Pareto's Doubtful 80/20 Rule

William H. Press
Oden Institute of Computational Engineering and Science
The University of Texas at Austin

February 19, 2025

In 1897 (except that some say 1907) the Italian economist and polymath Vilfredo Pareto observed that 80% of the land in England (some say that he said Italy) was owned by 20% of the people—and that imbalances of this nature (some say these particular proportions) embody a kind of universality.

Clearer in the historical record is the role played by management consultant Joseph M. Juran, who as early as 1941 and continuing through the 1950s popularized (or invented) the idea of a universal 80/20 "Pareto principle."[1, 2]. It is hard to find any modern reference to the principle that does not run back through Juran. From so obscure a seed, the 80/20 principle has blossomed into a a verdant cultural meme, invoked everywhere:

- 80% of the work accomplished in a typical office is done by 20% of the people.
- 80% of a company's sales come from 20% of its products.
- 80% of healthcare resources are expended on 20% of the patients.
- 80% of your success will come from the best 20% of your ideas.

Are these all true? Who knows! The Pareto principle has spawned multiple self-help books [3, 4, 5]. And more: New York City debt-collection defense attorney Jesse Langel ("We Defend Against Debt Lawsuits, Bank Restraints, and Wage Garnishments") highlights 100 examples of the principle on the firm's web site [6]. One puzzler there is

- 80% of people marry within 20% of the local population,

seemingly implying that each individual in the desirable 20% has four spouses.

Evidently not all invocations of the principle can be fact. Many seem not even "truthy", much less true. Still, we all observe situations in which some people work four times as hard as others, or are four times needier. So perhaps we should not dismiss the principle completely.

But, is it four?

Where did "four times" come from? Well, 80 is four times 20, but let's look more closely. Suppose an office has a total productivity of 1 (in some units), and is populated by a fraction f of ambitious workers with productivity x_{high} and a fraction $1-f$ of lazy workers with productivity x_{low}. We can write,

$$\begin{aligned} f\,x_{\text{high}} + (1-f)\,x_{\text{low}} &= 1 \\ f\,x_{\text{high}} &= (1-f) \end{aligned} \tag{1}$$

whose solution is

$$x_{\text{high}} = \frac{1-f}{f}, \quad x_{\text{low}} = \frac{f}{1-f} \tag{2}$$

Substituting $f = 0.2$ (i.e., 20%), we get $x_{\text{high}} = 4$ (there's our "four"), but $x_{\text{low}} = 1/4$. In other words the hard workers have to work 16 times as hard as the slackers. If the hard workers are putting in an eight-hour day, then the rest of us are doing about half an hour of work spread throughout the day and spending the rest of our time on, I guess, our social media. Or polishing our résumés. I'm sure there are some offices like this, but it seems doubtful that this is the typical case.

The discrete charm of the false-unity rule

If exactly 20% of people are something, then it stands to reason that the remaining exactly 80% are something else, because $0.8 + 0.2$ must add to 1.0. Not so fast! In the 80/20 rule, the "20%" refers to people or "input", while the "80%" refers to production or "output". There is no mathematical reason that the numbers must add to one. It could be that 20% of the people produce 63.24% of the production, so that the remaining 80% (of people) produce the remaining 36.76% (of output). What must add to one is people-and-people or production-and-production, but not people-and-production. In this particular case we would have an 63.24/20 rule, not likely to have gone viral as a meme.

It is a matter of charm, not mathematics, that the 80/20 rule seems to add up to one. A 90/10 rule would be equally charming—if there were any reason for it to be true. We will call these "false-unity rules". Below, we answer the question of whether, in any particular case, some false-unity rule always exists and, if so, is it always something close to 80/20. (No spoilers here.)

Distribution functions of productivity

Instead of just two values of productivity x, it is more natural to take a distribution of values $p(x)$, the probability density of workers' productivity and thus normalized,

$$\int_0^\infty p(x)\,dx = 1 \tag{3}$$

We might just as easily have taken the lower limit to be $-\infty$, to include workers whose presence subtracts from overall productivity. We all know people like that. But, as a courtesy, we'll assume that none such exist. (Or they have been fired.)

For convenience, we can normalize the total productivity to 1 (as was done in equation 1), expressed as,

$$\int_0^\infty x\,p(x)\,dx = 1 \tag{4}$$

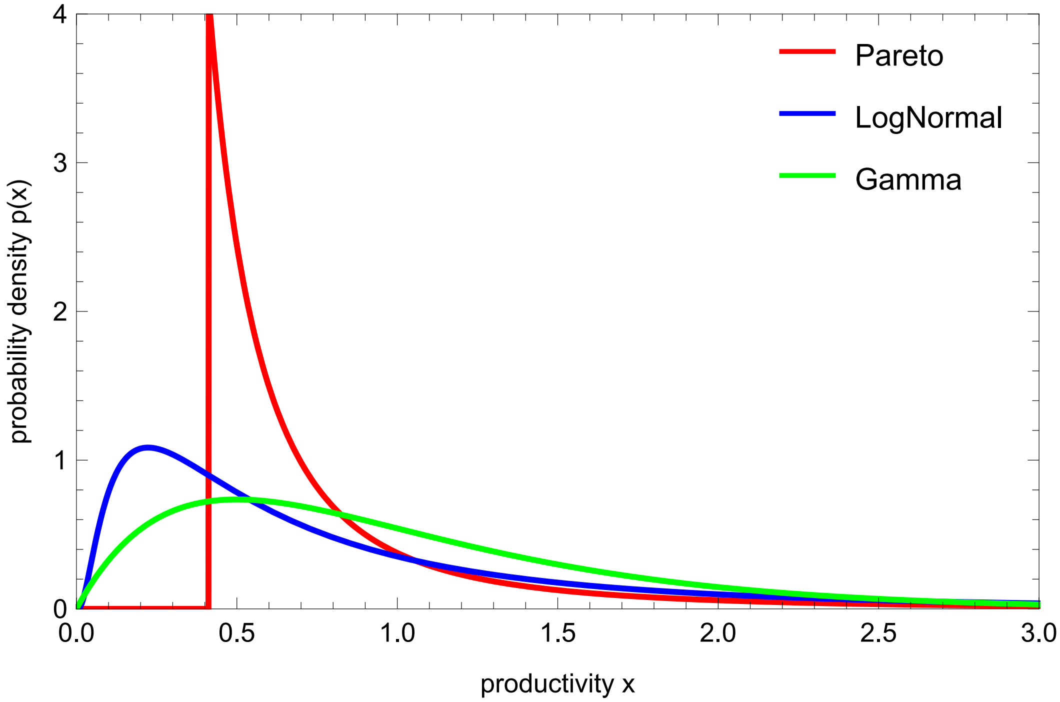

Figure 1: Examples of some standard distributions over the positive real line with means equal to 1.

This can also be interpreted as fixing the mean of the distribution to 1, meaning that any investigation of the 80/20 principle is an investigation of unit-mean distributions.

The distribution $p(x)$ might have any form, but we consider first some standard, "named" distributions over the positive real numbers: Pareto, LogNormal, and Gamma. Figure 1 shows one example of each type—not necessarily, as we will see, an example that satisfies the 80/20 rule.

Pareto

The eponymous Pareto has his own distribution [7], often referred to simply as a power-law or power-law-tail distribution:

$$p(x) = \begin{cases} 0, & x < x_0 \\ \dfrac{\alpha\, x_0^\alpha}{x^{\alpha+1}}, & x \geq x_0 \end{cases} \tag{5}$$

with $x_0 > 0$ and (for our purposes) $\alpha > 1$. Here α is a shape parameter, while x_0 adjusts the scale of x. The requirement of unit mean imposes a particular scale parameter value,

$$x_0 = \frac{\alpha - 1}{\alpha}, \tag{6}$$

yielding a one-parameter family of distributions with parameter α.

LogNormal

This distribution is Normal in the variable $\log(x)$. In the variable x it is

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left(-\frac{1}{2}\left[\frac{\log(x/x_0)}{\sigma}\right]^2\right) \tag{7}$$

As above, the unit-mean condition imposes a particular scale, now,

$$x_0 = \exp(-\sigma^2/2) \tag{8}$$

leaving σ as the free shape parameter.

Gamma

The Gamma distribution can be written as

$$p(x) = \frac{x_0^{-\alpha}}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-x/x_0} \tag{9}$$

(Most references write β for our x_0, but some confusingly write $1/\beta$ for our x_0, so we will stick with x_0.) The equation describes a distribution that grows as a power of x and then decays as a negative exponential.

The unit-mean condition is simply

$$x_0 = 1/\alpha \tag{10}$$

ROC curves

Suppose we write $C(x)$ for the fraction of people (or inputs) with productivity $\geq x$, and $E(x)$ for the fraction of total productivity that they produce. (The mnemonics are C for "cause" and E for "effect", with upper-case letters because these are cumulative distribution functions.) Then, with reference to the normalization equations (3) and (4), we have

$$\begin{aligned} C(x) &= \int_x^\infty p(x')\,dx' \\ E(x) &= \int_x^\infty x'\,p(x')\,dx' \end{aligned} \tag{11}$$

Equation (11) can be viewed as a parametric form for the function $E(C)$, that is, what fraction of output (effect) E is produced by the most productive fraction of input (cause) C. A plot of this is conventionally termed a ROC (receiver operating characteristic) curve, named for a quite different use of the same idea in World War II [8].

Figure 2 shows the ROC curves thus obtained for members of the Pareto (power-law) distribution family (a range of values of α) and members of the LogNormal distribution family (a range of values of σ). The figure for the Gamma distribution family (not shown) is qualitatively similar.

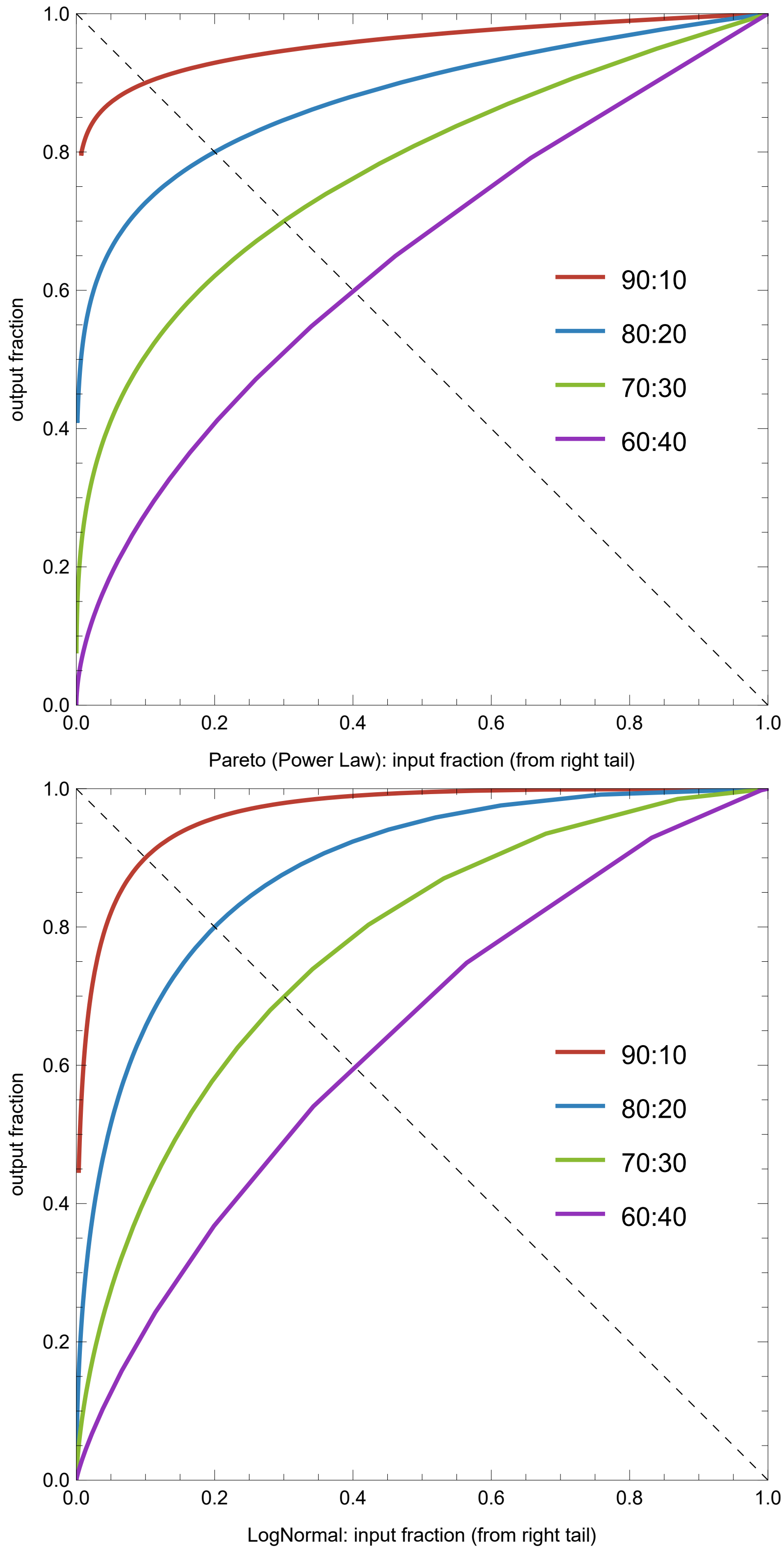

Figure 2: ROC curves for (top) Pareto distributions with different values of the tail exponent α, and (bottom) for LogNormal distributions with different values of the width parameter σ. See text for how the values of α and σ have been chosen.

Some general observations can be made:

1, The curves are monotonic, as follows from the non-negativity of the integrands in equation (11).

2. Since all the curves start at $(0, 0)$ and end at $(1, 1)$, and are monotonic and continuous, they must cross the upper-left-to-lower-right (that is, $(0, 1)$ to $(1, 0)$ line exactly once. That point is the distribution's unique false-unity rule, where some particular fraction f input produces a fraction $1-f$ of the output. But f is not necessarily 20% or any other universal value.

3. The ROC curves are always concave downward, as seen by (from equation 11),

$$\begin{aligned}\frac{d^2E}{dC^2} &= \frac{d}{dx}\left(\frac{dE}{dC}\right)\Bigg/\frac{dC}{dx} = \frac{d}{dx}\left(\frac{dE/dx}{dC/dx}\right)\Bigg/\frac{dC}{dx} \\ &= \frac{d}{dx}\left(\frac{-x\,p(x)}{-p(x)}\right)\Big/[-p(x)] = -\frac{1}{p(x)} \le 0\end{aligned} \tag{12}$$

4. The 50/50 rule ROC curve (not shown in the figures) is a straight line connecting $(0, 0)$ and $(1, 1)$, because that is the only (edge-case) concave-downward curve that also passes through $(0.5, 0.5)$. And, this is possible only in the limiting case where all workers have identical productivity.

5. No distribution ever has $f < 0.5$ in its unique false-unity rule, by a similar convexity argument. (If the best 60% people can do only 40% of the work, then the remaining lazy 40% fraction of people can't possibly do 60% of the work.)

The magical shape parameters to give 80/20

The magic in Figure 2 lies in choosing the specific values of the shape parameters α (top figure) or σ (bottom figure) that give memorable false-unity rules: 90/10, 80/20, 70/30, and 60/40; rather than something like 73.24/26.76. For any $p(x; \alpha)$, where α is a shape parameter, one can proceed by nested numerical root finding: An inner function, given α, finds x such that $E(x) = 1 - C(x)$ (equation 11) and returns the value $E(x)$. (Observation 2. above shows that there is always exactly one solution.) An outer function finds α to give the desired rule $E(x) = f = 0.2$ (say). For the case of the Gamma distribution, equation (9), we do exactly this. The value $\alpha = 0.244829$ there yields the 80/20 rule.

Of no fundamental importance, but fun, the Pareto and LogNormal distributions yield analytic solutions. It is convenient in both cases to let the scale parameter x_0 "float", because the answer will depend only on shape, not scale. We then have two simultaneous equations,

$$C(x) = f, \qquad E(x) = 1 - f \tag{13}$$

to solve for the two unknowns, shape parameter and x. In the case of Pareto, equations (13) are

$$\left(\frac{x_0}{x}\right)^{\alpha} = f, \qquad \left(\frac{x_0}{x}\right)^{\alpha-1} = 1 - f \tag{14}$$

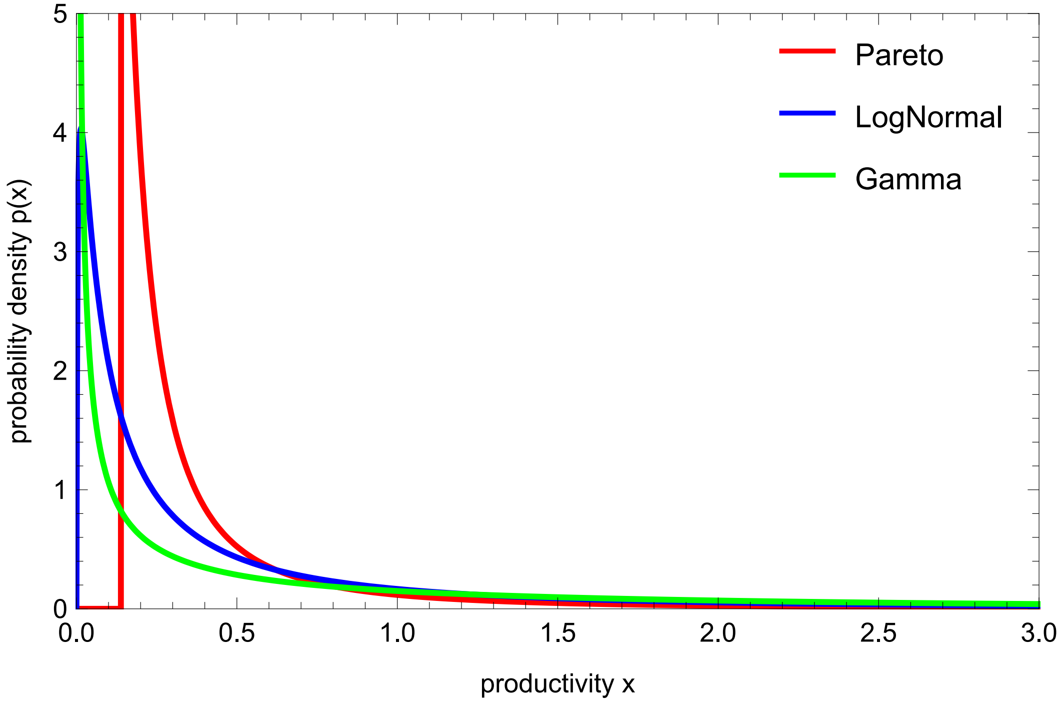

Figure 3: Same families of distributions as was shown in Figure 1, but now setting their shape parameters to have 80/20 rules. The tails extend far to the right of the figure. (The distributions shown have means equal to 1.)

whose solution is easily found to be

$$\alpha = \frac{\log(f)}{\log(f) - \log(1-f)} \tag{15}$$

The result $x = (1-f)x_0/f$ is also found, but is not needed for anything. The 80/20 rule for the Pareto distribution thus has exponent $\alpha = \log 5 / \log 4 = 1.16096$.

For LogNormal, with some help from Wolfram Mathematica [9], we find the equations,

$$\frac{1}{2}\,\mathrm{erfc}\left(\frac{\log x}{\sqrt{2}\sigma}\right) = f, \qquad \frac{1}{2}\,\mathrm{erfc}\left(\frac{\log x - \sigma^2}{\sqrt{2}\sigma}\right) = 1 - f \tag{16}$$

where erfc is the complementary error function. The solution is, rather amazingly,

$$\sigma = 2\sqrt{2}\,\mathrm{erf}^{-1}(1-2f) \tag{17}$$

where erf^{-1} is the inverse error function. The 80/20 rule for the LogNormal distribution then has shape parameter $\sigma = 1.68324$.

All 80/20 distributions are very long-tailed

Figure 1 showed textbook-typical Pareto, LogNormal, and Gamma distributions. Figure 3 now shows the unique members of those families that yield 80/20 rules. They are nothing like typical! Indeed, the figure can hardly do them justice: The functions shown all have mean 1, but by eye, the preponderance of their weight seems to be at values smaller than 1. The explanation, of course, is that their tails extend far to the right and decay only slowly.

We can prove this as a general property of any distribution $p(x)$ with an 80/20 rule. In fact, we already almost did prove this in equations (1) and (2). The revised argument is this: We are given an arbitrary $p(x)$, scaled and normalized to satisfy equations (3) and (4). For a fixed f, define x_m as the productivity value such that

$$\int_{x_m}^{\infty} p(x)\,dx = f, \quad 0 < f < 0.5 \tag{18}$$

Let $\langle x_{\text{hi}} \rangle$ be the average productivity x of all workers with productivity $x > x_m$, that is,

$$\langle x_{\text{hi}} \rangle = \int_{x_m}^{\infty} xp(x)\,dx \Bigg/ \int_{x_m}^{\infty} p(x)\,dx = \int_{x_m}^{\infty} xp(x)\,dx \Big/ f \tag{19}$$

Similarly $\langle x_{\text{lo}} \rangle$ is the average of workers with $x < x_m$,

$$\langle x_{\text{hi}} \rangle = \int_{0}^{x_m} xp(x)\,dx \Big/ (1-f) \tag{20}$$

The equation

$$f\ \langle x_{\text{hi}} \rangle + (1-f)\ \langle x_{\text{lo}} \rangle = 1 \tag{21}$$

is immediately seen to be the normalization equation (4). Since $f\ \langle x_{\text{hi}} \rangle$ is the total productivity $> x_m$ (that is, fraction of workers times their average productivity), the condition for a false-unity rule is

$$f\ \langle x_{\text{hi}} \rangle = 1 - f \tag{22}$$

equations (21) and (22) are exactly the same as equation (1), but with mean quantities now replacing assumed point values. The solution for the means is thus the analog of equation (2),

$$\langle x_{\text{hi}} \rangle = \frac{1-f}{f}, \quad \langle x_{\text{lo}} \rangle = \frac{f}{1-f} \tag{23}$$

Substituting $f = 0.2$ to get an 80/20 rule, we see that that for any functional form of $p(x)$, the average productivity of the top 20% is exactly 16 times the average productivity of the bottom 80%. It follows that, for smooth-tailed distributions (i.e., not magically massed around this unique multiple 16), a significant fraction of the over-achieving 20% must lie at $\sim$ 20 or $\sim$ 30 times the mean of the other 80%. Such hyper-skewed distributions are a fact of life in some realms, such as income inequality [10] or healthcare costs [11]. It seems unlikely that they are typical of worker productivity in the average office. The reader can now judge the plausibility of other claimed 80/20 applications case by case. What about this one:

- 80% of *true* 80/20 assertions come from 20% of the websites making them.

References

[1] J. A. DeFeo and J. M. Juran, *Juran's Quality Handbook*, 7th ed. McGraw Hill, July 11 2023.

[2] I. Juran Global, "Pareto principle (80/20 rule) & Pareto analysis guide," n.d., last accessed: 2024-12-29. [Online]. Available: https://www.juran.com/blog/a-guide-to-the-pareto-principle-80-20-rule-pareto-analysis/

[3] R. Koch, *The 80/20 Principle: The Secret of Achieving More With Less.* New York: Bantam Doubleday Dell Publishing, 1998.

[4] Z. Lawson, *Pareto Principle: Unleash the True Power of the Pareto Principle (The Secret Strategy to Optimizing Every Area of Your Life).* Zoe Lawson, 2022.

[5] A. E. Oberdorfer, *Why Less Is More: The Pareto Principle in a Nutshell: Boost Your Productivity by Working Less.* Independently published, May 20 2021.

[6] L. Jesse Langel, Esq., "100 short examples of Pareto's 80/20 rule with 35 key insights," August 23 2018, last accessed: 2024-12-29. [Online]. Available: https://www.thelangelfirm.com/debt-collection-defense-blog/2018/august/100-examples-of-the-80-20-rule/

[7] Wikipedia contributors, "Pareto distribution," https://en.wikipedia.org/wiki/Pareto_distribution, n.d., accessed: 2024-12-29.

[8] ——, "Receiver operating characteristic," https://en.wikipedia.org/wiki/Receiver_operating_characteristic, n.d., accessed: 2024-12-29.

[9] Wolfram Research, Inc., *Wolfram (Mathematica), Version 14.1*, Wolfram Research, Inc., Champaign, IL, 2024. [Online]. Available: https://www.wolfram.com/mathematica/

[10] T. Piketty and E. Saez, "Income inequality in the United States, 1913–1998," *Quarterly Journal of Economics*, vol. 118, no. 1, pp. 1–41, 2003.

[11] D. Pritchard, A. Petrilla, S. Hallinan, D. H. Taylor Jr, V. F. Schabert, and R. W. Dubois, "What contributes most to high health care costs? Health care spending in high resource patients," *Journal of Managed Care & Specialty Pharmacy*, vol. 22, no. 2, p. 102, February 2016.