

Microeconomia do Desenvolvimento

Comparação entre as distribuições de renda de São Paulo e Curitiba

Ricardo S. Castro, Yudi S. Tujisoki, Juan B. Villegas

1º Semestre 2021

Introdução

Neste trabalho, analisamos a distribuição de renda das cidades de São Paulo e Curitiba por meio das óticas de desigualdade, pobreza e bem-estar social. Usamos para isso os dados provenientes da PNADc no primeiro trimestre de 2020.

Adotamos como variável principal a medida de renda familiar per capita, construída a partir dos rendimentos dos integrantes de cada domicílio. Apresentamos estatísticas descritivas da amostra para as cidades nas tabelas 1 e 2, bem como no gráfico de distribuição de renda, mas deixamos uma análise mais aprofundada para as seções seguintes.

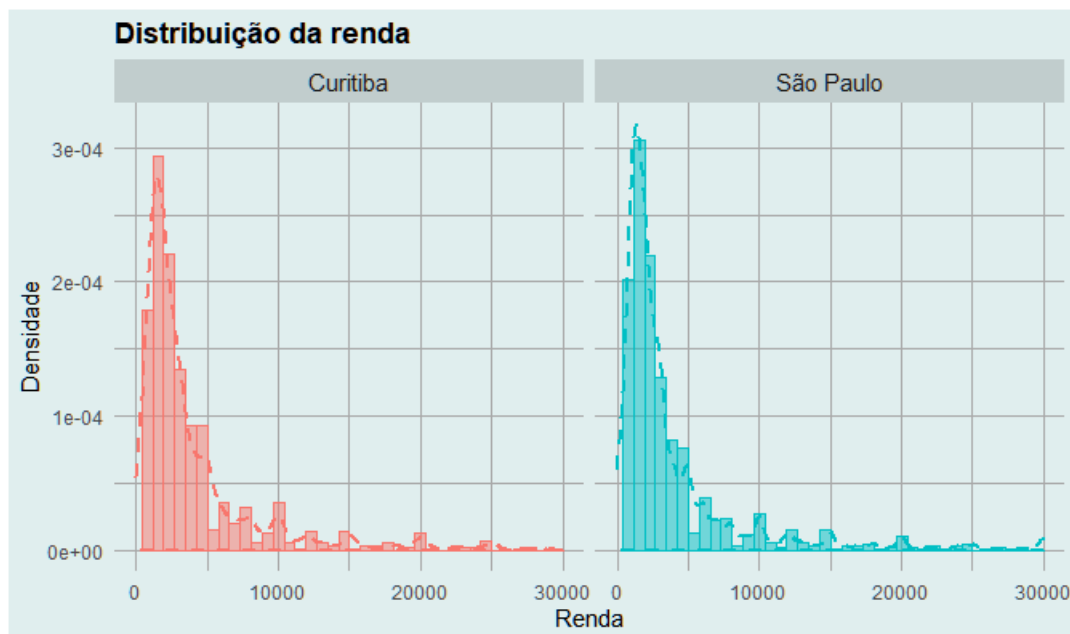
Tabela 1: Medidas descritivas

	Média	Desvio Padrão	Máximo	N
São Paulo	R\$1.802,95	3.476,33	R\$100.000	4416
Curitiba	R\$1.663,41	2.743,69	R\$300.000	2261

Tabela 2: Renda acumulada por percentil

	Percentil mais pobre			Percentil mais rica			
	5%	10%	50%	1%	5%	10%	50%
São Paulo	0.001	0.006	0.106	0.145	0.384	0.531	0.894
Curitiba	0.003	0.008	0.126	0.116	0.324	0.472	0.874

O trabalho está organizado em quatro seções: metodologia, comparação de desigualdade, comparação de pobreza e comparação de bem-estar.



Metodologia

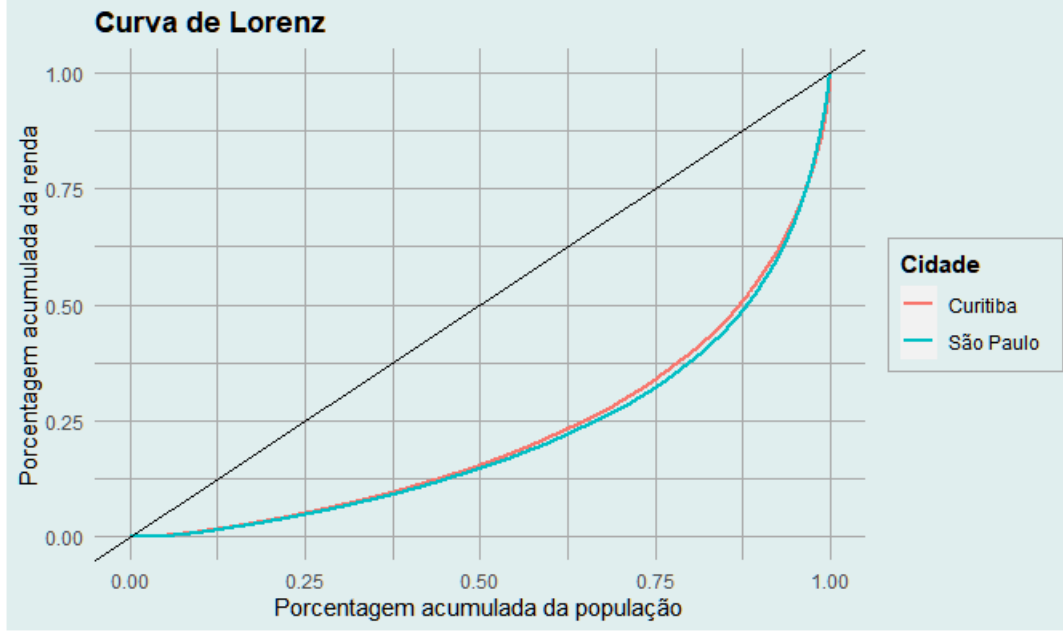
Para este trabalho, escolhemos usar a base de dados da Pesquisa Nacional por Amostra de Domicílios Contínua, PNAD Contínua, do primeiro trimestre de 2020. A PNADc é uma pesquisa feita pelo IBGE que tem como objetivo produzir informações acerca do mercado de trabalho associadas a características educacionais e demográficas. As informações têm representatividade para os níveis geográficos: Brasil, Grandes Regiões, Unidades da Federação, 20 Regiões Metropolitanas que contêm Municípios das Capitais e Região Integrada de Desenvolvimento da Grande Teresina. O universo populacional das distribuições de renda são domicílios.

A PNADc é um base de dados bastante completa, porém, é comum famílias não reportarem a renda precisamente (arredondando, ou alterando o valor real), logo sofremos com erro de medida. Além disso, só levamos em conta famílias que reportaram renda.

Comparando desigualdade

Inicialmente queremos comparar a desigualdade das distribuições de renda das duas cidades abordadas em nosso trabalho, São Paulo e Curitiba, para tal julgamos mais interessante fazer a análise em termos relativos, indo na direção da maioria da literatura sobre o tema. Mais especificamente, faremos a análise das curvas de Lorenz, e aplicaremos medidas Lorenz-consistentes para aprofundar a discussão.

As curvas das duas cidades são bastante próximas até os 30% mais pobres, com Curitiba ligeiramente mais próximo da linha de 45 graus. Porém, São Paulo tem uma quantidade



maior de riqueza concentrada na mão dos 5% mais ricos, em contraste com a maior concentração na “população média” da outra cidade. Como as curvas se cruzam e, portanto, como indica o terceiro critério de Lorenz¹ não podemos fazer nenhuma análise de dominância.

Entretanto, iremos utilizar medidas de desigualdade consistentes com a ordenação de Lorenz, para que - no que tange as medidas escolhidas - seja possível ordenar a igualdade entre os municípios. Usaremos os seguintes índices: Gini, a primeira e a segunda medidas de Theil (Theil 1 e 2, como consta na tabela [3]), e o índice de Atkinson com parâmetros 0.5 e 2 (Atk. 0.5 e Atk. 2).

Tabela 3: Medidas de desigualdade

	Gini	Theil 1	Theil 2	Atk. 0.5	Atk. 2
São Paulo	0.610	0.549	0.557	0.254	0.641
Curitiba	0.627	0.534	0.546	0.249	0.647

O coeficiente de Gini, por ser o mais utilizado na literatura, tem o maior peso em nossa análise e consequentemente consideraremos São paulo (Gini = 0.557) como mais desigual que Curitiba (Gini = 0.546). Ainda assim, é interessante observar as outras medidas.

Ambas as medidas de Theil indicam uma distribuição mais desigual quanto mais próximo de 1, e a primeira é mais sensível à diferenças no topo em oposição à segunda. Ambas indicam São Paulo como mais desigual, mas a razão entre o valor associado às duas cidades é menor na medida mais sensível no topo da distribuição:

$$\frac{\text{Theil } 1_{\text{São Paulo}}}{\text{Theil } 1_{\text{Curitiba}}} = 1.028 > 1.020 = \frac{\text{Theil } 2_{\text{São Paulo}}}{\text{Theil } 2_{\text{Curitiba}}}$$

¹Ver Fields (2001, p. 23)

O índice de Atkinson indica o ganho (normalizado) de utilidade social ao redistribuir perfeitamente a renda da cidade - sob algumas hipóteses restritivas acerca da função de utilidade da cidade - além disso, o parâmetro do índice carrega a ideia de maior aversão à desigualdade. Olhando para os resultados, vemos que com o parâmetro 0.5 São Paulo teria um maior ganho de utilidade com uma redistribuição de renda (mais desigual), mas após aumentá-lo para 2, vemos o contrário. Entretanto, nossa conclusão se manterá na linha do indicado pelo Gini, onde São Paulo é mais desigual.

Comparando pobreza

Seguindo, realizamos uma comparação em relação à pobreza entre as cidades em questão, contudo, antes de partir para a análise propriamente dita, é necessário definirmos uma linha de pobreza. Existem diversas formas de construir uma linha de pobreza. Ela pode ser determinada a partir de critérios científicos, políticos e até mesmo subjetivos. Um método bastante comum é definir a linha de pobreza a partir da abordagem de custo das necessidades básicas. Dentro dessa, o critério mais satisfatório funciona da seguinte maneira: (1) determinar uma cesta de consumo considerado adequado envolvendo componentes alimentares e não alimentares; e (2) estimar o custo dessa cesta de bens e serviços para cada subgrupo da população (urbano e rural, por região etc.). Na prática, para definir essa cesta, escolhe-se uma necessidade nutricional para uma boa saúde (2100 kcal/dia, por exemplo, que é um padrão proposto pela FAO). Em seguida, estima-se o custo de atender a essa necessidade de energia alimentar, usando uma dieta que reflita os hábitos das famílias que se encontram abaixo da linha de pobreza. Estima-se também o custo dos bens não alimentares – como espaço habitacional e eletricidade – e finalmente soma-se com o custo do componente alimentar da cesta, resultando na linha de pobreza.

No entanto, a partir dos trabalhos de Amartya Sen, há novos esforços para medir a pobreza a partir de uma perspectiva multidimensional, como a abordagem de capacidades (*capability approach*). Para Sen, a pobreza é vista como a privação de certas capacidades básicas, não só como insuficiência de renda. Estas capacidades podem variar, como Sen argumentou, das físicas elementares (como ser bem nutrido), estar adequadamente vestido e protegido, evitar morbidade evitável e assim por diante, às conquistas sociais mais complexas, como participar da vida da comunidade, poder aparecer em público sem vergonha, e assim por diante. Uma possível expansão de nosso trabalho seria criar uma linha de pobreza que levasse em conta o trabalho de Sen.

Neste trabalho, trabalhamos com duas opções para a definição de linha de pobreza: a definida pelo IBGE no ano de 2019 (R\$ 436) e o salário mínimo (R\$ 1100). Denotaremos o salário mínimo como a linha máxima de pobreza, $z_{máx}$.

Com o fim de comparar a pobreza nas cidades de São Paulo e Curitiba, empregamos o Índice de Pobreza de Sen e as medidas da classe P_α , explicados em detalhe a seguir.

Índice de Pobreza de Sen

Sejam, respectivamente, z , \bar{y}_p , $\bar{I} = (z - \bar{y}_p)/z$ e G_p a linha de pobreza, a renda média dos pobres, o déficit médio (normalizado) de renda entre os pobres e o coeficiente de Gini de desigualdade de renda entre os pobres. Temos que o índice de pobreza de Sen, P_{Sen} , é dado por

$$P_{Sen} = H[\bar{I} + (1 - \bar{I})G_p]$$

É fácil ver que, tudo o mais constante: (1) quanto maior a proporção de pobres, maior é o índice de Sen; (2) quanto maior a renda média dos pobres, menor é o P_{Sen} ; e (3) quanto maior a desigualdade de renda entre os pobres, maior o P_{Sen} .

Note que o índice de Sen satisfaz aos axiomas de anonimato (**A**), homogeneidade (**H**), monotonicidade (**M**) e sensibilidade distributiva (**DS**), além dos axiomas básicos de reflexividade (**R**) e transitividade (**Tr**).

A classe P_α

Considere primeiro a função de brecha normalizada do indivíduo i :

$$p(y_i, z) \begin{cases} (z - y_i)/z, & \text{se } y_i < z \\ 0, & \text{caso contrário} \end{cases}$$

As medidas da classe P_α são definidas da seguinte forma:

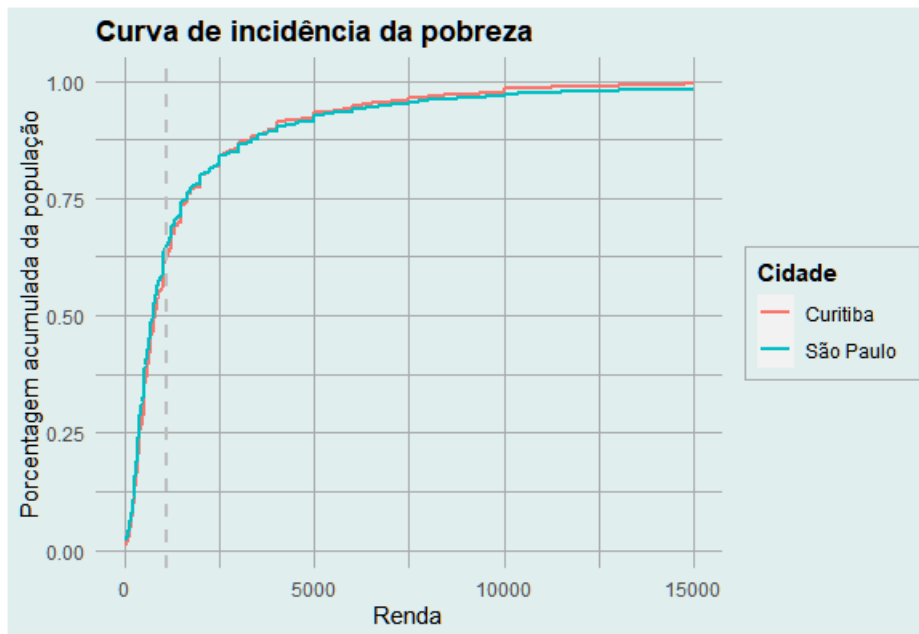
$$P_\alpha = \frac{1}{n} \sum_{i=1}^q [(z - y_i)/z]^\alpha,$$

onde n é o total de pessoas da distribuição de renda, q é o número de pessoas abaixo da linha de pobreza e α , um parâmetro natural (incluindo o zero). Note que quando $\alpha = 0$ temos a proporção do número de pessoas na pobreza (*poverty headcount ratio*). Já quando $\alpha = 1$, tem-se a brecha (normalizada) de renda per capita (*per capita income gap*). Quando $\alpha = 2$, P_2 pondera cada função de brecha normalizada do indivíduo por ela mesmo – isto é, se a renda de um indivíduo é 10% abaixo da linha de pobreza, ela é ponderada por 10%. Logo, à medida que α aumenta, P_α é mais sensível a rendas extremamente baixas.

Os axiomas satisfeitos pelas medidas da classe P_α variam de acordo com o valor do parâmetro α . Quando $\alpha = 0$, P_0 satisfaz aos axiomas de anonimato (**A**) e de homogeneidade (**H**), apenas. Para $\alpha \geq 1$, P_α satisfaz aos axiomas de anonimato, homogeneidade, monotonicidade (**M**) e sensibilidade distributiva (**DS**). Além disso, todas as medidas desta classe satisfazem aos axiomas básicos de reflexividade (**R**) e transitividade (**Tr**).

Resultados

A partir dos teoremas supracitados, podemos realizar a análise de dominância estocástica para obter resultados sobre como os indicadores P_α se comportarão. Começamos traçando a curva de incidência da pobreza (CIP) - a CDF da distribuição:



Vemos que as curvas apenas se cruzam depois da linha de pobreza máxima definida por nós², o que - intuitivamente - indica a maior quantidade de pessoas em níveis baixos de renda para São Paulo. De acordo com Fields (2001, p 89), podemos então definir dominância de primeira ordem, uma vez que as curvas não se cruzam até z_{max} , e aplicando o teorema de que se São Paulo domina em primeira ordem Curitiba, a pobreza (medida com a classe P_α , com $\alpha \geq 1$) de São paulo será maior. Veremos que de fato isso é verdade, mas antes, para robustez, se a definição de z_{max} fosse criticada e fosse feito o argumento de que não podemos falar em dominância de primeira ordem, analisaremos a de segunda ordem.

Montamos a curva de deficiência da pobreza, integrando a CIP, e vemos que elas de fato nunca se cruzam, intuitivamente isso significa que a quantidade da renda da população acumulada em pessoas com renda menor que um x , para qualquer x é maior em São Paulo. Com isso temos dominância de segunda ordem e aplicamos o teorema de que, se São Paulo domina Curitiba em segunda ordem, a pobreza (medida com a classe P_α , com $\alpha \geq 2$) de São paulo será maior. Partimos agora para a análise dos resultados.

Agora analisando os resultados, vemos que de fato todos os indicadores da classe P_α mostram São Paulo como mais pobre, para as várias linhas de pobreza diferentes que foram testadas. Inclusive o indicador de Sen, que não tinha sido tratado em nossa análise estocástica reportou resultados similares. Especificamente para o salário mínimo, temos os resultados na tabela 4.

²Se cruzam para uma renda de, aproximadamente, R\$9360,00

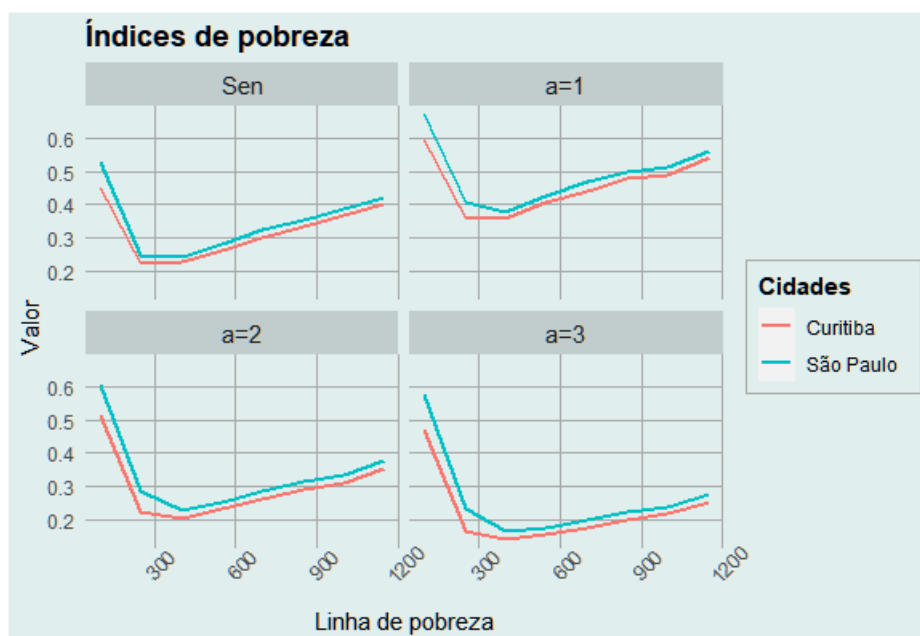
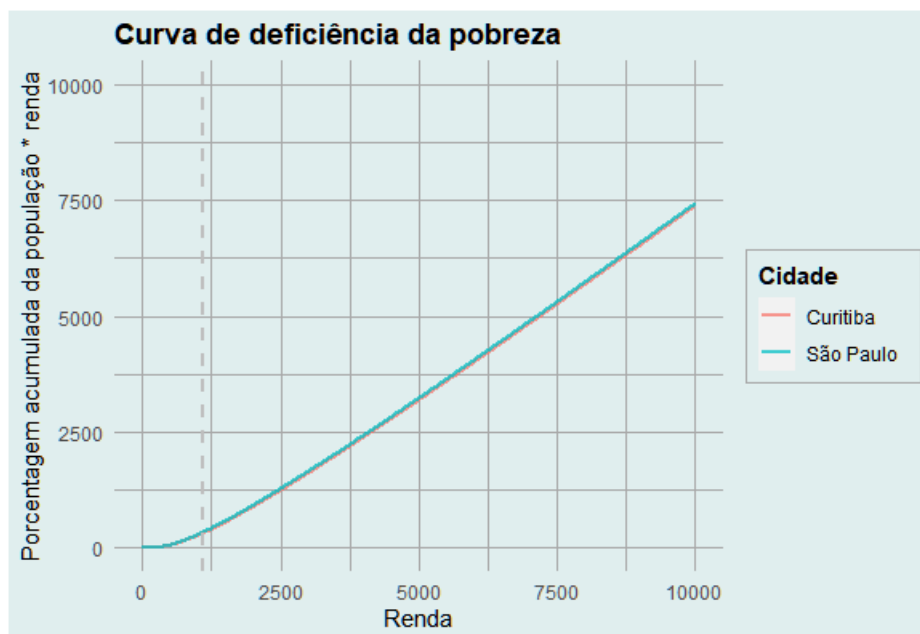


Tabela 4: Medidas de pobreza com salário mínimo

	Sen	$P_{\alpha=1}$	$P_{\alpha=2}$	$P_{\alpha=3}$
São Paulo	0.4039	0.5459	0.3625	0.2633
Curitiba	0.3856	0.5269	0.3404	0.2416

Comparando bem-estar

Em último lugar, temos a comparação das distribuições de Curitiba e São Paulo em relação ao bem-estar. Quando pensamos nesse tipo de análise, logo partimos para o uso das utilidades individuais. Entretanto, no mundo real, é complicado obter dados sobre as preferências das pessoas, principalmente partindo de distribuições de renda. Felizmente, existem funções de bem-estar baseadas na rendas individuais, as chamadas *nonwelfaristic*.

Para conseguirmos responder a pergunta de qual das duas cidades em análise possui a distribuição de renda apresenta maior bem-estar, usar as funções *nonwelfaristic* não é o suficiente. Primeiramente, caso as médias das distribuições fossem iguais, poderíamos usar os conceitos de dominância de Lorenz para a comparação desejada. Todavia, a partir da Tabela 1, é possível observar que a média das rendas em São Paulo é igual a R\$1.802,95, enquanto a de Curitiba é R\$1.663,41, ou seja, são diferentes, portanto não é possível utilizar esse método. Assim sendo, podemos usar os teoremas de dominância de bem-estar como um método alternativo.

Dadas as distribuições de renda de ambas as cidades, o teorema de dominância de primeira ordem define

$$X(p) = \inf x | F(x) \geq p \quad (1)$$

Em que X representa a distribuição, x uma renda individual, p percentil da população, tal que $p \in [0, 1]$, e F a função distribuição acumulada. Construindo os respectivos gráficos para São Paulo e Curitiba obtemos a Figura 1. Apesar de limitarmos a renda máxima para R\$50,000.00, a fim de facilitar a observação dos gráficos, ainda é possível concluir que as curvas se cruzam. Essa fato além de tornar a análise ambígua, já que dependendo da referência utilizada alguns podem dizer que o bem-estar em São Paulo é maior que em Curitiba e outros podem dizer o contrário, quebra uma das exigência da dominância de primeira: para que uma distribuição domine em primeira ordem outra, temos que todas os valores de uma curva tem ser maiores ou iguais da segunda.

Devido ao resultado inconclusivo do primeiro teorema o segundo pode nos ajudar a obter uma resposta. Esse define a curva de Lorenz Generalizada, que corresponde à curva de Lorenz de uma distribuição de renda multiplicada pela respectiva média, isto é,

$$GL_{X(p)} = \mu_X \cdot L_{X(p)} \quad (2)$$

Aplicando esse conceito para as distribuições de São Paulo e Curitiba obtemos a Figura 2. Assim como na dominância de primeira ordem, para a dominância de segunda ordem ser conclusiva precisamos que a curva de Lorenz Generalizada de uma distribuição esteja acima e no máximo coincidindo com a outra. Entretanto, isso não é observado dado que elas se cruzam nos níveis mais altos de renda. Portanto, não é possível fornecer uma resposta concreta em relação ao bem-estar das distribuições.

Embora não seja possível obtermos uma resposta com certeza sobre o bem-estar, dando enfoque à parcela não rica da população, dado que as curvas se cruzam no valor de p corres-

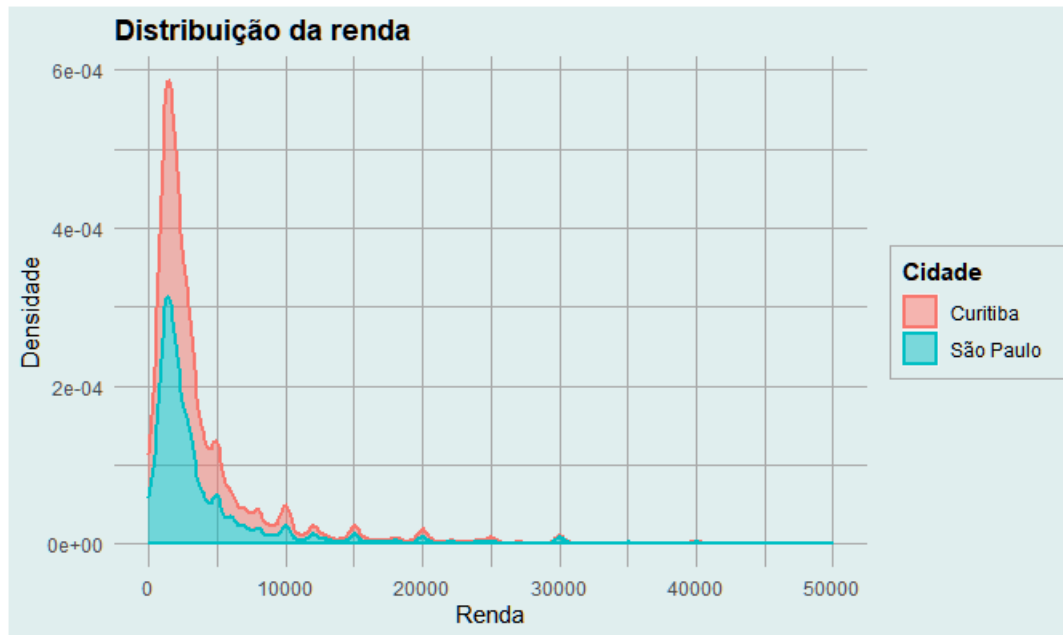


Figura 1: O eixo x foi limitado por R\$50,000.00

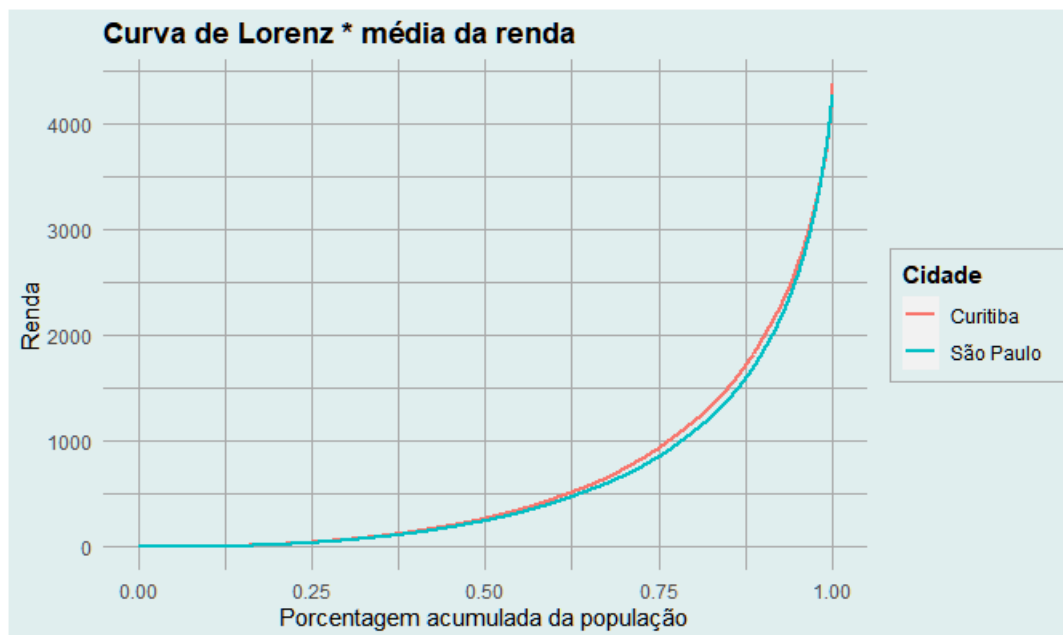


Figura 2: As curvas se cruzam em $x = 0.96$

pondente às pessoas que têm renda bem acima da média, poderíamos afirmar que o bem-estar em relação à distribuição em Curitiba é maior que em São Paulo, já que a curva de Lorenz Generalizada da primeira está acima da segunda, excerto quando nos aproximamos de $p = 1$. Essa referência levada em consideração é plausível em nosso ponto de vista por dois motivos: primeiro, em nossas análises sempre estamos mais interessados com a população mais pobre; e, segundo, alterar em um valor pequeno a renda de um indivíduo rico não afeta o bem-estar individual tanto quanto se alterássemos a renda de alguém pobre, já que 10 reais valem muito mais para o segundo do que para o primeiro.

Em conclusão, usando a teoria disponível de comparação de bem-estar não conseguimos escolher entre São Paulo e Curitiba, mas dando foco àqueles não tão favorecidos, Curitiba estaria em uma melhor posição.

Referências

FIELDS, G. S. *Distribution and Development: a New Look at the Developing World*. [s.n.], 2015. OCLC: 1227492700. ISBN 978-0-262-27260-5. Disponível em: <https://doi.org/10.7551/mitpress/2465.001.0001i>.