

# Coeficientes de Gini para distribuições com renda negativa, uma interpretação geométrica

Seme Gebara Neto UFMG  
Luiz Fernando Saraiva UFF  
João Felipe Cury Mathias UFRJ

## Resumo

O coeficiente de Gini para distribuições com rendas negativas exige uma normalização. Chen et al. (1982) propuseram um ajuste, mas a desigualdade máxima ainda não estava bem definida. Raffinetti et al. (2014) propuseram outra normalização, tornando o coeficiente bem definido. Na análise de um exemplo com dados da nossa pesquisa, o coeficiente de Raffinetti não respeitou Pigou-Dalton, enquanto o coeficiente de Chen sim. No estudo deste caso foi possível definir uma normalização que ao mesmo tempo está bem definida e respeita o princípio da transferência.

Palavras-chave: renda negativa, Gini normalizado, Pigou-Dalton.

Submetido à área 3: HISTÓRIA ECONÔMICA, DO PENSAMENTO ECONÔMICO E DEMOGRAFIA HISTÓRICA

## 1 Introdução

O coeficiente de Gini  $G$  é um reconhecido indicador de desigualdade de renda, quando as famílias têm, todas, renda não negativa. Usaremos as palavras “família” e “renda”, em vez de unidade e valor, mas a discussão é geral, para qualquer distribuição de valores.

Ao se tratar com sociedades escravocratas, considerando as pessoas escravizadas como unidades familiares, um modelo possível é imputar a elas renda negativa, o que constitui parte do nosso conjunto de dados. Este modelo representa a renda gerada por pessoas escravizadas sendo expropriada delas. Para distribuições com rendas negativas, a interpretação de  $G$  fica comprometida, pois a desigualdade máxima corresponde a uma distribuição para a qual uma família percebe toda a renda e todas as outras nada, e pode ocorrer  $G > 1$ .

Para lidar com distribuições que admitem rendas negativas, Chen, Tsaur e Rhai (1982) [CTR] ajustam a normalização do coeficiente de Gini propondo outra situação de desigualdade máxima. Dada uma distribuição  $(Y_1, \dots, Y_M)$ , com valores ordenados de maneira não decrescente e com a soma de todas as rendas ainda positiva, os autores consideram a situação em que os primeiras  $M - 1$  famílias tenham rendas negativas e positivas que somem zero e a última tenha renda igual à soma total. Normalizando desta maneira definem o coeficiente  $G^* \in [0, 1]$  e que coincide com  $G$  na ausência de rendas negativas, tornando possível análises que ainda não havia.

Raffinetti, Siletti e Vernizzi (2014) [RSV] encontram problemas na definição de  $G^*$ , mostrando três distribuições significativamente diferentes com mesmos número de famílias e soma total de rendas, todas com  $G^*$  muito próximo de 1. Ou seja, a situação de desigualdade máxima proposta por [CTR] não determina a curva de Lorenz (isto é, pode haver mais de uma distribuição que represente aquela situação). Raffinetti et al. propõem considerar, como situação extrema, a distribuição na qual uma família recebe a soma das rendas negativas, uma família recebe a soma das rendas positivas e as outras  $M - 2$  famílias têm renda zero, e normalizam mais uma vez o coeficiente de Gini, definindo  $G^P \in [0, 1]$ , que coincide com  $G$  na ausência de rendas negativas e que diferencia as distribuições apresentadas não distinguidas por  $G^*$ .

A notação  $G^P$  é devida a De Battisti, Porro e Vernizzi (2019) [BPV] e faz referência a “polarizado” (pois na desigualdade máxima há os polos positivo e negativo). De Battisti et al. adicionam novos exemplos aos presentes em [RSV], indicando cinco distribuições distintas para as quais o valor de  $G^*$  atinge seu máximo. Apresentam uma discussão sobre se esses índices respeitam o princípio de Pigou-Dalton. Eles definem a “compensação mínima” de uma distribuição com rendas negativas como sendo a redistribuição que transforma em zero todas as rendas até que a soma acumulada das rendas seja zero, preservando as restantes rendas positivas. Ou seja, a compensação é feita com as menores rendas positivas possíveis. Provam que no caso da compensação mínima  $G^*$  satisfaz o princípio da transferência (Pigou-Dalton), mas isso não é automático para  $G^P$ : existe uma condição que depende da razão entre os coeficientes de Gini (mesmo eventualmente maiores que 1) da distribuição original e da redistribuição, e também das somas totais das rendas positivas e negativas. Os autores apresentam, em um exemplo com dados reais, o cálculo de  $G$ ,  $G^*$  e  $G^P$ , assim como uma análise desses números.

Recentemente adicionamos à classe de exemplos existentes o cálculo de  $G^P$  para a distribuição de renda em Juiz de Fora na década de 1870, considerando escravizados como

peçoas, como famílias com renda negativa, a partir da metodologia de atribuição de renda aos escravizados discutida em Saraiva et al. (2022) [SMGG]. Conforme sugerido em [RSV] entendemos que o valor  $G^* = 0,94$  calculado para a distribuição analisada, embora não surpreendente, demandava, para uma análise mais completa, também o cálculo de  $G^P$ . O resultado  $G^* = 0,94$  não é surpreendente pois a curva de Lorenz cruza o eixo  $x$  próximo de  $0,93$  e isto significa que a soma das rendas dos 93% mais pobres ainda é zero. E a maximização de Chen et al. é justamente obtida quando a curva de Lorenz da redistribuição considerada cruza o eixo  $x$  mais à direita possível, entre  $x = 1$  e  $x = 1 - 1/M$ . Obtivemos, para nosso conjunto de dados,  $G^P = 0,76$ , mostrando que a inclusão de pessoas escravizadas na conta da renda resulta numa desigualdade ainda alta, mas não tão próxima da situação extrema como  $G^*$  sugeria.

Ao considerar distribuições com os escravizados na conta da renda, há outros cenários de interesse: a imputação de renda positiva, estimando como seria a sociedade se esses trabalhadores fossem livres e remunerados, e a imputação de renda zero, estimando como seria a sociedade se fossem livres, mas sem nenhum poder de compra. Em tais cenários, o coeficiente a ser calculado é  $G$ , pois neles não há rendas negativas. No cenário de imputação de renda zero, obtivemos  $G = 0.77$ , muito próximo de  $G^P = 0.76$ . Ambos estes cenários, com a imputação de rendas zero e negativa, são redistribuições de uma mesma renda total com o mesmo número famílias, da primeira para segunda diminuindo a renda dos escravizados de zero para negativa e transferindo esses valores como renda positiva à renda dos trabalhadores livres (proporcionalmente à sua própria renda), ou seja, movimentando renda dos mais pobres para os mais ricos, sendo portanto o segundo cenário intuitivamente mais desigual. Embora tais distribuições não sejam comparáveis no sentido de [RSV], ambos os índices  $G$  e  $G^P$  são tentativas de se medir desigualdade de renda, assim, a proximidade obtida não era esperada. Raffinetti et al. identificam problemas na definição de Chen et al., mas enfatizam a intenção inicial explicitada por [CTR] de “tentar reformular e normalizar o coeficiente de Gini para que se possa comparar distribuições sem rendas negativas e distribuições com algumas rendas negativas” (literalmente: *attempt to reformulate and normalize the Gini coefficient so that comparability can be attained between the distributions without negative incomes and the distributions with some negative incomes*). Assim os valores obtidos com nossos dados nos deixaram curiosos em relação às condições para que estes índices satisfaçam o princípio de Pigou-Dalton.

O objetivo deste trabalho é aprofundar a análise de distribuições com a presença de renda negativa, a partir das ideias geométricas presentes nas definições e propriedades dos coeficientes de Chen et al. e Raffinetti et al., propondo uma interpretação dos nossos dados que responda satisfatoriamente as questões apontadas. A seção 2 apresenta as definições, comentários e interpretação geométrica do coeficiente de Gini e suas normalizações. A seção 3 detalha a construção do cenário de renda negativa como uma transferência regressiva a partir do cenário de renda zero e introduz um coeficiente normalizado que tanto está bem definido quanto respeita Pigou-Dalton no exemplo em questão.

## 2 Definições e interpretação geométrica

Para justificar nossos cálculos e ilustrar nossa análise, ambos baseados nas interpretações desses coeficientes por áreas determinadas pela curva de Lorenz, faz sentido visitar a equivalência de suas definições usuais, seja a partir da soma das diferenças absolutas, seja a partir das áreas.

Para um vetor de distribuição de renda  $(Y_1, \dots, Y_M)$ , tal que  $Y_j \leq Y_{j+1}$  para todo  $j$  e  $\sum_{j=1}^M Y_j > 0$ , estão definidos a soma das diferenças absolutas

$$S = \sum_{i,j=1}^M |Y_j - Y_i| = 2 \sum_{j=1}^M \sum_{i<j} (Y_j - Y_i),$$

e o coeficiente de Gini  $G$ , como a razão entre  $S$  e o maior valor que  $S$  pode atingir para uma redistribuição que preserva o número de famílias e a soma total das rendas. Se consideramos apenas distribuições com rendas não negativas, é conhecido que  $S$  assume seu máximo em  $(0, \dots, 0, \sum Y_j)$ , portanto

$$G = \frac{S}{2(M-1) \sum_1^M Y_j}. \quad (1)$$

Uma observação sobre a definição do coeficiente de Gini é que se pode considerar (1) ou

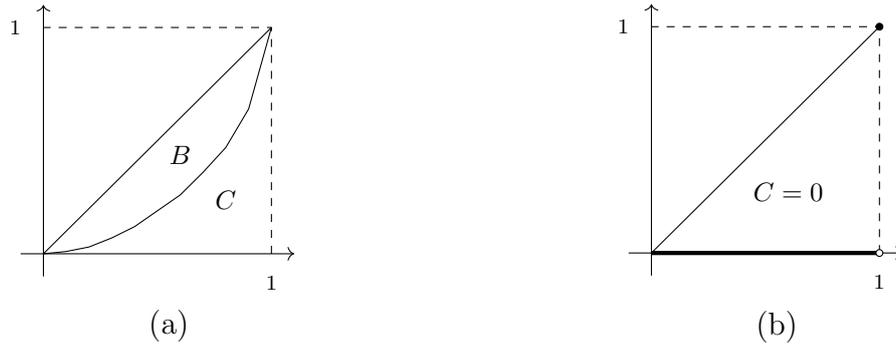
$$G = \frac{S}{2M \sum Y_j}. \quad (2)$$

Usando (1), teremos  $G(0, \dots, 0, \sum Y_j) = 1$ ; usando (2), sempre teremos  $G < 1$ , pois na situação de desigualdade máxima temos  $G = 1 - 1/M$ . As definições ficam mais próximas entre si na medida em  $M$  cresce. No caso do nosso conjunto de dados, como o número de observações é  $M = 27985$ , a diferença será desprezível (erro menor que  $10^{-4}$ ). De acordo com [BPV], embora (1) seja a definição original de Corrado Gini em 2014, ela não satisfaz o princípio da população de Dalton. A pequena diferença nos permite trabalhar com (2), a exemplo de [CTR].

A curva de Lorenz de uma distribuição é o gráfico de uma função linear por partes, com vértices  $(a_0, b_0) = (0, 0)$  e, para  $j = 1, \dots, M$ ,  $(a_j, b_j) = (\frac{j}{M}, \sum_1^j y_i)$ , onde  $y_i = \frac{Y_i}{\sum_1^M Y_j}$  é a contribuição da  $i$ -ésima família em relação à soma total das rendas. Seu último vértice é  $(1, 1)$ . Na ausência de rendas negativas, fica totalmente dentro do quadrado unitário do plano cartesiano.

A curva de Lorenz é uma linha poligonal, portanto áreas são facilmente calculadas e resulta que  $G$  é a razão entre a área de concentração da distribuição considerada e a área de concentração na desigualdade máxima. Se a distribuição não admite rendas negativas, a área de concentração, denotada por  $B$ , é a área entre a curva da perfeita igualdade e a curva de Lorenz. A área na desigualdade máxima é  $B + C$ , sendo  $C$  a região abaixo da curva de Lorenz e acima do eixo  $x$ . Assim,  $G = \frac{B}{B + C}$ .

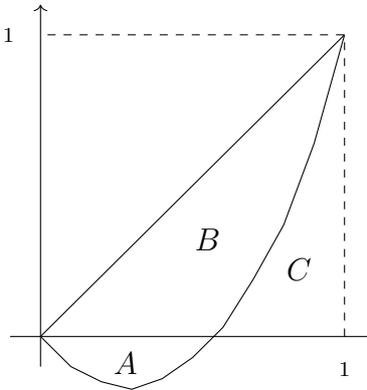
Como escolhemos trabalhar com  $G = \frac{S}{2M \sum_1^M Y_j}$ , então  $B + C = 1/2$ . Se tivéssemos escolhido a definição (1), teríamos  $B + C = \frac{1}{2} - \frac{1}{2M}$ . A mesma convenção deve ser feita para os coeficientes  $G^*$  e  $G^P$ .



**Figura 1.** Curva de Lorenz na ausência de rendas negativas: (a) um exemplo, (b) desigualdade máxima para distribuições contínuas ( $B + C = 1/2$ ).

**Fonte:** elaboração própria.

A presença de rendas negativas implica que curva de Lorenz determina áreas  $A$ ,  $B$  e  $C$ , onde  $A + B$  é a área de concentração, determinada pela curva de Lorenz e a linha da perfeita igualdade, e  $C$  é a área acima do eixo  $x$  e abaixo da curva de Lorenz. A área  $A$  é a parte da área de concentração que fica abaixo do eixo  $x$  (figura 2). Na ausência de rendas negativas,  $A = 0$ . O coeficiente de Gini passa a não ser mais um bom indicador de concentração, já que sua definição é equivalente à escrever  $G = \frac{A + B}{B + C} = 1 + 2(A - C)$ , portanto pode acontecer  $G > 1$ . Isso significa que a concentração máxima precisa ser repensada. De fato, no nosso conjunto de dados, calculamos  $G = 1.20$ .



**Figura 2.** Curva de Lorenz para uma distribuição com rendas negativas.

**Fonte:** elaboração própria.

O ajuste de [CTR] para a definição do coeficiente de Gini, isto é, para a área de concentração máxima de uma redistribuição de  $(Y_1, \dots, Y_M)$ , foi feito a partir do cálculo de  $G$  supondo, como hipótese simplificadora, que existe um número  $k \in \{1, \dots, M - 1\}$  para o qual  $\sum_1^k Y_j = 0$ . Esta hipótese não ocorre em geral, mas é importante para o raciocínio. Se esse é o caso, obtemos

$$\begin{aligned}
G &= \frac{1}{M} \sum_{j=1}^M \sum_{i < j} (y_j - y_i) \\
&= 1 + \frac{2}{M} \left[ \sum_1^k j y_j - \sum_{k+1}^M y_j (1 + 2(M - j)) \right].
\end{aligned}$$

Considerando redistribuições que preservam o número de famílias e a soma total das rendas, os autores consideram como caso extremo o vetor  $(Y_1, \dots, Y_{M-1}, \sum Y_j)$  onde  $\sum_1^{M-1} Y_j = 0$ , para a qual o cálculo do coeficiente de Gini é

$$G = 1 - \frac{1}{M} + \frac{2}{M} \sum_{j=1}^{M-1} j y_j,$$

sempre menor que  $1 + \frac{2}{M} \sum_{j=1}^{M-1} j y_j$ , e os autores então definem  $G^*$  como a razão entre o coeficiente de Gini  $G$  e tal cota superior:

$$G^* = \frac{G}{1 + \frac{2}{M} \sum_{j=1}^{M-1} j y_j} = \frac{1 + \frac{2}{M} \left( \sum_1^k j y_j - \sum_{k+1}^M y_j (1 + 2(M - j)) \right)}{1 + \frac{2}{M} \sum_{j=1}^{M-1} j y_j}. \quad (3)$$

Desse modo, na ausência de rendas negativas,  $k = 0$  e  $G^* = G$ , portanto  $G^*$  assume seu valor mínimo,  $G^* = 0$ , quando cada família tem a mesma renda ( $y_j = y_i$  para todo par  $i, j$ ). Como, mesmo com rendas negativas, o número  $\sum_{j=1}^{M-1} j y_j$  é positivo, na situação extrema  $k = M - 1$  tem-se

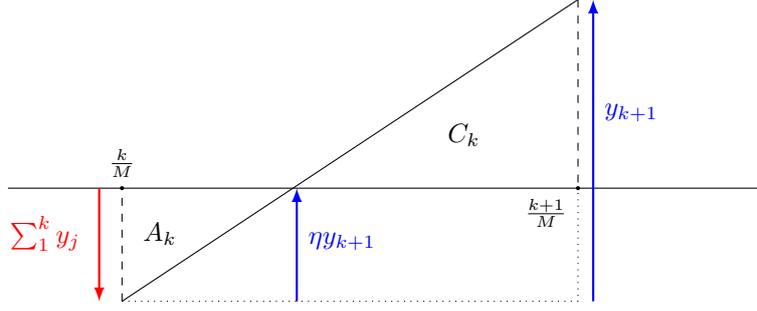
$$G^* = \frac{1 - \frac{1}{M} + \frac{2}{M} \sum_1^{M-1} j y_j}{1 + \frac{2}{M} \sum_{j=1}^{M-1} j y_j} < 1,$$

e podemos criar exemplos de distribuições que tenham  $G^*$  tão próximo de 1 quanto quisermos, pois  $M \rightarrow \infty \Rightarrow G^* \rightarrow 1$ .

A interpretação por áreas, comparando a área de concentração com a área na desigualdade máxima é, a exemplo de  $G$ , simples e intuitiva: a equação (3) implica que uma maneira equivalente de se calcular  $G^*$  é (lembrando que estamos considerando  $B + C = 1/2$ ),

$$G^* = \frac{1 + 2(A - C)}{1 + 2A} = \frac{G}{1 + 2A} = \frac{A + B}{A + B + C}.$$

Então,  $A + B + C$  representa a área de concentração na máxima desigualdade para [CTR]. A interpretação por áreas, por sua vez, determina como deve ser formulado, em função das contribuições  $y_j$  de cada família, o cálculo de  $G^*$  no caso geral, identificando em  $(Y_1, \dots, Y_M)$  o número  $k \in \{1, \dots, M - 1\}$  tal que



**Figura 3.** Interseção da curva de Lorenz com o eixo  $x$ .  
**Fonte:** elaboração própria.

$$\sum_1^k y_j < 0 \text{ e } \sum_1^{k+1} y_j > 0.$$

Para calcular a contribuição da área do tipo  $A$  para uma curva de Lorenz como essa, calcula-se a área entre os vértices  $(0, 0)$  e  $(\frac{k}{M}, \sum_1^k y_j)$  (trapézios), somada à área  $A_k$  (figura 3), limitada pelo eixo  $x$  e a curva de Lorenz entre os pontos  $x = \frac{k}{M}$  e  $x = \frac{k}{M} + \eta \frac{1}{M}$ . O parâmetro  $\eta$  é definido por semelhança de triângulos como sendo a proporção de  $y_{k+1}$  que somada a  $\sum_1^k y_j$  resulte zero, isto é,  $\eta = \frac{-\sum_1^k y_j}{y_{k+1}}$ . Consequentemente  $A_k = \frac{1}{2M} \eta^2 y_{k+1}$ .

Analogamente, a contribuição do tipo  $C$  depende da área  $C_k = \frac{1}{2M} (1 - \eta)^2 y_{k+1}$ . A partir desses cálculos, é possível incluir os termos correspondentes na equação (3). A fórmula indicada por Chen et al. tinha uma imprecisão, que foi corrigida por Berrebi e Silber (1985) [BS], mas não é necessário explicitá-la aqui, pois, como já dissemos, nossos cálculos são realizados a partir das áreas.

Em outro artigo notável, Raffinetti, Siletti e Vernizzi (2014) percebem que, na presença de rendas negativas, a situação de desigualdade máxima para uma distribuição com soma total  $\sum_1^M Y_j = T_p - T_n > 0$ , onde  $T_p$  é a soma total das rendas positivas ou nulas e  $T_n$  o valor absoluto da soma total das rendas negativas, é determinada pela redistribuição  $(-T_n, 0, \dots, 0, T_p)$ , desde que as redistribuições “permitidas”, além de preservarem a soma total das rendas e o número de famílias, sejam tais que as rendas de maior magnitude, tanto positiva quanto negativa, não ultrapassem  $T_p$  e  $-T_n$ , respectivamente. Com a mesma ideia de comparar a área de concentração da distribuição com a área de concentração na desigualdade máxima, isto é, dividir o coeficiente de Gini  $G$  pelo valor máximo que pode ser atingido, calculam a soma das diferenças absolutas para a redistribuição  $(-T_n, 0, \dots, 0, T_p)$ :

$$S = \sum_{i,j=1}^M |Y_j - Y_i| = 2(M - 1)(T_p + T_n),$$

o que leva os autores a definir o coeficiente  $G^P$ :

$$G^P = \frac{S}{2(M - 1)(T_p + T_n)}.$$

De Battisti et al. (2019) notam que, usando  $M - 1$  no denominador da definição de  $G$ , tem-se

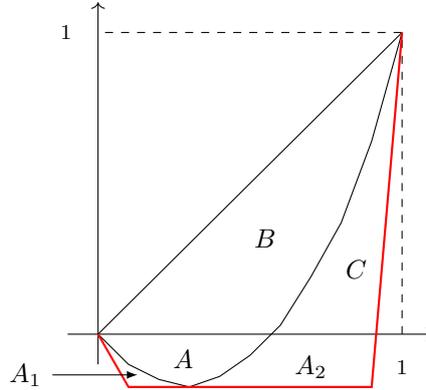
$$G = \frac{S}{2(M-1)\sum Y_j} = \frac{S}{2(M-1)(T_p - T_n)},$$

portanto  $G^P = G \left( \frac{T_p - T_n}{T_p + T_n} \right)$ . Para que essa relação seja válida no nosso caso, temos que usar como definição  $G^P = \frac{S}{2M(T_p + T_n)}$ .

Para entender o que isso significa em termos de áreas, lembramos que em [CTR] a área de concentração máxima foi caracterizada pela imposição que a curva de Lorenz cruze o eixo  $x$  cada vez mais próximo do ponto  $(1,0)$ , minimizando a área  $C$ , mas deixando invariante a área  $A$ . Já a desigualdade máxima em [RSV] maximiza a área do tipo  $A$ : uma definição equivalente para  $G^P$  é

$$G^P = \frac{A + B}{(A_1 + A + A_2) + B + C},$$

onde  $A_1 + A + A_2$  representa a área delimitada pela parte da curva de Lorenz da desigualdade máxima que fica abaixo do eixo  $x$  (figura 4).



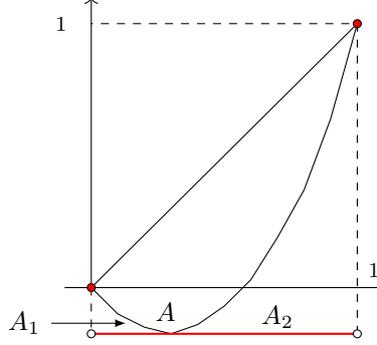
**Figura 4.** Curva de Lorenz de uma distribuição com rendas negativas e curva de Lorenz na desigualdade máxima segundo [RSV] (em vermelho).

**Fonte:** elaboração própria.

A figura 4 é um argumento geométrico de que  $S = 2(M - 1)(T_p + T_n)$  é o valor máximo que pode ser atingido por uma redistribuição com soma total das rendas  $T_p - T_n$ . A condição de comparabilidade entre redistribuições de [RSV] é equivalente a dizer que duas redistribuições quaisquer tenham curvas de Lorenz com mesmo valor mínimo.

Como  $G^P = G \left( \frac{T_p - T_n}{T_p + T_n} \right)$  e  $G^* = G \left( \frac{1}{1 + 2A} \right)$ ,  $G^P$  pode ser interpretado como uma correção do fator multiplicativo de  $G^*$ , considerando, em vez de  $A$ , a área máxima  $A_1 + A + A_2$ . De fato, quando consideramos a aproximação  $B + C = 1/2$ , igualando  $\frac{T_p - T_n}{T_p + T_n}$

a  $\frac{1}{1+2A}$ , obtemos  $A = \frac{T_n}{T_p - T_n}$ , que é precisamente a área do retângulo delimitado pelo eixo  $x$  e a linha  $y = -\frac{T_n}{T_p - T_n}$  entre  $x = 0$  e  $x = 1$ , que é a área do tipo  $A$  da curva de Lorenz para a situação extrema de uma distribuição contínua (figura 5).



**Figura 5.** Curva de Lorenz de uma distribuição com rendas negativas e curva de Lorenz na desigualdade máxima para distribuições contínuas.

**Fonte:** elaboração própria.

Observamos novamente que essas considerações podem ser feitas para distribuições discretas, usando-se  $M - 1$  no denominador, obtendo o valor exato dos coeficientes. Raffinetti et al. inclusive calcularam a área máxima de concentração nesse caso, obtendo

$$(A_1 + A + A_2) + B + C = \left( \frac{1}{2M} + \frac{M-2}{M} + \frac{1}{2M} \cdot \frac{T_n}{T_p} \right) \left( \frac{T_n}{T_n - T_p} \right) + \frac{1}{2} - \frac{1}{2M} \left( \frac{T_p - T_n}{T_p} \right).$$

Os autores comparam essa expressão com a área de concentração na desigualdade máxima em [CTR] e concluem que não tinha sido considerada a parcela  $-\frac{1}{2M} \left( \frac{T_p - T_n}{T_p} \right)$ . Isto de fato ocorre, mas é devido à definição em [CTR] usar, no denominador,  $M$  em vez de  $M - 1$ , o que acarreta em  $B + C = \frac{1}{2}$  e a parcela apontada por [RSV] desaparece. Porém não é isso o que diferencia os coeficientes  $G^*$  e  $G^P$ , e sim o fato de que as áreas  $A_1$  e  $A_2$  não foram consideradas por Chen et al. na desigualdade máxima.

O movimento da curva de Lorenz de uma dada distribuição para a desigualdade máxima, em [CTR], é levar a interseção da curva com o eixo  $x$  o mais à direita possível, preservando a área  $A$ . Já em [RSV], o movimento é puxar o ponto de mínimo da curva o mais à esquerda possível e também empurrá-lo o mais à direita possível (preservando o valor mínimo), trocando um ponto de mínimo por um intervalo de mínimo, maximizando o comprimento do intervalo.

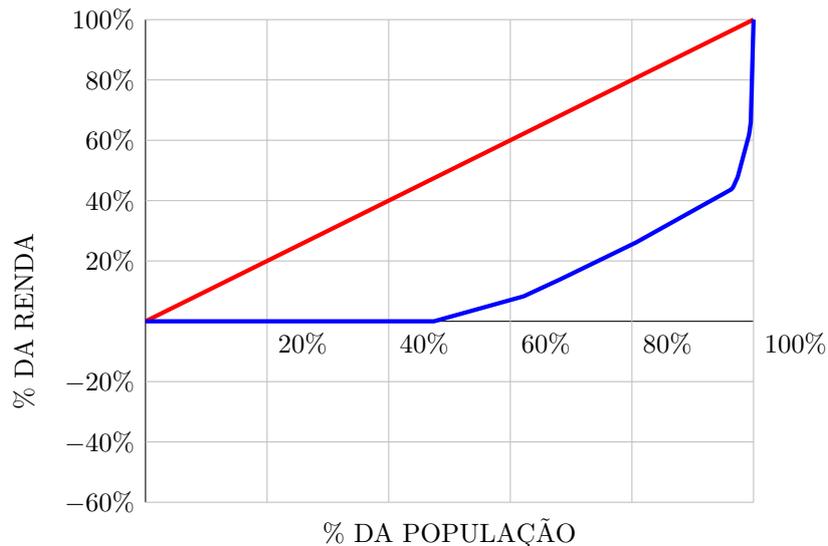
### 3 Soluções encontradas para nosso conjunto de dados.

Como já apontado, nosso conjunto de dados nasce da tentativa de incluir pessoas escravizadas na conta da renda e tentar estimar como seria a desigualdade de renda numa

sociedade escravocrata. Em [SMGG] está explicitada a metodologia que permite inferir a renda dos escravizados a partir de fontes como o Censo Geral do Império do Brasil de 1872, as Listas de Qualificação Eleitoral de 1876, inventários *post-mortem* da década de 1870 (totalizando 367 inventários) e periódicos locais, especialmente “O Pharol” (1876-1879). As informações contidas em cada uma dessas fontes são, claro, limitadas e necessitaram de certos ajustes, como os recortes para nosso local de análise, o município de Juiz de Fora - MG.

O vetor de distribuição de renda contempla  $M = 27985$  indivíduos (e não famílias, como é mais comum, devido ao exposto em Andrade (1995) [A], que menos de 1/3 da população escravizada pertencia a algum tipo de arranjo familiar). Desses, 13278 eram escravizados. Esta população está organizada em grupos de trabalhadores e a distribuição de renda, segundo [SMGG], está dividida em 33 grupos de trabalhadores livres e três grupos de escravizados. Por exemplo, há um grupo de 4128 pessoas *costureiras e especializadas em serviços domésticos*, com renda anual individual de 150\$000 réis, um grupo de 99 *especialistas em trabalhos em madeiras* e um grupo 12 *especialistas em trabalhos em edificações*, ambos com renda anual individual de 300\$000 réis. Os três grupos de escravizados são definidos em analogia a trabalhadores rurais, trabalhadores urbanos sem profissão definida (serventes, diaristas, etc.) e trabalhadores urbanos especializados (artesãos, serralheiros, carpinteiros, etc.). Escravizados trabalhadores rurais tiveram sua renda estimada a partir da média de rendas das fazendas de café e da produtividade anual estimada para escravizados. Trabalhadores urbanos não qualificados tiveram sua renda estimada a partir do valor de seu aluguel anual. Já trabalhadores especializados, a partir da renda de trabalhadores livres com mesma qualificação. Todos estes dados estão disponíveis, bastando solicitar.

Para o cenário de imputação de renda zero aos escravizados, não há dificuldade adicional, basta aumentar o número de observações no vetor de distribuição de renda, mantendo as rendas positivas dos trabalhadores livres. O índice a ser calculado é  $G$ , pois não há rendas negativas. Obtivemos  $G = 0,77$ . A curva de Lorenz desta distribuição está ilustrada na figura 6.



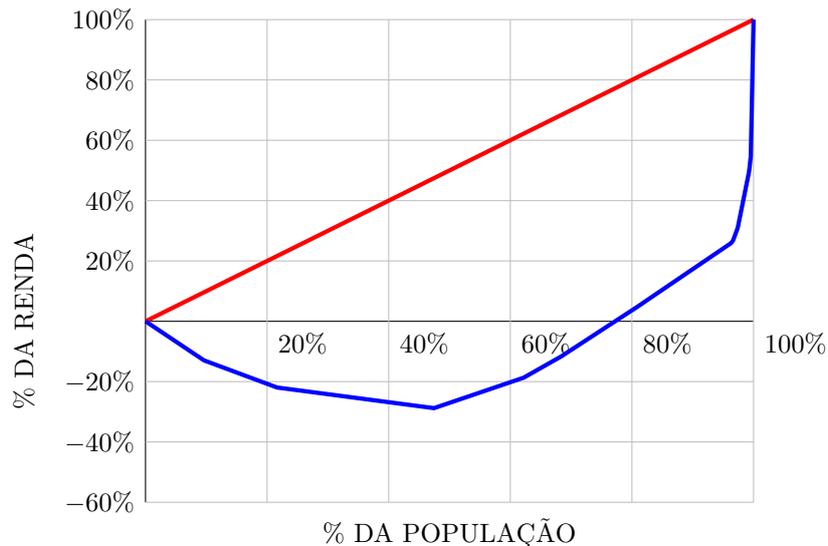
**Figura 6.** Curva de Lorenz para imputação de renda zero a pessoas escravizadas.  
**Fonte:** elaboração própria

Para construir os vetores de distribuição de renda com imputações de rendas positivas ou negativas aos escravizados, uma vez definidas as rendas anuais dos grupos de escravizados, uma compensação de sinal oposto deve ser feita aos trabalhadores livres, para que a soma total das rendas seja preservada.

A redistribuição com rendas positivas foi obtida imputando-se a cada escravizado uma renda proveniente de grupos de trabalhadores livres, sendo a compensação calculada em dois passos. Primeiro, usamos a hipótese que cada grupo de trabalhadores livres possuía escravizados trabalhadores urbanos (2º e 3º grupos descritos acima) proporcionalmente à renda do grupo e retiramos da renda desse grupo, igualmente entre os indivíduos, a proporção correspondente da soma das rendas dos escravizados urbanos. No segundo passo, nos restringimos aos escravizados trabalhadores rurais e aos lavradores proprietários de escravos, usando novamente a hipótese de proporcionalidade. No primeiro passo, todos os grupos de trabalhadores livres têm sua renda modificada para baixo, no segundo, apenas os lavradores.

Para construir o vetor de distribuição de renda com imputações de rendas negativas aos escravizados, que é o que nos interessa mais de perto no momento, a hipótese de proporcionalidade também se aplica, mas no sentido contrário: aos escravizados estão atribuídas rendas negativas e as compensações aos trabalhadores livres são positivas. Os coeficientes normalizados foram calculados, obtendo-se  $G^* = 0,94$  e  $G^P = 0,76$ . A curva de Lorenz para esta distribuição está ilustrada na figura 7.

As duas distribuições estudadas são redistribuições da mesma renda a um mesmo número de famílias (no nosso caso, indivíduos). O movimento do vetor resultado da imputação de renda zero para aquele que resulta da imputação de renda negativa é regressivo, isto é, as transferências de renda são dos mais pobres para os mais ricos. De certa maneira pode-se dizer que cada incremento de renda para um determinado grupo de trabalhadores livres é proveniente da expropriação da mesma quantidade de renda dos grupos de escravizados. Mesmo não sendo transferências individuais de renda, movimentos como esse



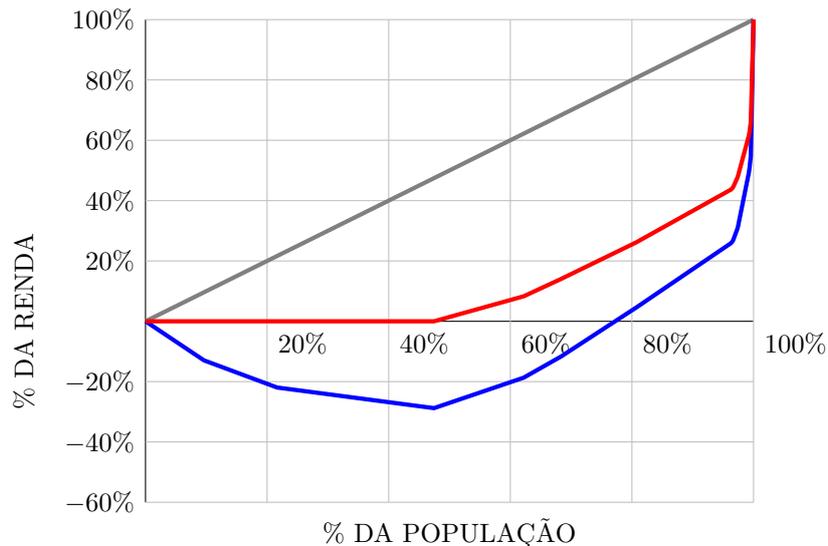
**Figura 7.** Curva de Lorenz para imputação de renda negativa aos escravizados.  
**Fonte:** elaboração própria

são composições de finitas transferências individuais (veja Castagnoli, E. e Muliere, P., (1989) [CP]). Neste trabalho, os autores destacam que vetores de distribuição de renda com mesma soma total de renda e mesmo número de observações têm uma ordem parcial natural: são comparáveis quando as respectivas curvas de Lorenz não se interceptam, sendo o vetor cuja curva de Lorenz está abaixo chamado de “mais desigual”. Os enunciados em [CP] são referentes a distribuições de rendas não negativas, mas a generalização é imediata. Note que esse é o nosso caso, veja a plotagem das duas curvas de Lorenz juntas na figura 8. Claro, se poderia perguntar se de fato as curvas não se cruzam, isto não é garantido pela figura, mas sim pelos dados disponíveis.

Mas ainda há um problema a ser tratado na comparabilidade dessas distribuições. O fato de que as transferências realizadas não preservam a ordem anteriormente existente. Especificamente, as mudanças de ordem após a transferência proporcional do cenário de renda zero para o cenário de renda negativa são apenas:

(a) O grupo de renda zero é levado em três grupos de rendas negativas, que naturalmente ficam ordenados pela ordem decrescente da magnitude de suas rendas. Isso não é uma preocupação, pois a ordem  $y_i \leq y_{i+1}$  é ainda preservada: no máximo, dois valores consecutivos que eram ambos nulos, depois da transferência são diferentes, com o menor à esquerda.

(b) A outra única mudança de posição acontece entre três grupos de trabalhadores que estão entre os 3% mais ricos da população. Precisamente entre os pontos de abscissa  $x = 0,9747$  e  $x = 0,9921$  da curva de Lorenz das distribuições. Especificamente, antes da compensação tínhamos 483 lavradores com renda individual anual 2.000\$000, três farmacêuticos com renda anual individual anual 2.037\$483 e quatro juizes com renda 2.100\$000, que após a compensação passaram a ter rendas respectivamente iguais a 2.678\$998, 2.484\$806 e 2.561\$049.



**Figura 8.** Curvas de Lorenz para as situações em estudo: vermelha para a imputação de renda zero aos escravizados, azul para imputação de renda negativa.

**Fonte:** elaboração própria.

Uma pequena modificação pode então ser feita nesses valores, a saber, somar a renda total desses três grupos e dividir pelo número de indivíduos, de modo que todos eles agora tenham a mesma renda, cujo valor é 2.676\$846. Este ajuste é pequeno suficiente tanto para não modificar a essência da compensação proporcional quanto para não modificar significativamente os coeficientes: para esta nova distribuição, os coeficientes de Gini normalizados são  $G^* = 0,93757$  e  $G^P = 0,764223$ , enquanto que naquela obtida pela compensação proporcional tínhamos  $G^* = 0,93758$  e  $G^P = 0,764224$ . As curvas de Lorenz são praticamente idênticas.

Então, se consideramos o movimento do cenário de renda zero para o cenário de renda negativa modificado como acima, temos uma transferência regressiva, portanto, pelo princípio de Pigou-Dalton, é esperado que um coeficiente que se proponha a medir desigualdade aumente (ou pelo menos não diminua - princípio “fraco” de transferência). E, como vimos, não é o que acontece com  $G^P$ . De fato, com cinco casas decimais, para o cenário de renda zero temos  $G = 0,77396$ . De qualquer maneira, talvez se possa dizer que, na prática, o coeficiente não diminui, o que significa que ao menos o princípio fraco de transferências estaria satisfeito. Porém são cenários claramente distintos e o esperado é que o de renda negativa represente uma desigualdade significativamente maior do que o de renda zero.

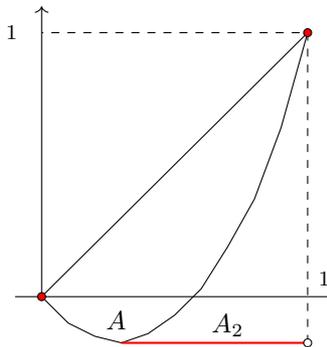
A ideia para propor uma solução para estes problemas, pelo menos nesse caso específico, é considerar os conceitos fundamentais presentes em [CTR] e [RSV]. A maximização proposta por Chen et al. tem essencialmente a mesma ideia da desigualdade máxima para rendas não negativas, isto é, deslocar o ponto  $x_z$ , a raiz positiva da função de Lorenz da distribuição, para a direita, levando-o o mais próximo de  $x = 1$  possível. Porém, ao preservar a área  $A$ , a maximização proposta por [CTR] não determina a curva de Lorenz, portanto não está bem definida. Já a situação de máxima desigualdade descrita em [RSV] é obtida a partir de uma distribuição dada transferindo todas as rendas positivas

ou nulas para a direita, no sentido dos mais ricos, até que um só indivíduo receba toda a renda positiva e o mesmo é feito no sentido oposto: as rendas negativas são transferidas para a esquerda, para os mais pobres, até que um só indivíduo “receba” toda a renda negativa. Ocorre que, no nosso conjunto de dados, rendas negativas são imputadas teoricamente, então faz sentido considerar como situação de desigualdade máxima apenas a movimentação interna das rendas positivas, deixando inalterada a parte da distribuição referente às negativas. Faz sentido não piorar ainda mais a desigualdade dentro do grupo das pessoas escravizadas.

A ideia de [RSV] aplicada apenas ao lado direito do ponto de mínimo da curva de Lorenz da distribuição leva a uma redistribuição que preserva a parte de rendas negativas e que bem poderia ter sido usada como a pior possível na normalização em [CTR], o que em termos de cálculo significaria definir o coeficiente de Gini normalizado por:

$$G^u = G \left( \frac{1}{1 + 2(A + A_2)} \right),$$

onde  $A + A_2$  representa a área delimitada pela parte da curva de Lorenz da desigualdade máxima que fica abaixo do eixo  $x$ , mas acima da curva de Lorenz entre  $x = 0$  e o primeiro ponto de mínimo da função. (figura 9). Devido à natureza da construção da situação máxima, escolhemos o índice  $u$ , fazendo referência à palavra unilateral.



**Figura 9.** Curva de Lorenz de uma distribuição com rendas negativa: a desigualdade máxima obtida por “deslocamento unilateral” do ponto de mínimo.

**Fonte:** elaboração própria.

Algebricamente, o se o vetor de distribuição de renda é  $(Y_1, \dots, Y_\ell, Y_{\ell+1}, \dots, Y_M)$ , onde  $\ell$  é o índice tal que  $Y_\ell < 0$  e  $Y_{\ell+1} \geq 0$ , então o vetor correspondente à desigualdade máxima está definido por  $(Y_1, \dots, Y_\ell, 0, \dots, 0, T_p)$ . Geometricamente fica claro que qualquer outra redistribuição que fixe as rendas negativas terá sua curva de Lorenz totalmente acima ou coincidente com a da desigualdade máxima. Em outras palavras, o coeficiente  $G^u$  fica bem definido.

Também é claro a partir da definição que, na ausência de rendas negativas,  $G^u$  coincide com  $G$ . Os valores que obtivemos no nossos conjuntos de dados foram  $G^u = 0,907738$  após a compensação proporcional e  $G^u = 0,907736$  após a modificação que preservou a ordem crescente de renda dos indivíduos. Portanto, o coeficiente  $G^u$  está bem definido,

respeitou Pigou-Dalton nesse caso pois cresceu após uma transferência regressiva de renda e tem um valor que a nosso ver representa melhor a situação que  $G^P$ .

Assim, ao mesmo tempo que acreditamos ter avançado um pouco na discussão sobre coeficientes de Gini normalizados para distribuições com rendas negativas, entendemos que esse estudo está longe de ser definitivo.

## Referências

ACEMOGLU, Daron; GARCÍA-JIMENO, Camilo; ROBINSON, James A. Finding Eldorado: Slavery and long-run development in Colombia. **Journal of Comparative Economics**, v. 40, n. 4, p. 534-564, 2012.

ALMICO, Rita. **Fortunas em movimento: as mudanças ocorridas na riqueza pessoal? Juiz de Fora/1870-1920**. 2001. Tese de Doutorado. Dissertação de Mestrado. Campinas, UNICAMP.

ALMICO, Rita. A Dança da Riqueza: Variações na Fortuna Pessoal em Juiz de Fora 1870/1914. **Revista Eletrônica de Economia (Juiz de Fora)** 1: 7-38, 2003.

ALVARENGA, Thiago. Desigualdade Econômica na Cidade do Rio de Janeiro nas Últimas Décadas do Império do Brasil. In: MATHIAS, J. F. C. M.; SARAIVA, L. F. **Igual-Desigual. História e Economia das Desigualdades Antes, Durante e Após a Pandemia**, São Paulo, Hucitec, 2020.

[A] ANDRADE, Rômulo Garcia de. **Limites impostos pela escravidão a comunidade escrava e seus vínculos de parentesco: Zona da Mata de Minas Gerais, século XIX. 2v. São Paulo: USP**. Tese de Doutorado, 1995.

[BS] BERREBI, Zeev M.; SILBER, Jacques. The Gini coefficient and negative income: a comment. **Oxford Economic Papers**, v. 37, n. 3, p. 525-526, 1985.

BOTELHO, Tarcisio R. Censos e construção nacional no Brasil Imperial. **Tempo social**, v. 17, p. 321-341, 2005.

BRAZIL. 1875. **Collecção das Leis do Imperio do Brazil de 1875**. Typographia Nacional, 1875.

CARDOSO, Ciro Flamarion Santana. O Trabalho da Colônia. In: LINHARES, Maria Yedda Leite; CARDOSO, Ciro Flamarion Santana. **História Geral do Brasil**, p. 69-88. Campus, 1990.

[CM] CASTAGNOLI, Erio; MULIERE, Pietro. A Note On Inequality Measures And The Pigou-Dalton Principle Of Transfers. In: OAGUM, Camilo; ZENGA, Michele (Eds.). **Studies in Contemporary Economics**, p. 171-182, Springer-Verlag, 1989.

CHALHOUB, S. **Visões de Liberdade: Uma História das Últimas Décadas da Escravidão na Corte**. Companhia das Letras, 1990.

[CTR] CHEN, Chau-Nan; TSAUR, Tien-Wang; RHAI, Tong-Shieng. The Gini coefficient and negative income. **Oxford Economic Papers**, v. 34, n. 3, p. 473-478, 1982.

[BPV] DE BATTISTI, Francesca; PORRO, Francesco; VERNIZZI, Achille. The Gini coefficient and the case of negative values. **Electronic Journal of Applied Statistical Analysis (EJASA)**, v. 12, n. 1, p. 85-107, 2019.

DELL, Melissa. The persistent effects of Peru's mining mita. **Econometrica**, v. 78, n. 6, p. 1863-1903, 2010. ENGERMAN, Stanley L.; SOKOLOFF, Kenneth L. Factor endowments, inequality, and paths of development among new world economics. **NBER Working Paper No. 9259**, 2002.

FRAGOSO, J.; PITZER, R. R. Barões, Homens-Livres Pobres e Escravos: Notas sobre o uso de inventários post-mortem na pesquisa histórica. **Arrabaldes** (Rio de Janeiro) 1 (2): 29-52.

GANDRA, Rodrigo Mendes. O debate sobre a desigualdade de renda no Brasil: da controvérsia dos anos 70 ao pensamento hegemônico nos anos 90. **História Econômica & História de Empresas**, v. 8, n. 1, 2005.

GRAHAM, Richard. **Clientelismo e Política no Brasil do Século XIX**. Rio de Janeiro: Editora da UFRJ, 1997.

GUIMARÃES, Elione Silva. **Violência entre parceiros de cativo: Juiz de Fora, segunda metade do século XIX**. Annablume, 2006.

GUIMARÃES, Elione Silva. **Terra de preto: usos e ocupação da terra por escravos e libertos (Vale do Paraíba mineiro, 1850-1920)**. Niterói, EDUFF, 2009.

GUIMARÃES, Elione Silva, SARAIVA, Luiz Fernando, SARAIVA, Paulo. Desigual entre os Desiguais. Apontamentos para uma História das Desigualdades Raciais no Brasil: Juiz de Fora na Década de 1870. In: MATHIAS, J. F. C. M.; SARAIVA, L. F. **Igual-Desigual. História e Economia das Desigualdades Antes, Durante e Após a Pandemia**, São Paulo, Hucitec, 103?133, 2020.

LACERDA, Antônio Henrique Duarte. **Os padrões das alforrias em um município cafeeiro em expansão: Juiz de Fora, Zona da Mata de Minas Gerais, 1844-88**. Annablume, 2006.

LACERDA, Antônio Henrique Duarte. **NEGÓCIOS DE MINAS: Família, fortuna, poder e redes de sociabilidades nas Minas Gerais ? A Família Ferreira Armonde (1751-1850)**. Intermeios, 2013.

MCDONALD, Jill M., HOPPE, Robert A., BANKER, David. Structural Change in US Agriculture. In: **Policy Reform and Adjustment in the Agricultural Sectors of Developed Countries**, 2006.

MUSSE, Christina Ferraz. A imprensa e a memória do lugar: Juiz de Fora (1870/1940). In: **XII Congresso Brasileiro de Ciências da Comunicação da Região Sudeste, Juiz de Fora**. In: COUTINHO, Inês; LEAL, Paulo Roberto Figueira E-papers, 2009.

NUNN, Nathan. Slavery, Inequality, and Economic Development in the Americas. **Institutions and Economic Performance**, 15: 148?180, 2008.

PEÇANHA, Vinicius. Origens Históricas da Desigualdade e Futuro Pós-Covid-19. In: MATHIAS, J. F. C. M.; SARAIVA, L. F. **Igual-Desigual. História e Economia das Desigualdades Antes, Durante e Após a Pandemia**, São Paulo, Hucitec, 82-90, 2020.

PIRES, Anderson. **Capital Agrário, investimento e crise da cafeicultura de Juiz de Fora (1870-1930)**. Niterói: UFF/ICHF, 1993. Dissertação de mestrado.

PIRES, Anderson. **Café, Finanças e Indústria. Juiz de Fora - MG**. Funalfa Edições, 2009.

PUNTONI, Pedro. Os recenseamentos gerais do Brasil no século XIX: 1872 e 1890. **CEBRAP, Sao Paulo**, 2004.

[RSV] RAFFINETTI, Emanuela et al. Inequality measures and the issue of negative income. In: **47th SIS Scientific Meeting of the Italian Statistica Society Proceedings**. CUEC (Cooperativa Universitaria Editrice Cagliariitana), 2014. p. 1-7.

SARAIVA, Luiz Fernando; ALMICO, Rita. Demografia Escrava e Produção Econômica na Zona da Mata Mineira. **História e Economia**, v. 12, n. 1, p. 85-102, 2014.

SARAIVA, Luiz Fernando; CASTRO, Pedro G.; NABARRO, Wilson. **Atlas Histórico Econômico do Brasil no século XIX**. Ed. do Senado, 2024.

[SMGG] SARAIVA, L. F.; MATHIAS, J. F. C. M.; GUIMARAS, E. S.; & GEBARA NETO, S. (2022). Desigualdades em sociedades escravocratas: uma reflexão metodológica. In F. A. M. de Mattos, J. Hallak Neto, & F. G. Silveira (Eds.), **Desigualdades: visões do Brasil e do mundo (Vol. 3)**. Hucitec.

SCHUTZ, Robert R. On the measurement of income inequality. **The American Economic Review**, v. 41, n. 1, p. 107-122, 1951.

SOARES, Rodrigo Goyena. Estratificação profissional, desigualdade econômica e classes sociais na crise do império. Notas preliminares sobre as classes imperiais. **Topoi (Rio de Janeiro)**, v. 20, p. 446-489, 2019.

SODRÉ, Nelson Werneck. **História da imprensa no Brasil**. Mauad Editora Ltda, 1998.

SOUZA, Pedro Ferreira. **Uma história de desigualdade: a concentração de renda entre os ricos, 1926-2013**. Hucitec Editora, 2018.

UNITED NATIONS (2009). System of National Accounts. Disponível em: <http://unstats.un.org/unsd/>

VILLA, Carlos Eduardo Valencia. **Ao longo daquelas ruas: a economia dos negros livres em Richmond e Rio de Janeiro, 1840-1860**. Paco Editorial, 2016.

VITTORETTO, Bruno Novelino. **Fronteiras do Café na Zona da Mata Mineira (1870-1940)**. UFJF, Juiz de Fora, Tese de Doutorado, 2016.

WOODWARD, Kyle. "Timing: Ex Ante, Interim, Ex Post." Disponível em: <https://kylewoodward.com> 2021.