household Surveys Alone". The Economic Journal. vol. 112, no. 476. pp. 51-92. 2002.

Rodrigues, C. F. (1993). Measurement and Decomposition of inequality in Portugal, 1980/81-1989/90. University of Cambridge, Department of Applied Economics.

Rodrigues, C. F., Figueiras, R. e Junqueira, V. (2016). Desigualdade do Rendimento e Pobreza em Portugal: As Consequências Sociais do Programa de Ajustamento", Fundação Francisco Manuel dos Santos, Lisboa. ISBN: 9789898863461.

Rodrigues, C. F., e Andrade, I. (2019). Redistribution Policy and Inequality in Portugal. Apresentação na Escola de Inverno do 2.º Encontro anual da Associação Portuguesa de Economia Política (Coimbra).

Conceição, P., Galbraith, J. K., & Bradford, P. (2001). The Theil index in sequences of nested and hierarchic grouping structures: Implications for the measurement of inequality through time, with data aggregated at different levels of industrial classification. Eastern Economic Journal, 27(4), 491-514.

Cowell, F. A. (1980). Generalized entropy and the measurement of distributional change. European Economic Review, 13(1), 147-159.

Cowell F. A. (2000). "Measurement of Inequality", in A B Atkinson / F Bourguignon (Eds): Handbook of Income Distribution, Amsterdam.

Cowell, F. A. (2003). Theil, inequality and the structure of income distribution. LSE STICERD Research Paper, (67).

Dalton, H. (1920). The measurement of the inequality of incomes. The Economic Journal, 30(119), 348-361.

Decoster, A., & Van Camp, G. (1998). The unit of analysis in microsimulation models for personal income taxes: fiscal unit or household?. CES-Discussion paper series DPS 98.33, 1-35.

Fellman, J. (2018). Income Inequality Measures. Theoretical Economics Letters , 8, 557-574.

Formby, J. P., Seaks, T. G., & Smith, W. J. (1981). V/I LEY. The Economic Journal, 91(364), 1015-1019.

Galbraith, J. K. (2008). Inequality, unemployment and growth - New measures for old controversies. The Journal of Economic Inequality, 7(2), 189–206.

Gini, C. (1912). Variabilità e mutabilità. Reprinted in Memorie di metodologica statistica (Ed. Pizetti E, Salvemini, T). Rome: Libreria Eredi Virgilio Veschi.

Heshmati, A. (2004). Inequalities and their measurement.

Kakwani, N.C. (1977). Measurement of Tax Progressivity: An International Comparison, Economic Journal 87: 71-80.

Kiefer, D. W. (1984). Distributional tax progressivity indexes. National Tax Journal (pre-1986), 37(4), 497.

Kuznets, S. (1955). Economic growth and income inequality. The American economic review, 45(1), 1-28.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. Publications of the American statistical association, 9(70), 209-219.

Musgrave, R.A. and Thin, T. (1948). Income tax progression 1929-48, Journal of Political Economy 56: 498-514.

OECD, Nicolas Ruiz and Nicolas Woloszko (2016). What do households surveys suggest about the top 1% incomes and inequality in OECD countries?. Economics Department Working Papers No. 1265

Palma, J. G. (2011). Homogeneous middles vs. heterogeneous tails, and the end of the 'inverted-U': It's all about the share of the rich. development and Change, 42(1), 87-153.

Pechman, J. A., & Okner, B. A. (1974). Who Bears the Tax Burden? Studies of Government Finance Series.

Piketty, T., Saez, E., & Zucman, G. (2017). Distributional national accounts: methods and estimates for the United States. The Quarterly Journal of Economics, 133(2), 553-609.

Reynolds, M. and Smolensky, E. (1977). Public Expenditures, Taxes, and the Distribution of Income: The United States, 1950, 1961, 1970, Academic Press, New York.

Sen, A. (1997). On Economic Inequality. Oxford: Clarendon Press

Shannon, C. E. (1948). A mathematical theory of communication. Bell system technical journal, 27(3), 379-423.

Shorrocks, A. F. (1980). The class of additively decomposable inequality measures. Econometrica: Journal of the Econometric Society, 613-625.

Solt, Frederick. (2019). Measuring Income Inequality Across Countries and Over Time: The Standardized World Income Inequality Database." SWIID Version 8.1, May 2019.

Stroup, M. D. (2005). An index for measuring tax progressivity. Economics Letters, 86(2), 205-213.

Stroup, M. D., & Hubbard, K. E. (2013). An improved index and estimation method for assessing tax progressivity (No. 13-14). Mercatus Center Working Paper.

Suits, Daniel B. 1977. "Measurement of Tax Progressivity," The American Economic Review, 67(4):747-752.

Theil, H. (1967). Economics and information theory (No. 04; HB74. M3, T4.).

Anexos

Quadro 3 – Síntese de comandos básicos e principais *packages* desenhados para análises de distribuição, dos *softwares* estatísticos *Stata* e *R*.

Quadro 3.I				**Stata**				
instalar		ssc install *nome*		Ctrl+D para executar linha				
detalhes		help *nome*						
básicos	count if *var*	sum *var*, de	drop if *var* ==0	gen *newvar* = expressão		replace *vartochange*=expr.		
nome	Gini	Percentis	Fatias	Concentração	Redistrib.	Progressi.	Atkinson (Yede)	Entropia Gen.
ineqdeco	✔	☐	✘	✘	✔	✘	✔✔	✔✔
pshare	✔✔	✔	✔✔	✔	☐	☐	✘	✘
sumdist	✘	✔✔	✔	✘	☐	✘	✘	✘
sumup	☐	✔	☐	✘	☐	✘	✘	✘
progres	✔	✘	✘	✔✔	✔✔	✔✔	✘	✘
dasp	É um programa extremamente desenvolvido que permite calcular todos estes indicadores e muito mais							
✔✔	melhor comando para este tipo de indicador							
✔	comando/programa calcula automaticamente este tipo de indicador							
✘	comando/programa não permite calcular este tipo de indicador							
☐	comando/programa oferece informação suficiente para calcular este tipo de indicador							
				Manual				
Percentis				egen perc_99 = pctile(rend_bruto), p(99)				
Fatias				gen d_top1 = 0 replace d_top1 = 1 if rend_bruto >= perc_99				
	collapse (sum) rend_bruto (mean) mean_bru=rend_bruto (max) percentil=rend_bruto, by(d_top1)							

Quadro 3.II			**R**	
instalar	install.package("*nome*")		library(*nome*)	Ctrl+Enter para executar linha
detalhes	?*nome*			
básicos	head(*basedados*)	summary(*basedados* $rend_bruto)		estudar *tidyverse*
nome				
ineq	Vários indices de concentração; Gini, Reynolds-Smolensky; Atkinsons; Entropia Generalizada; Coeficiente de variação. Curvas de Lorenz e concentração.			
IC2	Ginis; concentração, curvas de Lorenz e concentração, Decomposição de Atkinsons e Entropia Gen.			

Tabela 1A – Bases de dados internacionais de indicadores de desigualdades de rendimento, riqueza e fiscalidade.

Database	Variables / case for Portugal	Country Coverage	Period	Reference
EU-SILC Statistics On Income And Living Conditions Eurostat public version but also provides micro-data to researchers.	Quantile's cutoffs and shares: Quartiles, Quintiles, Deciles, Percentiles; Several mean and median equivalised net incomes decomposed by: sex, age, household types, activity status, source, education, tenure; Mean and median incomes before social transfers (including or not pensions) by: household types, age, sex, degree of urbanisation; Share of people with income above or equal to a threshold by: urbanisation, education, age and sex; S80/S20 by sex, age and NUTS 2 regions; Gini for equivalised disposable income, gini for eq. disp. inc. before social transfers (including or not pensions). That is, market income gini (after personal income tax and social security contributions) including or not pensions;	28 EU Member States and some other european countries	2004-2018	https://ec.europa.eu/eurostat/web/microdata/european-union-statistics-on-income-and-living-conditions
LIS - Luxembourg Income Study	Deciles All household characteristics; Labor, capital and transfers incomes; Disposable, factor and pension incomes; Social Security Redistribution, Private redistribution. Ginis and Atkinson's coefficients; Labor intensity (types of contract, industry, hours worked etc.); Consumptions, Assets/Liabilities transactions. **Portugal missing**	50 countries	I-X Waves from 1980-2016	http://www.lisdatacenter.org/our-data/lis-database/
LWS - Luxembourg Wealth Study	Deciles All hh characteristics Total income, consumption, dispo_inc Labor intensity (types of contract, industry, hours worked etc.) Assets/Liabilities stock and Total net worth and disposable net worth, decomposed (inheritance, equity/credit) Savings (very decomposed, from whom, for what etc.) **Portugal missing**	14 countries	IV-IX Waves from 1995-2013	http://www.lisdatacenter.org/our-data/lws-database/
Leiden Budget Incidence Fiscal Redistribution Dataset 2017 (LIS based)	Median and Mean equivalized income All the following for Total pop. and Working pop.: Ginis for Primary/Original income, Gross income (after transfers) and Disposable income (after "" and taxes) Absolute and Relative fiscal redistribution , Share of fiscal redistribution (all from transfers and from taxes) Budget size (% of hh gross inc form each): transfers and taxes + their Efficiency/Targeting (*) **Portugal Missing**	47 countries	Varies alot, according to LIS availability	http://www.lisdatacenter.org/wp-content/uploads/files/resources-other-llbifr-doc.pdf http://www.lisdatacenter.org/wps/liswps/724.pdf

Tabela 1B – Bases de dados internacionais de indicadores de desigualdades de rendimento, riqueza e fiscalidade.

Incomes Across the Distribution Database Stefan Thewissen, Brian Nolan, Max Roser (2016) (LIS based)	Deciles means and cut-offs for real (PPP2011) and nominal of equivalised disposable income Ginis, absolute and relative redistribution, decile ratios (90/10, 90/50, 50/10). **Portugal missing**	50 countries	Varies alot, according to LIS availability	http://www.lisdatacenter.org/wp-content/uploads/resources-other-idd.pdf
OECD Income Distribution Database (IDD) mostly based on EU-SILC	Total population, elderly, working age; Mean (for various cohorts; capital, self-emp) and median disp. inc., transfers paid (taxes and SS)/received; gini disp inc, gini market, gini's standard errors; Palma ratio (S90/S40), P90/P50, P90/P10, P50/P10, S80/S20. (**)	OECD countries	PT effective: 2004-2015 All PT: 1979-80, 1990, 1995, 2000, 2004 (Household Budget Survey - INE), 2005, 2006, 2007, 2008 and 2009 (SILC)	http://www.oecd.org/social/income-distribution-database.htm
Comparative Welfare States Data Set, 2014 (David Brady, Evelyne Huber, and John D. Stephens)	Myriad of variables, based on other major datasets (OECD IDD, SWIID, LIS, WID) on: Wage, Salary, and Income Distribution Data (top 1%, 0.1%, ratios etc.); Social Spending, Revenue, and Welfare State Institutions Data Social Policy, Education, Labor Force and Labor Institutions Data Demographics and Macro **Portugal missing**	27 rich countries	1960-2014 mostly post-2000	Codebook: http://www.lisdatacenter.org/wp-content/uploads/CWS-codebook.pdf http://www.lisdatacenter.org/wp-content/uploads/CWS-data.xlsx
The Chartbook of Economic Inequality Atkinson, Hasell, Morelli, and Roser (2017)	**Portugal** (based on Farinha 1996 and Pereirinha 1988, top shares Alvaredo 2008): No wealth data Top income shares since 1936 Gini since 1967 Poverty and dispersion of earnings 1980/2 (Atkinson 2008)	25 countries	top shares: 1940-2014 ginis and poverty: 1980-2014	https://chartbookofeconomicinequality.com/wp-content/uploads/Chartbook_Of_Economic_Inequality_complete.pdf
AllGinis (World Bank) Branko Milanovic	Compiles only Ginis from 9 survey sources (LIS, SILC, WYD, OIVCAL, WIDER etc.) **Portugal:** 1973-2008 but only 11 observations - 1973/80/90/91 WIDER, 1989/94/97 WYD, 2005-8 SILC	166 countries	1950-2015	http://pubdocs.worldbank.org/en/728601472744842249/Description-of-AllGinis-Oct2014.pdf

Tabela 1C – Bases de dados internacionais de indicadores de desigualdades de rendimento, riqueza e fiscalidade.

World Inequality Database (WID)	Huge project aimed at producing Distributional National Accounts (DINA) started by calculating top income shares, than income, than wealth. (no ginis) Combines Macro, Fiscal and Survey data. **Portugal:** No wealth data yet Pre-tax national income - Top 1% and 1'% inc share 1976-2005 but Top 0.1% 1938-2005 Fiscal and pre-tax income for all and various percentiles (avg and shares 1938-2005, cutoff and beta 1989-2005) All from Alvaredo 2008/09	54 countries	Varies alot	http://wid.world/wid-world/
SWIID The Standardized World Income Inequality Database	Uses multiple-imputation to fill in missing observations, based on a myriad of datasets, aiming at the LIS values (standard), to became the largest income inequality database. Gini net, gini market, absolute and relative redistribution. **Portugal:** able to calculate the **ginis for 1968-2015** and the **redistributions for 1975-2015**	192 countries	1960-2015	https://fsolt.org/swiid/
UNU-WIDER (WIID) The World Income Inequality Database	Ginis, quintiles, deciles and 5/95 ventiles, mean/median incomes from several sources with survey Quality indicators, tax units, household level, expenditure/consumption/income, etc. => duplicates **Portugal: 1973, 80, 90, 95-99, 2000, 01, 2004-2005**	182 countries	pre-1960 to 2015	https://www.wider.unu.edu/database/world-income-inequality-database-wiid34
Lakner-Milanovic	Panel of country-deciles cutoffs and means for comparision. **Portugal** source is World Income Distribution (WYD) database: 1989, 94, 98, 2005, 2008	160 countries	5-in-5 years 1988-2008	https://www.gc.cuny.edu/CUNY_GC/media/LISCenter/brankoData/LaknerMilanovic2013WorldPanelIncomeDistributionLMWPIDDescription.pdf

(*) "applying Kakwani's (1986) 'index of concentration' to transfers. This index takes on the value of -1.0 if the poorest person gets all transfer income, 0 if everybody gets an equal amount, and +1.0 if the richest person gets all transfer income (cf. Korpi & Palme, 1998: 684)"