# Online Supplement

Do the Wealthy Earn Higher Returns? Evidence from U.S. Survey Data

Christophe Van Langenhove

September 2026

## Contents

## SA.A The replication method

I replicate each study on my PSID or SCF data at its own measurement configuration. Where a study leaves a decision unstated, I estimate the return gradient under each option (Table SA.1). Every replication is an annual real return over the study's own denominator. Its numerator adds the annualized capital gain to the survey year's capital income. The replication of Sundram (2025) stays nominal because his deflation is unstated and his wealth levels are nominal. A first home purchase enters as a flow in every replication because no study addresses it, and two estimates below book it as a capital gain instead.

The crosswalk's rebuilt capital income holds the PSID's business income as one amount, so an asset-part numerator is approximated by booking that income to business owners or by halving it. My PSID wealth panel begins in 2001 and my SCF is the standard SCF from 1989 on heads aged 20 to 80, without the SCF+ history. The PSID reports an interest rate on the primary mortgage alone, so the rest of the cost of debt is imputed. The private-business income adjustments of Xavier (2021) cannot be rebuilt, so the two SCF replications use the SCF's reported income and enter the synthetic yields as a variant.

## SA.B The seven studies

### SA.B.1 Cao and Luo (2017)

Both headline facts of Cao and Luo (2017) reproduce on my data. Their adjacent-window return correlation is 0.22 to 0.30 without capital gains and near zero with them, and mine falls from 0.24 to 0.07. Their return dispersion rises from 11.3 to 27.1 percent once capital gains enter, and mine rises from 21.4 to 32.8 percent. Mine starts from a higher base because a biennial wave smooths less than their five-year window.

I compound each household's annual returns over the panel as they do, on my biennial waves from 2001, which reach none of their 1984 to 1999 waves.

The top wealth decile over the fifth at their configuration gives $+0.9$, inside the range of the baseline return gradients on the gross and net returns on assets. A single wave, a nominal return and a wealth ranking move it between $-4.2$ and $+1.5$ (Table SA.1).

### SA.B.2 Xavier (2021)

Xavier (2021) reports a return gradient of $+4.7$ on the net return on equity, and the replication gives $+4.9$ at her configuration. Its bootstrap interval of ($+4.0$, $+5.7$) covers her number. The top level is $+11.2$ percent against her $+8.3$ and the reference level $+6.3$ against her $+3.6$.

I rebuild her return at the bin level because it is a ratio of aggregates within each wealth bin and wave. Aggregate income plus aggregate index capital gains less aggregate debt service divides by aggregate wealth. As she does, I drop the negative-wealth bottom fifth, compute no private-business return below the 70th wealth percentile, and keep pensions outside her seven components.

Her return gradient rests on the bin-level dollar weighting and on her exclusions. Restoring business income to the bins she excludes turns it to $-9.1$ because the reported business income of small-business owners is large relative to their wealth. Replacing her aggregates with household means under a 0.5 percent outlier trim moves it to $+10.7$.

### SA.B.3 Gaillard and Wangner (2023)

The $+9.6$ of Gaillard and Wangner (2023) does not reproduce. At their configuration the replication gives $+4.7$ with a bootstrap interval of ($+1.2$, $+8.3$), and their headline falls just above it. The PSID's top 1 percent behind that interval holds 25 household-years per wave.

I rebuild their net return on assets at the two-wave midpoint and compare the top 1 percent against the 40th to 50th wealth percentile bin. Imputed rent enters at 5 percent less 2 percent depreciation, and mortgage interest is deducted at a 33 percent tax rate. Their stated denominator adds half the net investment flows to the midpoint, and no variant here adds that term. I trim the wealth aggregate at the household level because their 0.5 percent trim per asset class shapes the class distributions they calibrate and leaves the aggregate unstated.

The ranking basis and its timing, the business-income attribution, the rent yield and the tax treatment move the replication between $+2.9$ and $+15.5$, and their headline lies inside that range. An end-of-period ranking of gross wealth alone reaches $+15.5$, and their paper does not state its ranking timing.

The reference group matches on their terms once a first home purchase counts as a flow, at $+2.6$ against their $+3.6$ percent. Without that treatment a first-time buyer books half the house value as a return, which overstates the housing-heavy reference group. The top 1 percent gives $+7.3$ against their $+13.3$. My construction trims extreme wave-to-wave jumps in reported asset values and keeps only the PSID's SRC subsample, and their paper states neither restriction.

### SA.B.4 Daminato and Pistaferri (2024)

Daminato and Pistaferri (2024) report the steepest headline in the literature, $+30.0$, and it reproduces only under an end-of-period wealth ranking. The replication gives $+6.6$ ranked on beginning-of-period wealth and $+29.6$ ranked on end-of-period wealth with a 0.5 percent outlier trim, and their paper states neither timing.

I keep negative wealth and apply no wealth floor and no outlier trim, because their paper states none of the three. I divide their numerator by beginning-of-period wealth plus half the reported net investment flows, as they do. I drop the low-income oversample as they do, keep heads aged 25 to 65 who are not mostly self-employed, and compare the top wealth percentile against the median wealth group. My files keep only the SRC subsample, so the sample is stricter than theirs on the immigrant and Latino samples. My files cannot build their enrolment and retirement screen, so the sample is wider on that margin.

Under the lagged ranking the return of the median wealth group is $-4.9$ and matches their $-5$ percent. The return of the top wealth percentile is $+1.7$ percent and falls short of their $+26$. The PSID's top wealth percentile holds 25 household-years per wave with a thin right tail. An end-of-period ranking sorts households with a large transitory capital gain into that group, because the capital gain raises wealth rank and measured return in the same wave. Snudden

(2024) ranks on lagged wealth for that reason. Ranked at the end of the period and trimmed at 0.5 percent, the return gradient reaches +29.6 and covers most of their headline.

Left untrimmed, as their rule leaves it, the end-of-period ranking gives −132.3, because a few small-denominator returns in the middle wealth deciles drive the return gradient. Every other unstated decision keeps the lagged ranking and moves the return gradient at most to +13.4, under capital gains without the owner-occupied home. The rest vary an interest-only cost of debt, an un-annualized two-year capital gain, a nominal return and the beginning-of-period denominator. A 0.5 percent outlier trim and their footnote's two-year change in holdings complete the set. Their headline is consistent with an end-of-period ranking their paper does not state, and under a lagged ranking the shared PSID holds no return gradient of that size.

### SA.B.5 Snudden (2024)

The declining profile of Snudden (2024) reproduces at −6.7 against the −6.5 points his Figure 4 implies. The bootstrap interval of (−8.9, −4.6) at his configuration covers his number. His Figure 4 plots predictive margins by lagged wealth ventile with household and time fixed effects. It places the top wealth decile at +3 to +4 percent and the 50th to 60th wealth percentile near +10. His mean return to wealth of +8.2 percent declines with wealth past the 30th wealth percentile.

I rebuild his net return on equity with the household's reported mortgage interest deducted at a flat 25 percent in place of the marginal rate he takes from TAXSIM. I take unconditional group means because his Figure 4 plots fixed-effects margins, so the replication reproduces his profile without his estimator. My wealth is vehicle-free and fully lagged, where his includes vehicles and is wave-averaged for the asset classes without reported flows. I impose neither his restriction to heads born after 1920 nor his stable head-and-spouse requirement, so the sample is wider than his on those two margins.

I apply a single pass of his studentized 99 percent rule, whose standard deviation comes from the distribution still holding the outliers. He states the rule on the level and on the change of asset returns, and adding a change-based pass moves the return gradient by less than a point. Every version of his rule leaves the return gradient declining, at −3.5 when I iterate the rule to a fixed point. With no outlier trim it is −6.8, and at 0.5 percent per tail −7.8. The mean return level under his stated rule is +18.7 against his +8.2, and cleaning steps beyond his stated rule are the likely source of that gap. One such step is his treatment of assets reported at zero

without a sale. His empirical-Bayes estimator of household-specific returns lies beyond this replication, and his descriptive facts recur on my data apart from the mean's level.

### SA.B.6 Hayati (2026)

Hayati (2026) reports +1.7 after 1980 and −1.4 before. The standard SCF gives +4.2 for the post-1980 window, which has the same sign at more than double the size. Every gross-asset estimate across the synthetic yields, the rent yield and the outlier trim stays above his number.

I rebuild his net return on assets on my 1989 to 2022 waves, which lie inside his post-1980 window. His denominator adds half the net investment flows to beginning-of-period gross wealth. A single cross-section has no flow term, so the replication takes gross wealth alone. His Section 2.2 follows the accounting return on assets and his footnote 5 sets it against the wealth denominator of Xavier, which is why the base is gross.

His SCF+ waves extend back past my first wave, and a longer window averages lower. Within my 1989 to 2022 waves the wealth-weighted return gradient falls across the decades. He also harmonizes his income and debt-cost series across seven decades. His debt rates are the households' own reported ones, where mine come from a credit-spread schedule set by wealth decile, and both fall with wealth. A wealth denominator gives +1.8 and matches his number. His own footnote rejects that denominator, because the bottom 90 percent hold most of the leverage that separates the two bases. His positive wealth-weighted return gradient is the existing estimate closest to the compositional effect of Section SA.G.

### SA.B.7 Sundram (2025)

Sundram (2025) reports a return gradient of +4.0 on the gross return on assets, from +1 percent in his bottom gross-asset quintile to almost +5 in the top. The replication gives +3.5 at his configuration once I book business income to business owners, and its bootstrap interval of (+2.7, +4.2) covers his number. Under that booking the top level matches at +5.5 against his almost +5. The bottom level is +2.0 against his +1, and his Figure 9 tabulates no further values.

I rebuild his gross return on assets with no imputed rent and no cost of debt. The denominator lags gross assets for the classes with reported net investment and averages them across waves for the rest. The denominator holds the PSID's net vehicle stock, which differs from the reconstructed gross value of measurement choice 8. I trim at the 0.5 and 99.5 return percentiles, drop the low-income oversample, and compare mean returns across gross-asset quintiles of lagged

gross assets. My files cannot build his screen against students and retirees among heads aged 25 to 65, so the sample is wider than his on that margin. The crosswalk's rebuilt capital income does not split business income into its labor and asset parts. Business income restricted to the households that hold business wealth is therefore the closest available version of his asset-part numerator.

The bottom gross-asset quintile holds almost no gross assets, so where business income is booked decides its return. The crosswalk's rebuilt capital income holds the labor and asset parts of business income as one amount, and my processed capital income assigns every household that amount in full. Counting the self-employment income of the few households there with no business wealth turns the mean return positive and leaves the median at zero. Booking a first home purchase as a capital gain gives his lowest estimate at $-57.4$ (Table SA.1). A first-time buyer in the bottom gross-asset quintile then books half the house value as a return on little prior gross wealth.

## SA.C The common comparison

No headline falls inside the interval of its common comparison (Table SA.1). The common comparison recomputes each study's return type as the mean return of the top wealth decile against the fifth. It applies a $10,000 wealth floor and a 0.5 percent outlier trim. The net return on equity with no imputed rent, as Daminato and Pistaferri (2024) take it, gives +11.1. The net return on equity with imputed rent, as Snudden (2024) takes it, gives −0.1. Common data narrow the spread of the headlines from 36.5 to 13.3 points, and a common configuration narrows it to 11.5.

**Table SA.1: Headlines, the range over the unstated decisions, and the common comparison**

| | | | Unstated decisions | | Common comparison | |
|---|---|---|---|---|---|---|
| Study | Return type | Headline | Lowest | Highest | Estimate | 95% interval |
| Cao and Luo (2017) | $r_A^g$ | – | −4.2 | +1.5 | +6.1 | (+4.5, +7.8) |
| Xavier (2021) | $r_E^n$ | +4.7 | −9.1 | +10.7 | −0.4 | (−0.8, 0.0) |
| Gaillard and Wangner (2023) | $r_A^n$ | +9.6 | +2.9 | +15.5 | +6.2 | (+5.4, +7.0) |
| Daminato and Pistaferri (2024) | $r_E^n$ | +30.0 | −132.3 | +29.6 | +11.1 | (+9.7, +12.7) |
| Snudden (2024) | $r_E^n$ | −6.5 | −7.8 | −3.5 | −0.1 | (−1.9, +1.5) |
| Hayati (2026) | $r_A^n$ | +1.7† | +1.8 | +4.5 | +3.2 | (+3.0, +3.5) |
| Sundram (2025) | $r_A^g$ | +4.0 | −57.4 | +3.5 | +6.0 | (+5.2, +6.7) |

**Notes:** All entries are in percentage points. Headline is the study's reported return gradient. Lowest and highest are the replicated return gradient under the options of the decisions the study leaves unstated, which its subsection names. The common comparison recomputes the study's return type on the study's own survey as the mean return of the top wealth decile against the fifth. It applies a $10,000 wealth floor and a 0.5 percent outlier trim. $r_A^g$, $r_A^n$ and $r_E^n$ are the gross return on assets, the net return on assets and the net return on equity. The interval is a 95 percent bootstrap interval over whole households, within survey year in the SCF. The lowest values of Daminato and Pistaferri (2024) and Sundram (2025) are an untrimmed end-of-period ranking and a first home purchase booked as a capital gain. Cao and Luo (2017) report no return gradient. †Hayati (2026) reports +1.7 on the SCF+ extension after 1980.

## SA.D Asset coverage

Every specification keeps every asset class but vehicles in the asset base, and Cao and Luo (2017) alone exclude owner-occupied housing from it, on the view that the home is a consumption good. Dropping owner-occupied and other real estate from the asset base steepens the return gradient on the median and destabilizes it on the mean. For that reason the specifications keep housing in the asset base throughout.

Dropping housing steepens the median return gradient on the gross return on assets from $+2.2$ points on total wealth to $+2.5$. Housing yields less than the business and equity the top wealth decile holds. This check builds the return gradient from a stock-weighted aggregate of the per-asset-class returns, so the two figures are comparable with each other. Its total-wealth level of $+2.2$ stands beside the headline $+2.8$ of the main text, which uses the household-level return.

On the mean the effect is unstable, from $+4.5$ to $+11.3$ points, because I floor and rank households on total wealth. A housing-heavy household in a low wealth decile meets the floor on its home while holding almost no other assets. Once housing leaves the asset base, its return divides by a small remainder and turns extreme. The mean over the wealth decile follows that household and the median ignores it. Without housing, the median return gradient is stable and the mean return gradient is not.

## SA.E Variable definitions and external series

The housing capital gains and rental yield, the listed-equity capital gains and dividend yield, and the Treasury-bill rate are external series. They come from the long-run U.S. series of Jordà et al. (2019) and from the Federal Reserve Economic Data (FRED), and the replication package gives the exact survey variable codes. Annual FRED series supply the three interest rates of the cost-of-debt imputation. The imputation charges the mortgage debt without a reported interest at the thirty-year fixed mortgage rate and investment-property debt at that rate plus 0.625 points. It charges consumer debt at a consumer-credit rate and vehicle loans at a separate new-car finance rate (Section SA.F).

The consumer price index also comes from FRED, and I deflate each nominal return by the index's change over its return period. That period is the return year for the single-wave SCF return and the average annual change over the two-year wave for the PSID wave return. The deflator corrects level comparisons such as the top 1 percent return.

## SA.F Vehicle data and the interest proxy

The PSID records vehicle wealth net of any vehicle loan, as the amount a household states it would net from selling its vehicles. A specification that includes vehicles adds this net figure to wealth and, for the asset-based returns, adds back the loan proxied below. The SCF records the gross value and the installment loan separately, so it needs no reconstruction and its cost of debt already charges the vehicle loan.

The PSID records no vehicle-debt stock, so I infer the balance from the annual vehicle-loan payment it does record. For a level-payment loan of term $\tau$ years at annual interest rate $r_v$, the average balance over its life is $b(r_v, \tau)$ times the annual payment:

$$b(r_v, \tau) = \frac{1}{r_v}\left[1 - \frac{1 - (1 + r_v)^{-\tau}}{\tau \ln(1 + r_v)}\right]. \tag{SA.1}$$

A household making a payment owes that average balance and pays $r_v$ on it, at the FRED finance rate on new-car loans. The specification takes a five-year term, and at that term $b$ is 2.2. The vehicle interest is then 13 percent of the annual payment. At a three-year term it is 8 percent and at seven 18 percent. The PSID's other consumer-debt stock excludes the vehicle loan, so this cost double-counts nothing.

The baseline excludes vehicles under both measurement choices, because a vehicle is a depreciating consumption durable that earns no return. The vehicle cost of debt also rests on this proxy and on no observed rate.

## SA.G The studies on common data

The return type and the horizon set the sign and the size of the baseline return gradient (Table SA.2). The declining return gradient of Snudden (2024) replicates because leverage produces it. The wealth denominator shrinks in the middle wealth deciles and raises the measured returns there.

One compositional effect survives every measurement choice, since the listed equity and private business the wealthy hold earn more than the housing of the middle wealth deciles. With no rent imputed, the return gradient averaged across the outlier-trimmed SCF specifications is $+3.3$ points. Imputing a 5 percent yield to the housing-heavy middle wealth deciles flattens it to $+0.7$. That remaining effect bounds the composition premium from below, because the synthetic return cannot detect a within-class business premium. The realized PSID return shows a return gradient of $+2.8$ points on the median.

**Table SA.2: The return gradient by return type and horizon**

| | PSID horizon | | SCF |
|---|---|---|---|
| Return type | Single wave | Long run | Single wave |
| Gross return on assets ($r_A^g$) | $+3.1$ | $-1.4$ | $+0.8$ |
| Net return on assets ($r_A^n$) | $+5.5$ | $+0.7$ | $+3.2$ |
| Gross return on equity ($r_E^g$) | $-19.9$ | $-18.2$ | $-8.5$ |
| Net return on equity ($r_E^n$) | $-5.4$ | $-7.1$ | $-0.4$ |

**Notes:** Each entry is the mean return gradient of the top wealth decile against the fifth, in percentage points, annual and real. Every entry uses imputed rent, a \$1,000 wealth floor and a 0.5 percent outlier trim. The PSID single-wave column is the annual wave return and the long-run column averages a household's waves. The SCF column is a single annual wave.

## SA.H Specification diagnostics

### SA.H.1 Clustering the specifications

A $k$-means clustering of the outlier-trimmed specifications separates the 4 groups on the return type, the PSID or SCF source, the reference wealth decile and the PSID horizon (Table SA.3). The clustering runs on the indicator-coded measurement choices and the standardized return gradient, with the return gradient weighted three to one against each indicator. The steep group reproduces the pattern of the steepest specifications, a low reference wealth decile and a single wave (Section SA.I).

### SA.H.2 The variance decomposition

The return type and the summary statistic lead the pooled one-way shares of the cross-specification variance and the PSID Shapley shares (Table SA.4). The two decompositions rank the measurement choices the same way. I call the thirteen measurement choices, the return type and the PSID-or-SCF source the decisions. For each decision I fit the return gradient on that decision alone and report the share of the total sum of squares it explains. Those one-way shares do not sum to one because the models omit interactions and the decisions are not orthogonal. The reference wealth decile and the imputed rent move the median most and have small variance shares, so their importance appears in the per-return-type effects the main text reports.

**Table SA.3: Clusters of specifications by their measurement choices**

| Cluster | Share | Mean return gradient | Survey | Return type | Horizon | Reference wealth decile |
|---|---|---|---|---|---|---|
| 1 | 5% | −13.5 | PSID | gross return on equity | single wave | fourth |
| 2 | 45% | −0.1 | PSID | net return on equity | long run | sixth |
| 3 | 1% | +2.3 | SCF | net return on equity | – | third |
| 4 | 48% | +5.7 | PSID | net return on assets | single wave | third |

**Notes:** A $k$-means clustering of the 588,288 outlier-trimmed specifications on their indicator-coded measurement choices and their standardized return gradient, up-weighted three to one, with $k = 4$. Clusters are ordered from the lowest to the highest mean return gradient, in percentage points. Each entry after the mean return gradient is the cluster's modal value of that measurement choice. A dash marks the SCF cluster, where no horizon choice arises.

The mean-based specifications without an outlier trim dominate the pooled cross-specification variance.

Both decompositions run on the core specifications, so each remaining share measures one decision alone. The core excludes vehicles and the vehicle cost of debt and pins the business-income attribution at baseline. The core holds 103,036 PSID specifications, 11.9 percent of the reported PSID specifications, and 49.9 percent of the SCF's. The three decisions the core fixes are measured on the slice where each of them varies.

The Shapley value $\phi_c$ of decision $c$ averages its marginal contribution to the explained variance over all orderings of the set $\mathcal{C}$ of decisions that vary within the survey. A subset $S \subseteq \mathcal{C}$ explains the variance share $R^2(S)$:

$$\phi_c = \sum_{S \subseteq \mathcal{C} \setminus \{c\}} \frac{|S|!\,(|\mathcal{C}| - |S| - 1)!}{|\mathcal{C}|!} \left[ R^2(S \cup \{c\}) - R^2(S) \right]. \tag{SA.2}$$

The Shapley shares sum to the full-model $R^2(\mathcal{C})$ of 0.36 in the PSID and 0.45 in the SCF. A main-effects model over these decisions therefore leaves most of the cross-specification dispersion unattributed. In the PSID the return type and the summary statistic lead at 0.12 and 0.10, where the ranking timing leads the PSID-specific decisions at 0.05. In the SCF the return type leads at 0.30 and imputed rent follows at 0.07, so imputed rent matters more in the SCF than in the PSID. The outlier trim and the wealth floor move the mean-based specifications and leave the median-based ones in place, because a median ignores the extreme returns a trim removes.

Reducing the wealth-floor options to the two extremes and recomputing the one-way shares leaves the ranking stable. The rank correlation is 1.00 in the PSID and 0.98 in the SCF, so the ranking does not depend on the number of options per measurement choice.

Table SA.5 reports the same one-way shares split by survey, beside the effect of each measurement choice pooled over the mean-based and median-based specifications. That effect is the spread of the median return gradient across the measurement choice's options for each return type.

### SA.H.3 Equal weight on every option

Giving every option of each measurement choice that applies to every return type equal weight lowers the median across specifications from $+2.5$ to $+1.5$ points. The share above 10 points falls from 2.4 to 1.9 percent. The average across specifications falls from $+1.4$ to $-0.3$. Leverage pushes the gross return on equity below zero, so equal weight on it lowers the center. A majority of specifications stays positive at 63 percent.

**Table SA.4: Share of the return gradient's cross-specification variance explained by each decision**

| | One-way share | Shapley share | |
|---|---|---|---|
| Measurement decision | Pooled | PSID | SCF |
| Shared decisions | | | |
| Return type | 0.12 | 0.12 | 0.30 |
| Survey | 0.00 | – | – |
| Cost of debt: which debts | 0.00 | 0.00 | 0.03 |
| Cost of debt: interest rate | 0.00 | 0.00 | 0.01 |
| Imputed rent | 0.03 | 0.03 | 0.07 |
| Vehicle scope | 0.02 | – | – |
| Wealth floor | 0.00 | 0.00 | 0.00 |
| Outlier trim | 0.04 | 0.04 | 0.00 |
| Summary statistic | 0.09 | 0.10 | 0.00 |
| Reference wealth decile | 0.00 | 0.00 | 0.04 |
| Dataset-specific decisions | | | |
| Business income (PSID) | 0.00 | – | – |
| Vehicle cost of debt (PSID) | 0.01 | – | – |
| Horizon (PSID) | 0.00 | 0.00 | – |
| Denominator timing (PSID) | 0.00 | 0.00 | – |
| Ranking timing (PSID) | 0.05 | 0.05 | – |

**Notes:** Each entry is the share of the return gradient's cross-specification variance that one decision explains. The one-way share pools the two surveys, and the Shapley share is computed within one survey. The survey entry is the share the PSID-versus-SCF split explains. One-way shares do not sum to one because the decisions are not orthogonal. The Shapley shares sum to the full-model $R^2$ of 0.36 and 0.45 up to rounding at two decimals. A dash marks a decision the core fixes or that does not arise in that survey.

I reweight by iterative proportional fitting because the specifications are not drawn from a population. The center of their distribution depends on how many specifications take each option of each measurement choice. The imbalance that matters is in the return type, because the two gross returns omit the cost of debt and with it the vehicle cost of debt. The two net-of-

debt returns therefore hold several times as many specifications, and the unweighted center over-represents them.

Reweighting the four return types alone gives +1.6 points, close to the fully reweighted figure. The reweighted median is +1.5 in the PSID and +1.9 in the SCF. Weighting the two surveys equally in place of their specification counts gives +1.7. Steep specifications stay rare under every weighting.

**Table SA.5: The measurement choices ranked by importance, by survey**

| Measurement choice | Variance share | Effect by return type (pp) | | | |
|---|---|---|---|---|---|
| | | $r_A^g$ | $r_A^n$ | $r_E^g$ | $r_E^n$ |
| Panel A: PSID | | | | | |
| Summary statistic | 0.10 | 1.5 | 1.6 | 8.5 | 6.7 |
| Ranking timing | 0.05 | 3.6 | 3.5 | 5.2 | 5.5 |
| Outlier trim | 0.04 | 0.8 | 0.8 | 3.9 | 2.8 |
| Imputed rent | 0.03 | 2.2 | 1.8 | 8.1 | 7.1 |
| Vehicle scope | 0.02 | 0.8 | 0.5 | 3.5 | 2.0 |
| Vehicle cost of debt | 0.01 | – | 0.2 | – | 0.4 |
| Denominator timing | 0.00 | 1.6 | 1.5 | 2.2 | 1.6 |
| Cost of debt: rate | 0.00 | – | 0.0 | – | 0.1 |
| Reference wealth decile | 0.00 | 3.7 | 2.7 | 12.2 | 6.5 |
| Horizon (one vs. several waves) | 0.00 | 0.7 | 0.8 | 1.3 | 1.5 |
| Business income | 0.00 | 0.2 | 0.2 | 0.6 | 0.6 |
| Cost of debt: which debts | 0.00 | – | 0.4 | – | 1.3 |
| Wealth floor | 0.00 | 0.4 | 0.4 | 0.2 | 0.9 |
| Full return-gradient range | | $[-97.4, +21.0]$ | $[-94.9, +22.8]$ | $[-100.0, +18.4]$ | $[-99.9, +26.1]$ |
| Panel B: SCF | | | | | |
| Imputed rent | 0.07 | 2.1 | 1.6 | 4.2 | 3.8 |
| Reference wealth decile | 0.04 | 2.7 | 2.3 | 5.8 | 3.5 |
| Vehicle scope | 0.03 | 0.5 | 0.6 | 2.4 | 1.0 |
| Cost of debt: rate | 0.01 | – | 0.1 | – | 0.6 |
| Cost of debt: which debts | 0.01 | – | 1.3 | – | 2.7 |
| Outlier trim | 0.00 | 0.2 | 0.4 | 1.5 | 0.5 |
| Wealth floor | 0.00 | 0.1 | 0.3 | 0.1 | 0.5 |
| Summary statistic | 0.00 | 0.8 | 1.2 | 3.5 | 0.7 |
| Full return-gradient range | | $[-0.1, +6.0]$ | $[+0.5, +8.0]$ | $[-64.2, +6.4]$ | $[-30.1, +42.2]$ |

**Notes:** Variance share is the share of the return gradient's cross-specification variance a measurement choice explains in a one-way analysis of variance, pooled over the four return types within each survey. Effect is the spread of the median return gradient across the measurement choice's options at a fixed return type, in percentage points. $r_A^g$, $r_A^n$, $r_E^g$ and $r_E^n$ are the gross return on assets, the net return on assets, the gross return on equity and the net return on equity. A dash marks a measurement choice the return type does not use. The horizon, the two timings, the vehicle cost of debt and the business-income attribution apply to the PSID only. Shares and effects hold the vehicle measurement choices and the business-income attribution at baseline and measure those three where each varies. Shares pool the mean-based and median-based specifications and include those without an outlier trim. Full return-gradient range is the lowest to the highest return gradient across every specification of that return type.

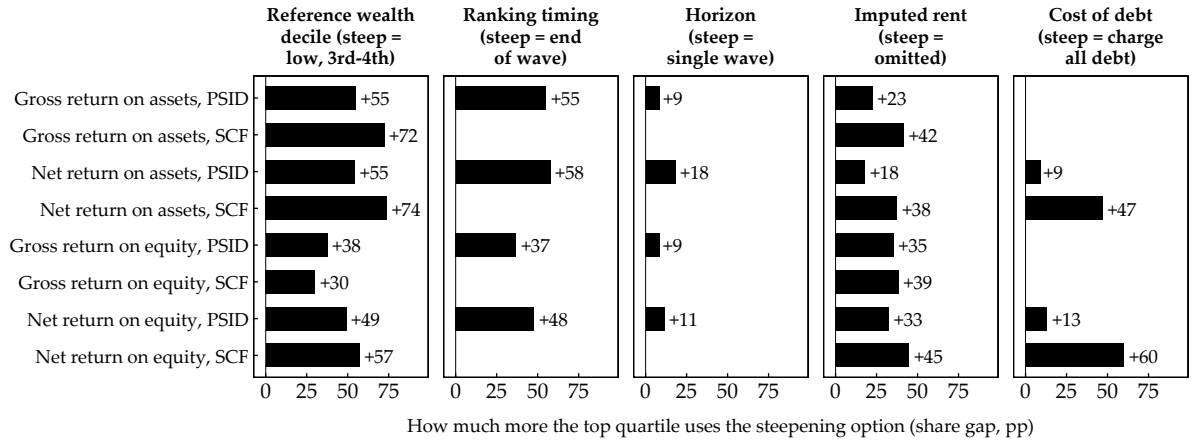
## SA.I The highest and lowest return gradients

Five measurement choices separate the top quartile of return gradients from the bottom quartile within each survey and return type (Figure SA.1). I split the outlier-trimmed specifications into the two quartiles. For each measurement choice I report how much more often the top quartile takes the steepening option than the bottom quartile does.

A reference wealth decile at the third or fourth appears in most top-quartile specifications and divides the quartiles in every panel, least in the SCF gross return on equity. In the PSID an end-of-wave ranking and a single wave both do the same in every return type. Omitting imputed rent divides the quartiles in every panel, more in the SCF than in the PSID. Charging all debt in place of mortgage debt alone matters for the two net-of-debt returns, mainly in the SCF.

The bottom quartile takes a high reference wealth decile, imputed rent, and in the PSID a start-of-wave ranking and a long-run average. For the two asset-based returns that bottom quartile lies near a zero return gradient. For the two equity-based returns it is negative, because leverage turns the return gradient downward there.

**Figure SA.1: Options separating the highest and lowest return gradients**




**Notes:** Outlier-trimmed specifications of each survey and return type, split into the top and bottom quartile of the return gradient. Each panel is one survey and return type, and each bar one measurement choice. A bar is the share of top-quartile specifications taking that measurement choice's steepening option less the share of bottom-quartile specifications taking it, in percentage points. The steepening options are a third or fourth reference wealth decile, an end-of-wave ranking, the single-wave horizon, no imputed rent, and all debt charged. The ranking timing and the horizon apply to the PSID only, and the cost of debt enters only the two net-of-debt return types.

## SA.J Coverage of the wealth tail

The PSID has no oversample of the wealthy, so its top wealth decile holds 356 household-years per wave at the median and its top wealth percentile 25. The top wealth decile ranges from 282 to 395 household-years across the 12 waves. The SCF's oversample fills the same groups with 1,410 and 649 households per wave at the median across its 8 waves from 2001 onward. The two surveys share that period, and the fuller groups are why the SCF's intervals are narrower.

Dropping the top 1 percent by wealth rank leaves the PSID return gradient in place and removes the SCF's positive part. Measured to the top 1 percent, the mean return gradient is $+0.6$ in the PSID and $+1.9$ in the SCF. With the top wealth decile reduced to its 90th to 99th wealth percentile it is $+3.2$ and $+0.7$ at the mean. At the median it is $+2.9$ and $-0.2$, so the PSID return gradient rises and the SCF's falls to about zero. I drop the top 1 percent by rank because the SCF's sample-design flag is confidential and its oversample cannot be removed directly. The check therefore localizes any top premium in place of removing it.

Reweighting the PSID's thin top toward a heavier Pareto tail raises the median return gradient on the gross return on assets to between $+11.5$ and $+14.7$ points. The reweighting runs over tail indices from 1.3 to 1.8 around the fitted index of 1.46, which is in the range standard for U.S. wealth. The check varies the weight on the observed thin top and leaves the coverage of unobserved richer households as it is. That coverage is what the SCF oversample and Balloch and Richers (2023) address.

Weighting each PSID household-year by its longitudinal family weight gives $+2.4$ at the median and $+2.7$ at the mean. The equal-weighted baseline gives $+2.8$ at the median and $+3.1$ at the mean, so both differences are under half a point. The equal weighting rests on the SRC sample, which the screen keeps close to representative (Section SA.L). This check mirrors the drop of the SCF survey weight in Appendix A of the main text.

## SA.K Risk-adjusted returns

Adjusted for risk, the median Sharpe-ratio gradient in the PSID is +0.20 (+0.16, +0.23). The raw return gradient on the same gross return on assets is +2.8. I test the adjustment because richer households bear more portfolio risk, and a raw premium can vanish once risk is priced. I test it in the PSID alone because the synthetic SCF assigns a common return within each asset class and has no idiosyncratic risk. The one U.S. study that risk-adjusts directly finds the premium at the very top widening once risk is priced (Balloch and Richers, 2023). Fagereng et al. (2020) show their return gradient survives controls for portfolio shares and betas without risk-adjusting.

I measure portfolio risk from the dispersion of realized returns within each asset class. I first drop the returns beyond the Cao and Luo (2017) bound of 300 percent, which a near-zero denominator would otherwise make unbounded. The class risks are 18.5 percent for housing, 54.4 for equity, 68.6 for business and 14.9 for fixed income, so business and equity hold most of the risk. A household's portfolio risk is $\sigma_i = \sqrt{\sum_a \omega_{ia}^2 \sigma_a^2}$, where $\sigma_a$ is the risk of asset class $a$ and $\omega_{ia} = A_{ia}/A_i$ the household's share of gross assets in it. That risk rises with wealth, from a mean of 16.8 percent in the reference wealth decile to 18.4 percent in the top, as the rich tilt toward business and equity.

The dispersion behind $\sigma_i$ is realized, so it includes transitory variation as well as risk. A household's average return over its odd waves and over its even waves correlate at $-0.04$ (Section 5.2 of the main text), so the dispersion overstates risk. That overstatement inflates the denominator in both wealth deciles, and the return gradient it produces is unbiased only where the inflation is the same in each.

I take the risk-adjusted return as the Sharpe ratio $r^g_{A,i}/\sigma_i$ at a zero risk-free rate. I take the return gradient as the top wealth decile minus the fifth, as in the main text. The mean Sharpe-ratio gradient is outlier-sensitive, because a household whose modeled risky assets are a small share of its gross wealth has little measured risk and a large Sharpe ratio. Adding a common cross-class return correlation of 0.3 to the portfolio risk leaves the median Sharpe-ratio gradient at +0.16. Neither the PSID nor the SCF can detect the risk-adjusted premium from diversification and private-business access that Fagereng et al. (2020) and Balloch and Richers (2023) find, so this check does not speak to it.

## SA.L The PSID and SCF samples

### SA.L.1 Sample period and selection

The PSID sample runs over the 12 biennial waves from 2001 to 2023. The SCF sample runs over the 12 triennial waves from 1989 to 2022. Both samples keep household heads aged 20 to 80. The PSID estimation sample keeps the SRC core subsample and drops the low-income SEO oversample, so the equal weighting of Section SA.L.2 stays close to representative. The baseline wealth floor of $1,000 is the most common in the literature, and across specifications the floor ranges from none to $10,000.

A PSID household-year leaves the sample on a non-positive denominator stock, and an SCF observation on missing wealth or survey weight, a negative wealth denominator or a missing deflator. In each survey I trim the return within year and denominator after the floor, at 0.5 or 3 percent per tail. The two exclusions of the construction appendix of the main text apply under the option without an outlier trim.

The PSID composition rule runs upstream and sets aside the capital gains of a household-year whose wealth moves in or out with a change in household composition. It flags 480 household-waves or 0.9 percent at a threshold of 1.0 percent of prior wealth floored at $1,000. The asset-exit rule also runs upstream and removes only the capital gain of an asset class that disappears between waves. It sets aside 12,094 of 243,378 asset-class observations or 5.0 percent, counted against asset-class observations because the rule flags one class at a time.

In the SCF the head-age window removes 12,525 of 290,642 records, 4.3 percent unweighted and 5.2 percent weighted. The upstream bounds on the portfolio shares and the ratio of mortgage debt to assets remove 98 of 290,740 records or 0.0 percent.

### SA.L.2 Sample sizes, weighting, and imputation

Under the gross return on assets at the baseline the PSID contributes 27,482 household-year returns from 7,174 households after the floor and the outlier trim. The SCF contributes 223,068 observations across its waves. I weight PSID observations equally and let the bootstrap resample whole households, which are the panel's sampling unit. I weight SCF observations by the survey sampling weight, and the SCF supplies five imputation replicates per household. The effective sample is therefore about a fifth of the observation count. The SCF bootstrap resamples whole households within survey year, each drawn household with its five imputed

copies. The intervals combine the within-implicate resampling variance with the between-implicate variance by the rule of Rubin (1987). That rule widens the SCF intervals by a factor of 1.16 for the mean and 1.03 for the median.

### SA.L.3 Wealth deciles and top groups

Within each survey year I rank households into wealth deciles before the wealth floor applies. The PSID ranks unweighted on wealth at the wave's end under the baseline, and the SCF ranks on wealth weighted by the survey weight. The wave point of the PSID's ranking wealth is measurement choice 11, and other specifications rank on the wave's start and midpoint. One construction builds all three wealth-decile sets, which differ only in the wave point.

The return gradient compares the tenth wealth decile with a reference wealth decile from the third through the sixth, with the fifth as the common value. The top and reference wealth deciles come from one common ranking within each specification. I take the top 1 percent as the households above the within-year 99th wealth percentile. That wealth percentile is of mid-wave wealth in the PSID, which saves no end-of-wave wealth percentile, and of wealth in the SCF. I take the top wealth quintile as the ninth and tenth wealth deciles together. Under the long-run horizon I assign a household its modal wealth decile over the waves in which it appears.

### SA.L.4 Recovering other real estate before 2013

Through 2011 the PSID records other real estate net of the mortgage against it, and it reports the value and the debt separately only from 2013. The construction appendix of the main text needs the gross value for the denominator and the other-housing capital gain, and the mortgage for the cost of debt. I estimate a loan-to-value model on the SCF, where the value and the mortgage are both observed over 1989 to 2022. A first gradient-boosted stage classifies whether the property has any mortgage, and a second predicts the loan-to-value ratio conditional on one. Applying the model to the PSID splits the net figure into a gross value and a mortgage balance through 2011. From 2013 I use the reported gross value and mortgage.

## References


**Balloch, Cynthia, and Julian Richers.** 2023. "Asset Allocation and Returns in the Portfolios of the Wealthy." Financial Markets Group Discussion Paper 885, London School of Economics.

**Cao, Dan, and Wenlan Luo.** 2017. "Persistent Heterogeneous Returns and Top End Wealth Inequality." *Review of Economic Dynamics*, 26: 301–326.

**Daminato, Claudio, and Luigi Pistaferri.** 2024. "Returns Heterogeneity and Consumption Inequality over the Life Cycle." NBER Working Paper 32490.

**Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri.** 2020. "Heterogeneity and Persistence in Returns to Wealth." *Econometrica*, 88(1): 115–170.

**Gaillard, Alexandre, and Philipp Wangner.** 2023. "Wealth, Returns, and Taxation: A Tale of Two Dependencies." SSRN Working Paper 3966130.

**Hayati, Mojtaba.** 2026. "Scale-Dependent Returns and the Interest Rate." Working Paper, University of Zurich.

**Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M. Taylor.** 2019. "The Rate of Return on Everything, 1870–2015." *Quarterly Journal of Economics*, 134(3): 1225–1298.

**Rubin, Donald B.** 1987. *Multiple Imputation for Nonresponse in Surveys.* New York:John Wiley & Sons.

**Snudden, Stephen.** 2024. "Leverage and Rate of Return Heterogeneity among U.S. Households." LCERPA Working Paper 2024-8.

**Sundram, Jacob.** 2025. "Monetary Policy, Heterogeneous Returns, and Top Inequality." Job Market Paper, University of Copenhagen.

**Xavier, Inês.** 2021. "Wealth Inequality in the US: The Role of Heterogeneous Returns." Federal Reserve Board Working Paper. SSRN 3915439.