The Distribution of Household Net Worth in 2022

Joseph Nathan Cohen
Department of Sociology
City University of New York
March 7, 2024

Sociologists routinely discuss economic inequality in abstract terms, such as the "rich," "poor," "working class," "middle class," and so forth. Concretely, what do the finances of these groups look like? Who counts as rich or poor? How many people are actually broke? How much money do rich people have? Pursuing these questions empirically allows us to move past these two- or three-category typologies and towards a richer understanding of how economic resources and opportunities are distributed across society.

In this note, I use data from the most recent *Survey of Consumer Finances* (US Federal Reserve 2023) to explore the univariate distribution of household net worth (or "wealth") across the entire United States. *Net worth* is the monetary value of a household's assets less the value of its outstanding debts. A household's net worth is the accumulated economic value of their economic property. The figure signals a household's ability to marshal money to meet needs or wants.

Though we are all cognizant of the fact that America has rich and poor families (i.e., high- or low-wealth households), we often lack a concrete sense of how much money "rich" or "poor" people have, how many "rich" or "poor" people exist, and where to place situate people (including ourselves) in society's economic hierarchy (Cohen 2017). Empirical estimates of household net worth's univariate distribution can be helpful in these situations because they push us to develop tangible, well-defined, and non-imagined ideas about who has how much in contemporary America.

In addition to this analysis, readers may also wish to consult Aladangady *et al.*'s (2023) estimates published by the Federal Reserve. Readers may notice minor discrepancies between our estimates (on the order of fractions of a percent) and can consult this post to see my attempts to examine, test, and ultimately agree on the rigor of officially sanctioned This analysis replicates Aladangady *et al.*'s findings using the public-use microdata set, and uses these findings to offer a concrete, data-driven look at the distribution of wealth (and a dimension of general economic inequality) in contemporary America.

BASIC CONCEPTS

Economic inequality is very complicated. It manifests differently (e.g., income, wealth, opportunity). Economic outcomes and overall well-being often vary across geography and demographic groups, and positive economic outcomes do not automatically translate into better lives. Empirical social scientists deal with these complexities by developing concepts that describe one dimension of a complex idea like inequality, and then using empirical data analysis to establish a sense of how people's fortunes align with that dimension in question.

This analysis focuses on household net worth. As mentioned at the outset, *net worth* (or "wealth," used synonymously here) is the monetary value of a household's owned property, less their debts. A household's owned property is often referred to as its *assets* and typically includes things like bank

accounts, retirement savings, cars, homes, and other property that could be sold for money. A household's total assets represent an estimate of how much money it could acquire were it to sell everything it owns.

Debts are the amounts of money that a household would need to settle all outstanding obligations. We incorporate consideration of debts to differentiate those who have acquired assets through heavy borrowing from those that own much and owe little. It allows us to differentiate someone who owns a $1 million home with a $900,000 mortgage *versus* someone who owns that home outright. When all the economic dust is settled in a liquidation and debt settlement, we'd expect the former to be left with $100,000 in the bank and the latter with a million.

The resulting metric is measured empirically as assets minus debts, but what is the meaning of *net worth* in non-accounting terms? I think of it as a rough estimate of a household's ability to command money if needed. The estimates are rough in the sense that asset values are only estimates of how much money people would get, and people's ability to avoid (or avoid the ramifications of) debt default can also differ. As we will see below, it may be splitting hairs to dwell on the wellbeing differences of someone worth $50,000 *versus* $200,000, or $1 million *versus* $3 million, and other pairs of wealth scores that register near each other on the larger scale of America's net worth distribution. The empirical distribution looks less like an upper-, middle-, and lower-class system and more like what is commonly described as a "long right-tailed" distribution.

We assess wealth at the level of households. A *household* is a group of people who form a standard economic unit that sustains the financial livelihood of its members. We generally judge the economic conditions of households rather than individuals to account for the fact that many individuals earn no money. For example, if we judge all children on their wealth, they would almost all register as zero and be categorized as personally poor. However, people who live in the same household tend to share a common living standard, with livelihoods that avail themselves of communal household assets that are within the capacity of the household heads to acquire. Thus, a child of wealthy parents is expected to enjoy a different livelihood and level of well-being than a child of poor parents, even though both children have similar personal financial situations.

In discussions related to wealth, there is often a frame of "rich" and "poor" and much uncertainty over where to draw the line between either and "regular" folk. I have engaged this issue in my past analysis of this data (Cohen 2017). I sense that much of this boundary-drawing involves social signaling and people's dispositions to see themselves as typical. My preference is to think of the distribution of wealth in concrete quantitative terms and then move on to think about concrete ways in which these differences in financial situations could lead to concrete differences in well-being, livelihood, and opportunities. In this analysis, we will focus more squarely on describing the distribution of total wealth among U.S. households nationally.

DATA AND METHODS

Method

To describe the distribution of household net worth, one must encapsulate the economic circumstances of America's estimated 125 million-plus households. That is too much discrete information for the

human mind to process. We attempt to describe the distribution of their associated metrics to make these scores comprehensible. *Distributions* refer to the general contours of the scores registered along some metric. There are different ways to describe a distribution, both numerically and graphically. This analysis includes the following methods for describing the distribution of such a metric:

1. Bar Charts: A graphical depiction of the estimated distribution across different wealth levels.
2. Measures of Central Tendency: Scores that impart a sense of what is a normal or typical score. We will consider the mean and median.
3. Percentile Scores: Net worth thresholds beyond which households can be considered part of higher-wealth groups. These thresholds are based on the percentage of the population estimated to have less wealth than the household in question.
4. Frequency Tables: We will try to divide households up into meaningfully different levels of wealth and estimate how many households fall into each category.

```
# Generic session set up

# First, start your session with an empty memory
rm(list=ls())
gc()

# Set your working directory
directory <- "D:/Dropbox/Research/Household Finance/Net Worth Univariate"
setwd(directory)

# Load the libraries used in this session
library(httr)
library(jsonlite)
library(survey)
library(mitools)
library(haven)
library(scales)
library(ggplot2)
library(kableExtra)

# Set the random seed
set.seed(123)

# Turn off scientific notation
options(scipen=999)
```

Data

We use the net worth variable calculated in the summary data table of the *Survey of Consumer Finances* (SCF). In making parameter estimates, it is important to note that the Survey employs a stratified and clustered sample for which your analysis must account when making parameter estimates. The SCF's public-use data uses replicate weights to correct parameter estimates. My analysis used Lumley's (2011) 'survey' package.

The SCF also employs multiple imputations with randomness to estimate missing data. The analysis of such data requires special handling for both complex survey analysis and missing data imputation. This post walks the user through generating such estimates using *Survey of Consumer Finance* data. Readers can see my R code in the folds of this post or download a copy of this analysis' Markdown file from my OSF repository. Click here for a PDF version with code and without code.

```
# PART ONE: DOWNLOAD DATA
# Download Main 2022 Files
response_0 <- GET("https://www.federalreserve.gov/econres/files/scf2022s.zip",
      write_disk("scf2022s.zip", overwrite = TRUE))
response_1 <- GET("https://www.federalreserve.gov/econres/files/scfp2022s.zip",
      write_disk("scfp2022s.zip", overwrite = TRUE))
response_2 <- GET("https://www.federalreserve.gov/econres/files/scf2022rw1s.zip",
      write_disk("scf2022rw.zip", overwrite = TRUE))

rm(list=ls(pattern = "response")) # Clean up objects

# PART TWO: UNZIP DATA
unzip("scf2022s.zip")
unzip("scfp2022s.zip")
unzip("scf2022rw.zip")

file.remove("scf2022s.zip")  # Erase the zip files because I don't need them and they
 take up space.
file.remove("scfp2022s.zip")
file.remove("scf2022rw.zip")

# PART THREE:  CONVERT DATA TO R FORMAT
# The data are distributed in Stata format.  Below, I import data from its Stata form
at.
scf2022 <- read_dta("p22i6.dta")
scf2022s <- read_dta("rscfp2022.dta")
scf2022rw <- read_dta("p22_rw1.dta")

# PART FOUR: DATA QUALITY CHECK
# Data quality check: Ensure all three files have corresponding rows
stopifnot( nrow( scf2022 ) == nrow( scf2022rw ) * 5 )  # One RW score per household
stopifnot( nrow( scf2022 ) == nrow( scf2022s ) )
```

```r
#Confirm only the primary economic unit and the five implicate identifiers overlap:
stopifnot(all(sort(intersect(names(scf2022), names(scf2022s))) == c('y1', 'yy1')))
stopifnot(all(sort(intersect(names(scf2022), names(scf2022rw))) == c('y1', 'yy1')))
stopifnot(all(sort(intersect(names(scf2022s), names(scf2022rw))) == c('y1', 'yy1')))

# Convert column names to lower case in all sets, per Damico script
names(scf2022) <- tolower(names(scf2022))
names(scf2022rw) <- tolower(names(scf2022rw))
names(scf2022s) <- tolower(names(scf2022s))

# Per Damico script
# Remove implicate identifier from RW table, and add column of fives for weighting
scf2022rw[, 'y1'] <- NULL
scf2022[,'five'] <- 5

save(scf2022, scf2022s, scf2022rw, file = "SCF 2022 Raw Data Tables.RData")
# Save the data

# PART FIVE: MERGE MAIN AND SUMMARY DATA
# Merge Summary and Raw Data Tables by 'y1'
scf2022 <- merge(scf2022, scf2022s, by = "y1", sort = T)

# PART SIX: RECAST DATA AS FIVE SEPARATE IMPLICATES
# Splitting data set into five separate sets of individual implicates
scf_1 <- subset(scf2022,
       as.numeric(substr(scf2022$y1, nchar(scf2022$y1), nchar(scf2022$y1))) == 1)
scf_2 <- subset(scf2022,
       as.numeric(substr(scf2022$y1, nchar(scf2022$y1), nchar(scf2022$y1))) == 2)
scf_3 <- subset(scf2022,
       as.numeric(substr(scf2022$y1, nchar(scf2022$y1), nchar(scf2022$y1))) == 3)
scf_4 <- subset(scf2022,
       as.numeric(substr(scf2022$y1, nchar(scf2022$y1), nchar(scf2022$y1))) == 4)
scf_5 <- subset(scf2022,
       as.numeric(substr(scf2022$y1, nchar(scf2022$y1), nchar(scf2022$y1))) == 5)

# Clean Up Subject Identifier in Individual Implicates
for (i in 1:5){
  temp <- get(paste0("scf_", i))
  temp$yy1 <- temp$yy1.x
  temp$yy1.x <- NULL
  temp$yy1.y <- NULL
  assign(paste0("scf_", i), temp)
}

# Compile Individual Implicates to a List
scf_data_list <- list(scf_1, scf_2, scf_3, scf_4, scf_5)

# Removing unnecessary objects to save memory and space in Environment window
rm(scf_1, scf_2, scf_3, scf_4, scf_5)
rm(i, temp)
gc()
```

```r
# PART SIX: CLEANING DATA
# Replace missing replicate weights with zeros to prevent downstream bugs
scf2022rw[ is.na( scf2022rw ) ] <- 0

# Rescale weights, per documentation
scf2022rw[ , paste0( 'wgt' , 1:999 ) ] <-
    scf2022rw[ , paste0( 'wt1b' , 1:999 ) ] * scf2022rw[ , paste0( 'mm' , 1:999 ) ]

# Using Damico's strategy of storing as a data table with y1 and wgts*
scf2022rw <- scf2022rw[ , c( 'yy1' , paste0( 'wgt' , 1:999 ) ) ]

# Check if yy1 values match across the datasets and RW table
all(scf_data_list[[1]]$yy1 == scf2022rw$yy1)
all(scf_data_list[[2]]$yy1 == scf2022rw$yy1)
all(scf_data_list[[3]]$yy1 == scf2022rw$yy1)
all(scf_data_list[[4]]$yy1 == scf2022rw$yy1)
all(scf_data_list[[5]]$yy1 == scf2022rw$yy1)

# Survey Design Object following Damico
scf_design <-
  svrepdesign(
    weights = ~wgt,
    repweights = scf2022rw[ , -1 ] ,
        data = imputationList( scf_data_list ) ,
        scale = 1 ,
        rscales = rep( 1 / 998 , 999 ) ,
        mse = FALSE ,
        type = "other" ,
        combined.weights = TRUE
  )

scf_design_single <-
  svrepdesign(
    weights = ~wgt,
    repweights = scf2022rw[ , -1 ] ,
        data = scf_data_list[[1]] ,
        scale = 1 ,
        rscales = rep( 1 / 998 , 999 ) ,
        mse = FALSE ,
        type = "other" ,
        combined.weights = TRUE
  )
```

```
# Damico's function to combine implicates to give summary estimate from each of the f
ive sets.
scf_MIcombine <-
    function (results, variances, call = sys.call(), df.complete = Inf, ...) {
        m <- length(results)
        oldcall <- attr(results, "call")
        if (missing(variances)) {
            variances <- suppressWarnings(lapply(results, vcov))
            results <- lapply(results, coef)
        }
        vbar <- variances[[1]]
        cbar <- results[[1]]
        for (i in 2:m) {
            cbar <- cbar + results[[i]]
            # MODIFICATION:
            # vbar <- vbar + variances[[i]]
        }
        cbar <- cbar/m
        # MODIFICATION:
        # vbar <- vbar/m
        evar <- var(do.call("rbind", results))
        r <- (1 + 1/m) * evar/vbar
        df <- (m - 1) * (1 + 1/r)^2
        if (is.matrix(df)) df <- diag(df)
        if (is.finite(df.complete)) {
            dfobs <- ((df.complete + 1)/(df.complete + 3)) * df.complete *
            vbar/(vbar + evar)
            if (is.matrix(dfobs)) dfobs <- diag(dfobs)
            df <- 1/(1/dfobs + 1/df)
        }
        if (is.matrix(r)) r <- diag(r)
        rval <- list(coefficients = cbar, variance = vbar + evar *
        (m + 1)/m, call = c(oldcall, call), nimp = m, df = df,
        missinfo = (r + 2/(df + 3))/(r + 1))
        class(rval) <- "MIresult"
        rval
    }

save(scf_design, scf_design_single, scf_data_list, scf_MIcombine,
       file = "SCF01data.RData")
```

NATIONAL DISTRIBUTION OF HOUSEHOLD WEALTH

Household wealth varies greatly when looking at the country as a whole. These distributions will look very different where we focus on data from specific geographic regions or among different status groups. In this exercise, we are looking at comparative haves and have-nots in terms of monetary net worth on a national scale. They represent meaningful differences to the extent that, in the United States today, unequal access to money creates unequal access to essential goods and services, basic rights, and other facets of human well-being.

The Bankrupt. It is easy to quantify doing bad on this metric: there is a zero dollars or less floor. Approximately 8% of U.S. households had no wealth or a negative net worth.[1] This group seems unable to come up with much money if necessary. To the extent that money is required to access essential goods and services in a society, we would expect this group to forego more of such essentials. These households would require outside help from extended family, the state, or charities, to sustain high living standards.

This is my first look at the 2022 data, but I recall several observations made in past analyses (see Cohen 2017). First, this group is not uniformly poor. There are households with huge negative net worths, and the general contours of their balance sheets struck me as once high-wealth households (often businesspeople) whose asset values collapsed spectacularly. I was not entirely convinced that these people were now destitute. In any case, these are extreme rather than typical cases. A more common scenario might be a younger household with large education debts, a family that has been crushed by the costs of a medical problem, or a self-employed household whose business has failed (see Sullivan, Warren, and Westbrook 2001).

The Line Between Rich and Non-Rich. In my last engagement with this survey series, I wrestled with where to draw the line between rich and poor (Cohen 2017). As a mental exercise, I reverse-engineered an estimate by asking how much money would be needed to get a median, inflation-adjusted household income from an ultra-conservative investment portfolio, assuming that the general contours of inflation, the stock market, and the bond market held to their 100-year averages. The result was a wealth threshold of $1.4 million. That figure was received with credulity by people who argued that they could hardly live off that much money. However, their definition of a livelihood was not that of a person earning $47,000 per year household and renting their primary and only residence.

The Long Right Tail. Most households (about 86%) have some accumulated wealth but less than $3 million.[2] This still encompasses a very wide range of meaningfully different economic outcomes. Still, they are on the order of the differences that would exist in a local community between its wealth-holding upper-middle class and those with little savings. The $3 million upper limit is arbitrary. The insights conveyed in the findings below do not strongly depend on the upper limit. About 6.1% of households had more than $3 million in net worth. About 9.7% have more than $2 million. About 19% had more than $1 million. It is part of a pattern that is not faithfully imparted in a two- or three-category typology.

Figure 1 (below) depicts the national distribution of household wealth. The x-axis represents net worth categories between zero and $3 million in increments of $100,000:

```
results <- scf_MIcombine( with( scf_design , svymean( ~ networth_middle ) ) )
results <- data.frame(cbind(rownames(results$coefficients),
      results$coefficients))
names(results) <- paste("Percent")
results <- data.frame(categories = rownames(results),
```

[1] A household's net worth is considered to be negative when its debts exceed the value of its assets.

```
      percent = results$Percent)

desired_order <- c("networth_middleZero or Negative Net Worth",
                   paste0("networth_middle", 1:30),
                   "networth_middleOver $3 Million")
results <- results[match(desired_order, results$categories), ]
results$categories <- c("Negative", formatC(seq(0,2900000,100000),
      format = "f", big.mark = ",", digits = 0), "Over $3 Million")

# Convert 'categories' to a factor with levels in the order they appear in the data f
rame
results$categories <- factor(results$categories, levels = results$categories)

# Create the bar chart
fig1 <- # Create the bar chart with a title and a footnote
ggplot(results, aes(x = categories, y = percent)) +
  geom_bar(stat = "identity", width = 0.7) +
  scale_y_continuous(breaks = seq(0, max(results$percent, na.rm = TRUE),
      0.02), labels = scales::percent) +
  labs(x = "Net Worth Category", y = "Percent of US Households",
       title = "Distribution of U.S. Household Net Worth, 2022") +
  theme(axis.text.x = element_text(angle = 45, hjust = 1)) +
  annotate("text", x = Inf, y = -Inf, label = "Joseph Nathan Cohen, CUNY Queens Colle
ge", hjust = 1, vjust = -1, size = 3, colour = "gray40")
```

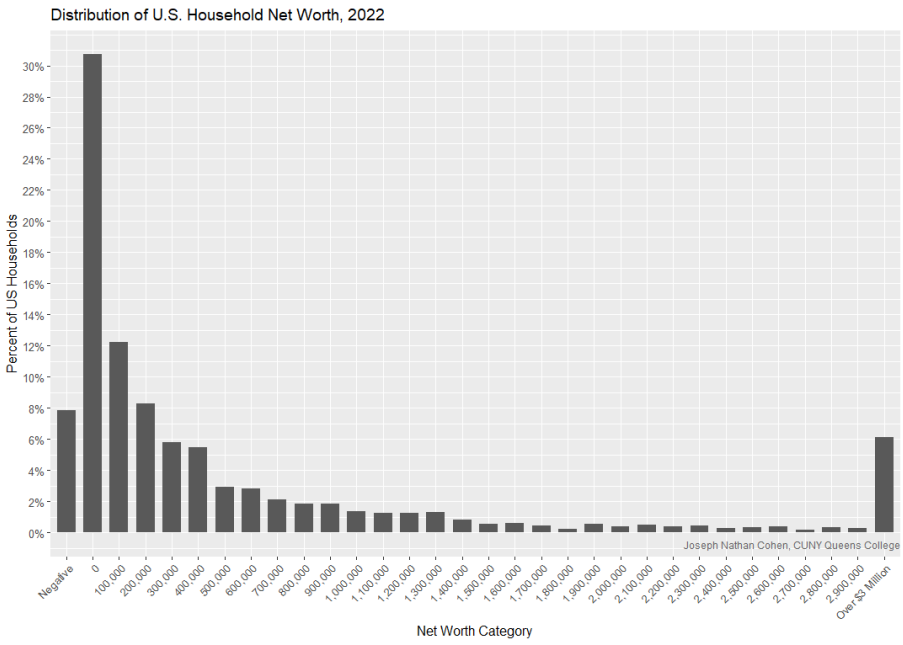

Figure 1: Distribution of U.S. Household Wealth, 2022

The distribution above is commonly described as having a "long right tail." This means that most scores register at the bottom of the distribution (here, people with say less than $100,000 or $200,000 in total net worth), with bigger scores getting rarer and rarer. We see that at least 1-in-100 households register in these categories up to that $1.4 million threshold (which has not been updated from 2017).

What is the implication of this long right tail, as distinct from the rich-not rich dichotomy? One problem is that we lose information when we force multiple shades of gray into a two-category system. The second is that it is hard to figure out where to draw concrete lines to sort people into these systems. If two married schoolteachers end their long careers with a combined $2 million in retirement savings, are they rich? Are they the same kind of rich, or are they in on the same game, as billionaires? Is the successful local businessperson – like the person who owns the locally-famous chain of restaurants or car dealerships in town – the same as the major industrialist with political lobbyists? It is possible to argue yes and no.

One way to think about the distribution of wealth is to think of there being a common circumstance, and departures from the norm that are rarer and rarer. We do not have to immediately assume that the world fits into neat categories, but rather that there are various degrees of departure from a baseline, "typical" outcome.

Where is the Bulk of Society? About 81% of households have less than $1 million in net worth. My sense of the overall economic situation of this group is similar in that they require some kind of socialized provision or subsidy of essential goods and services (e.g., education, healthcare, transportation, legal advice). This socialized provision can occur through the government and nongovernmental organization. The main point is that such households likely lack the means to personally finance what is widely considered to be "essential" goods and services today.

These households' fates are likely not tied permanently. It is comparatively uncommon for a household headed by someone in their twenties or thirties, when most households are at the start of their working lives. Some part of them will accumulate wealth over their careers and eventually find themselves in wealth-bearing classes. Many will not accumulate wealth. I'll leave concrete estimates for another analysis. My main point is that a group of low-wealth households will have some subset of young people who are on track to positions of wealth, with personal financials that do not yet reflect the returns to be gleaned from things like education, ability differentials, social and cultural capital, family help, and the many other factors that can help ensure the success of one's personal economic strategies. Monetizable wealth is an imperfect metric of overall economic advantage, and people with similar levels of wealth can be in dramatically different overall economic circumstances.

Different identity groups likely fare better in this marathon of wealth accumulation. In my analysis of the 2016 data, I found age to be the strongest point of division between low- and high-wealth households. Other predictors, in order of descending predicted impact: possession of a college education, unpaired status (i.e., no married or cohabiting), working status, race/ethnicity, and sex of household head. Things get more interesting as you delve into the weeds, but I will resist the temptation to answer everything in one analysis.

SITUATING A HOUSEHOLD IN THE RANKS OF HOUSEHOLD WEALTH

One way to understand where a household stands in the greater hierarchy of personal wealth is to consider their percentile score. This is the percentage of U.S. households who are expected to have less wealth than the household in question. The table below summarizes percentile wealth scores for U.S. households in 2022:

```r
p10 <- scf_MIcombine(with( scf_design ,
                           svyquantile(~networth, 0.1 , se = F)))$coefficients
p20 <- scf_MIcombine(with( scf_design ,
                           svyquantile(~networth, 0.2 , se = F)))$coefficients
p30 <- scf_MIcombine(with( scf_design ,
                           svyquantile(~networth, 0.3 , se = F)))$coefficients
p40 <- scf_MIcombine(with( scf_design ,
                           svyquantile(~networth, 0.4 , se = F)))$coefficients
p50 <- scf_MIcombine(with( scf_design ,
                           svyquantile(~networth, 0.5 , se = F)))$coefficients
p60 <- scf_MIcombine(with( scf_design ,
                           svyquantile(~networth, 0.6 , se = F)))$coefficients

p70 <- scf_MIcombine(with(scf_design,
                            svyquantile(~networth, 0.7 , se = F)))$coefficients
p80 <- scf_MIcombine(with(scf_design,
                           svyquantile(~networth, 0.8 , se = F)))$coefficients
p90 <- scf_MIcombine(with(scf_design,
                           svyquantile(~networth, 0.9 , se = F)))$coefficients
p95 <- scf_MIcombine(with(scf_design,
                           svyquantile(~ networth , 0.95 , se = F)))$coefficients
p99 <- scf_MIcombine(with(scf_design,
                           svyquantile(~ networth , 0.99 , se = F)))$coefficients
p999 <- scf_MIcombine(with(scf_design,
                  svyquantile(~ networth , 0.999 , se = F)))$coefficients

results <- data.frame(
  Percentile = c("10th", "20th", "30th", "40th", "50th", "60th", "70th", "80th", "90t
h", "95th", "99th", "99.9th"),
  Wealth = c(p10, p20, p30, p40, p50, p60, p70, p80, p90, p95, p99, p999))
results$Wealth <- round(results$Wealth, 0)
results$Wealth <- format(results$Wealth, big.mark = ",", decimal.mark = ".", nsmall =
 0, scientific = FALSE)

kable(results)
```

Table 1 (right) lists the estimated decile scores for household wealth in 2022. Some observations:

Table 1: Percentile Scores for U.S. Household Net Worth (2022)

Percentile	Wealth
10th	440
20th	13,528
30th	51,366
40th	110,314
50th	192,084
60th	312,622
70th	493,068
80th	891,750
90th	1,920,758
95th	3,779,600
99th	13,666,778
99.9th	61,827,166

First, about one-tenth of the country's households have less than $440 of wealth. Such a situation would leave a person at the edge of the monetized economy, and likely juggling a very tight cash flow. Those circumstances would place a household in a situation of a cash shortage in the event of any commonplace financially-costly emergency, like a car repair or a doctor's bill. Though we are discussing the 10th percentile score as the lowest level in our scheme, it is somewhat remarkable to take mental account of the fact that at least 10% of society's households are in that kind of situation.

Without looking at the data, I imagine that the next decile includes households with some high-depreciation assets (like a car) and perhaps a few thousand dollars in cash accounts. These are households that might be able to weather the consequences of a car repair or appliance replacement, but might still face major cash shortages and borrowing needs if faced with bigger but still commonplace disruptions, like a hospital bill or replacement car.

I expect the middle of this distribution to have personal wealth that is concentrated in the value of their owned primary homes, and perhaps a smaller subset of people accumulating business equity or retirement savings in lieu of home equity. People from all walks of life are represented in this group, though people begin this process earlier and run up larger stores of personally-earned wealth over their working lives.

In the top decile and above, you find households with larger and more diversified portfolios, and better access to debt. In my earlier analysis, I found that a lot of the households in the lower ranks includes educated, paired, employed, dual-income households who have accumulated or inherited a portfolio of home equity, retirement savings, and other forms of property. Near the upper-ranks, in the top percentile, business ownership becomes a much larger part of people's asset portfolios. When I last looked at the data, it seemed as if building or inheriting successful businesses was the main way that people became wealthy. I am confident that I will find the same thing when we look at the new data.