Realized International Returns Differentials[‡]

Alexandra Tabova
Francis E. Warnock

August 30, 2026

Abstract

We revisit estimates of the U.S. international returns differential using comprehensive security-level data on portfolio debt and equity holdings and multiple measures of returns. We find that security-level estimates of the returns differential are smaller than those based on aggregate data. The difference is most pronounced for realized returns, which account for the timing and magnitude of purchases and income streams. Over 2003-2022, the realized U.S. portfolio returns differential was *negative*. These findings challenge the large positive returns differential that features prominently in the literature on the U.S. exorbitant privilege.

JEL Classification: F30, G11, G12
Keywords: U.S. international returns, exorbitant privilege, bonds, equities

[*]Alexandra Tabova: International Finance Division of the Board of Governors of the Federal Reserve System (email: alexandra.m.tabova@frb.gov); Warnock: University of Virginia's Darden Business School and NBER (email: few9t@virginia.edu). The authors thank Gabriel Chodorow-Reich (co-editor) for invaluable guidance; Ben Stagoff-Belfort for excellent research assistance; and, for helpful comments, discussants Fernando Broner and Julia Schmidt; Ritt Keerati, Philip Lane, and Felipe Saffie; and participants at the CEPR International Macroeconomics and Finance Programme Meeting, the Banque de France Global Currencies Workshop, Bangko Sentral ng Pilipinas, the Central Bank of Chile, the ECB, Johns Hopkins SAIS, Keio University, Korea University, and the University of Tokyo. The views expressed in this paper are the sole responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System. A previous version of this paper was circulated under the title "Exorbitant Changes in Three Parts."

[†]Annual aggregates of all series in this paper's exhibits will be posted at https:/sites.google.com/view/francis-warnock/research/returns-differentials.

Introduction

Rogoff and Tashiro (2015, p. 44) define exorbitant privilege broadly as when "a country enjoys a cost advantage in any kind of borrowing or investment instrument, from equities to bonds (both government and private) to foreign direct investment." In this paper we focus on one component of this broad definition—the difference between returns on U.S. international portfolio claims and liabilities (henceforth, "returns differential" or RD). Early estimates of this differential were large.[1] The notion that "the interest that the U.S. must pay on its foreign liabilities is two to three percentage points less than the rate of return on its foreign investments" (Eichengreen 2011, page 4) persists and is incorporated in some form in recent theoretical models such as Maggiori (2017), Gopinath and Stein (2021), Kekre and Lenel (2024) and Itskhoki and Mukhin (2025).

We show that the RD on portfolio claims and liabilities over the past two decades has been smaller than previously thought, even negative. Our analysis updates the earlier estimates in three important ways. First, we use data from comprehensive and mandatory security-level surveys of U.S. cross-border portfolio investment through 2022. Second, we use these security-level data to carefully adjust for the timing of income streams and position data and to more accurately measure the flows associated with each asset. Third, we report RDs for three return measures that each answer a slightly different question about exorbitant privilege: arithmetic mean (AM), geometric mean (GM), and the internal rate of return (IRR) that takes into account the timing and magnitude of purchases and income streams.[2]

We use three measures of returns: the arithmetic mean (AM), which is commonly used

[1]Early estimates in what Curcuru et al. (2013) termed the "first wave" generally used aggregate data published by the Bureau of Economic Analysis (BEA); examples include Gourinchas and Rey (2007), Obstfeld and Rogoff (2005), Lane and Milesi-Ferretti (2007), and Meissner and Taylor (2008). For assessments, see Lane and Milesi-Ferretti (2009) and Curcuru et al. (2013).

[2]Our analysis covers 2003-2022, the period for which annual U.S. security-level portfolio surveys are available. We exclude direct investment—an important component of cross-border claims and liabilities—because we do not have access to firm-level FDI data. U.S. MNCs' international earnings have persistently exceeded those of foreign companies in the U.S. (Curcuru et al. 2013).

in the exorbitant privilege literature; the geometric mean (GM); and the IRR. Terms used in the literature for AM and GM are instructive. Dichev (2007) calls the GM the "securities" return because it measures the returns on the securities in a portfolio, not those earned by investors (unless no trading occurs during the sample). Bessembinder et al. (2025) refer to the AM as the "cash-account" return because earning it requires maintaining a constant investment amount: each period's returns must be swept into a cash account that earns a zero rate of return. Neither AM nor GM measures investors' realized returns.

IRR is the rate of return that makes the net present value of cash flows zero. In our setting, cash flows have four components: the initial and final values of the portfolio, interim net sales (i.e., negative of net purchases), and interim income streams (dividends and coupon payments). One contribution is to construct these components from security-level data. This had not previously been done for two components: U.S. international net purchases (capital flows) and income streams (part of primary income in the current account). Our estimates of these components, particularly of income streams, can improve the quality of published U.S. data.

A second contribution is to show that, for each of the three return measures, estimated U.S. RDs are smaller with security-level data than with aggregate data. This result is subject to an important caveat: all estimates—whether based on the annual return series used to calculate AM and GM or on the cash flows used to calculate IRR—draw on only 19 years of annual data. In addition, the magnitude of the difference between aggregate and security-level RDs depends on two choices for computing aggregate returns: scaling factor (the lagged position or the average of current and lagged positions) and as-of date (specifically, December, natural for calendar year data, or June). We compute aggregate RDs using all four combinations: those scaled by lagged positions are large (about 3%), while scaling by average positions yields more moderate differentials (1.5%-2.0%). All are higher than our security-level AM differential of 0.93%.

Our baseline uses as-of-June data for aggregate liabilities and, following Gohrband and Howell (2015), scales by average positions. The argument for June aggregate liabilities is straightforward. To compare aggregate and security-level RDs, the data should be synchronized. Claims data are synchronized: security-level surveys are conducted at end-December, and BEA data are measured by calendar year. Liabilities data are not synchronized because the security-level surveys are conducted at end-June.

AMs remain important because they are widely used, but our preferred measure of realized returns is the IRR, which we compute using synchronized data. The difference between aggregate and security-level IRRs is sizable. IRRs increase with net cash flows, which comprise the change from the portfolio's initial to final value, interim net sales, and interim income streams. Differences between aggregate and security-level IRRs occur both within asset classes and for the overall differential. Within-asset-class differences can be large. For example, the IRR on U.S. investors' foreign-equity portfolio is 7.87% with aggregate data but only 5.79% with security-level data. This gap primarily reflects reported U.S. income streams that are too large and reported U.S. flows into foreign securities that are smaller than our security-level estimates. These factors are also important for the overall differential, which aggregate data suggest is large (IRR of 1.41%). Cash flows computed from the security-level data provide a more accurate picture: a realized U.S. RD of -0.45%.

Overall, our results suggest a reassessment of the broad notion of exorbitant privilege. Theoretical papers embed a U.S. RD that has become much smaller—and, by some measures, negative—although it could reemerge. Focusing on the dollar amount of "valuation adjustments," Atkeson et al. (2025) argue that ex post privilege has ended. We make a similar point using security-level data and realized "dollar-weighted" returns.

The paper proceeds as follows. Section 1 discusses the construction of annual returns series using aggregate and security-level data. Section 2 presents AM RDs and explains why aggregate and security-level AMs differ. Section 3 presents GM returns and IRRs, shows

the sources of the substantial difference between aggregate and security-level IRRs, and highlights the composition and timing effects. Section 4 discusses some implications for the literature. Section 5 concludes. An Online Appendix (OA) provides information on BEA data series and the collection, quality checks, and cleaning of the security-level data, along with additional exhibits.

1 Construction of Annual Returns Series

Throughout this paper, we calculate annual returns using aggregate data from BEA (aggregate) and security-level data from Treasury surveys (security-level). Unless otherwise specified, annual data are measured by calendar year (CY). We use "June" to denote annual June-to-June data: end-June for stock variables (positions) and the 12 months ending in June for flow variables (e.g., net purchases and income streams). We refer to foreigners' U.S. portfolios as liabilities and to U.S. investors' foreign portfolios as claims (or assets).

1.1 Returns Series Using Aggregate Data

Gohrband and Howell (2015) create annual returns series from positions, income streams, and valuation (i.e., price and exchange rate) changes reported in BEA's International Investment Position (IIP). Specifically:

$$\text{RTOT}_t = (INC_t + VAL_t)/AVGPOS_t \tag{1}$$

where, for year t, RTOT_t is the total rate of return, INC_t is income, the valuation change VAL_t is the change in the investment position caused by changes in prices and exchange rates, and AVGPOS_t is the average of the investment positions (POS) in years t and *t-1* . Lagged positions (POS_{t-1}) provide another plausible scale factor for the dollar amount of

returns. In a rising market, however, scaling by lagged positions may overestimate RTOT if some of the year's dollar returns arise from purchases made during the year, .

Equation (1) can be computed using CY annual data, which alines with the end-December security-level *claims* surveys (discussed below). Because security-level *liabilities* surveys are conducted at end-June, we also create an annual June version of aggregate liabilities using BEA's quarterly Balance of Payments (BOP). INC is quarterly in BOP; creating an annual June-to-June version is straightforward. POS is available quarterly beginning in 2006. For earlier years, we use the annual June surveys to estimate June POS; this is reasonable because BEA has based IIP positions on those surveys since 2003. Quarterly IIP data do not report VAL until after our sample period, but we can construct a quarterly version. In the IIP, the valuation change equals the change in the position minus adjusted flows (reported flows plus other changes):

$$VAL_t = (POS_t - POS_{t-1}) - (FLOWS_t + OTCHG_t) \tag{2}$$

BOP flows are quarterly. Because OTCHG is available only on a CY basis, we use the average of its values in t and *t-1* as a reasonable estimate of annual June OTCHG. [3]

We create four measures of RTOT_t using two as-of months for aggregate liabilities (June and December) and two scale factors (AVGPOS_t and POS_{t-1}). Our baseline uses AVGPOS and data synchronized with the surveys (i.e., June liabilities data).

1.2 Returns Series Using Security-Level Data

Security-by-security data on U.S. claims and liabilities—the amount U.S. residents hold of each foreign security and the amount foreign residents hold of each U.S. security—are col-

[3]Equations (1) and (2) treat other changes (officially, "changes in volume and valuation n.i.e.") as distinct from valuation changes. This treatment is consistent with Gohrband and Howell (2015, p. 266), who explain that "other changes" are excluded because they do not accrue to investors during the year; such changes are more accurately treated as flows. In the rest of this paper, when we refer to aggregate flows, we mean reported flows plus other changes.

lected through annual surveys by the U.S. Department of the Treasury as part of the TIC reporting system.[4] The security-level data underlie the annual TIC reports on foreign holdings of U.S. securities and U.S. holdings of foreign securities; they also feed into official BEA data on the U.S. IIP. The main reporters are U.S.-resident custodians, including brokers and dealers. They must report all U.S. securities held on behalf of foreign residents—including through their own foreign subsidiaries and affiliates—and all foreign securities held on behalf of U.S. residents. Because reporting is mandatory, the holdings data are comprehensive; they capture the entire foreign portfolio of U.S. securities and the entire U.S. portfolio of foreign securities at the individual security level.

The security-level data have been collected annually since 2003; liabilities data are reported as of June 30, while claims data are reported as of December 31. Liabilities survey data distinguish between foreign holdings of U.S. Treasuries, U.S. Agencies, U.S. corporate bonds, and U.S. equities; claims data distinguish between U.S. holdings of foreign bonds and foreign equities. Survey data are reported based on the residence of the direct owner reported by the custodian, which is not necessarily the ultimate owner.

We conduct several checks, detailed in the OA. First, to confirm that all securities are included, we compare total holdings in each asset class with the surveys' published aggregates. Second, where necessary, we improve the surveys' price and payment-term data by comparing them with ICE Global Index constituent data for several U.S. and foreign bond indices, TreasuryDirect data on Treasury securities, Refinitiv data on foreign securities, the CRSP U.S. Stock Database for U.S. and foreign equities, and Bloomberg Back Office Security data for unmatched securities. This also helps with outlier entries and misreported pricing. Using the validated and cleaned survey data, we calculate the total return per security for each survey date, aggregate the returns to the asset-class level, and check that these are close

[4]See Tabova and Warnock (2026) on the origins of the TIC system.

to index returns for U.S. bonds, U.S. equity, and foreign equity.[5] The close correspondence provides confidence in the quality of the price and payment-term data.

To construct the annual return series from security-level data, we define security b's total rate of return (ROR) at time t:

$$\text{ROR}_{b,t} = \frac{C_{b,t} + P_{b,t} - P_{b,t-1}}{P_{b,t-1}} \tag{3}$$

where:

- $\text{P}_{b,t}$: security b's price at time t.
- $\text{C}_{b,t}$: security b's dividend or coupon income at time t.
  - Bonds: $C_{b,t} = \text{Coupon rate}_{b,t-1} * \text{Face value bond}_{b,t-1}$.
  - Equity: $C_{b,t} = \text{Dividend per share}_{b,t-1} * \text{Number of shares}_{b,t-1}$ or $C_{b,t} = \text{Dividend yield}_{b,t-1} *$ Equity market $\text{value}_{b,t-1}$.
- $\text{ROR}_{b,t}$: security b's total rate of return at time t.

We weight each of the N securities' returns by its share in the portfolio. Thus, the weighted-average rate of return from year *t-1* to year t, ROR_t, is:

$$\text{ROR}_t = \sum_{b=1}^{N} \omega_{b,t-1} \text{ROR}_{b,t} \tag{4}$$

where the weight $\omega_{b,t-1}$ equals holdings of security b at *t-1*, $\text{H}_{b,t-1}$, divided by total portfolio holdings at *t-1*:[6]

$$\omega_{b,t} = \frac{\text{H}_{b,t}}{\sum_{b=1}^{n} \text{H}_{b,t}} \tag{5}$$

[5] There is no readily available aggregate index to compare our foreign bond returns.

[6] Because weights are measured at *t-1*, a security must exist at *t-1* to be included in returns calculations.

This calculation yields a series of T annual returns. Because the annual surveys are conducted at different times, the liabilities portfolios run from June 2003 through June 2022, whereas the claims portfolios run from December 2003 through December 2022.[7]

2 Aggregate and Security-Level AM Returns Differentials

2.1 The Estimates

Table 1 presents aggregate-data RDs based on the AM (the simple average of annual returns) for four combinations of scaling factor (lagged or average positions) and as-of month for aggregate liabilities (June or December).[8] The first two rows of Panel A, which use December aggregate liabilities, are included for completeness because they rely on BEA's reported data and are therefore natural choices for researchers. For comparisons of aggregate and security-level differentials, synchronized (i.e., June) liabilities data are more appropriate. Results for June liabilities scaled by lagged positions are also in Panel A (third row); full details comparable to those in Panel B are in OA Table 1. Panel B (columns 4-6) presents detailed results from our baseline aggregate data set, which uses June aggregate liabilities and scales by average positions.

Focusing first on Panel A, compared to scaling by average positions (first row), aggregate AM RDs using lagged positions (the last two rows) are large—3.1% and 2.9%, respectively—because claims returns are high (8.7%), driven by high returns on equity claims (10.3%). Because scaling by lagged positions in a rising market could result in upwardly

[7]Identical sample periods for security-level portfolios are not possible given the timing of the annual surveys. The nonoverlapping periods—the second halves of 2003 and 2022—do not materially affect the results as our main points hold if we start the liabilities portfolio a year later (June 2004) or end the claims portfolio a year earlier (December 2021).

[8]Claims data, whether aggregate or security-level, are measured as of December throughout this paper.

biased returns estimates (when using equation 1), we regard the smaller (1.9% in row 1) differential based on average positions as more accurate.

Our baseline synchronized aggregate AM differential (Panel B, columns 4-6) is 156 basis points (bps), with U.S. claims earning 6.91% per year and U.S. liabilities yielding 5.36%. Security-level data (col. 1) indicate a smaller differential of 93 bps.[9]

While a large difference between RDs computed using aggregate and security-level data can arise from using reported aggregate CY data, scaling by POS_{t-1}, or both, our baseline aggregate measure produces a more modest difference of 63 bps. We next explore reasons for any difference.

2.2 Reasons Aggregate and Security-Level RDs Differ

We first explore whether the difference between aggregate and security-level RDs had persisted or emerged only recently by examining the year-by-year cumulative impact (each year's difference divided by 19 and then cumulated). The last bars in Figure 1 (top left), for 2022, correspond to differentials reported in Table 1 columns (1) and (4). For example, the last "Total" bar is the aggregate minus security-level total differential: 1.56% minus 0.93%, or 0.63%. The impact graph yields at least three findings: the full-sample differences between aggregate and security-level RDs were already evident by 2019, as the 2019 difference in total RDs, 0.55%, is close to the full sample result; much of the difference in total RDs occurred between 2013 and 2019 (aggregate RD remained below the security-level RD as late as 2012); and equity and bond dynamics through time are broadly similar to "total".

We next note that the difference between aggregate and security-level RDs is due to differences in returns on U.S. liabilities—security-level data indicate a 6.08% annual return compared to 5.36% using aggregate data—and in particular equity liabilities (11.04%

[9]Bertaut et al. (2024), using a similar security-level dataset, report an AM differential of 0.5% for 2004-2022, close to our estimate (0.62%) for the same period; see OA Table 4.

vs. 8.62%).[10] As shown in Figure 1 (bottom right), the higher security-level returns on equity liabilities are driven by a few years: 2009, 2014, and 2021 account for 1.4 of the 2.4-percentage-point difference. A composition effect discussed later also contributes: foreign investors held U.S. equities that earned higher returns than the aggregate data imply.

Notably, aggregate income returns in Table 1 are higher than their security-level counterparts for total, bond, and equity portfolios on both the claims and liabilities sides. We explore this in detail when we present IRRs.[11]

Finally, claims returns are more volatile than liabilities returns, as evident in Figure 1 and shown in the italicized *"var"* rows in Panle B of Table 1. This higher volatility is plausibly due to exchange rate fluctuations.[12]

3 Returns Differentials Using Realized Returns

3.1 Returns Measures: AM, GM, and IRR

AM is commonly used in many areas of research. For example, the exorbitant privilege literature, from Gourinchas and Rey (2007) and Curcuru et al. (2008) through Bertaut et al. (2024), has focused exclusively (to our knowledge) on AM. Bessembinder et al. (2025) discuss AM's widespread use in empirical finance research.

However, AM is not a useful measure of realized returns over long horizons. Bessembinder et al. (2025) describe AM as "cash-account" returns, because the returns produced by the AM require a constant investment amount; to begin each period with the same investment amount, each period's returns must be swept into a cash account that earns a zero rate of

[10] Returns on U.S. claims are similar: 7.01% with security-level data and 6.91% with aggregate data.

[11] See BEA (2024) for its methodology on computing returns and dividend yields. These calculations use one or more country-level indices, such as the S&P 500 for U.S. equity liabilities.

[12] We cannot verify this using survey data, which capture dollar amounts (essentially, price plus currency). But index data suggest the difference is due to exchange rate volatility: The S&P 500 is less volatile than the MSCI ex-U.S. index in U.S. dollars, consistent with our results, but more volatile than MSCI ex-US in local currency.

return (and any losses must be replenished from that account). Another measure of returns, the geometric mean (GM), is also known as the buy-and-hold return and is frequently used when reporting multi-year returns. Dichev (2007) calls GM "securities" returns, as they are the returns on the securities, but not returns earned by investors (unless they buy and hold for the entire sample).

AM and GM are limited measures of realized returns: the former produces a somewhat unhelpful "cash-account" return, while the latter measures returns on the securities rather than on the portfolio unless no trading occurs during the sample. That said, both are easily computed: For an annual returns series such as ROR_t, sum the T observations and divide by T to obtain the AM, or take the Tth root of the compounded return to obtain the GM:

$$\mathrm{GM} = \left(\prod_{t=1}^{T} (1 + \mathrm{ROR}_t) \right)^{\frac{1}{T}} - 1$$

Realized returns can be computed with the less commonly used "dollar-weighted" return, or IRR, that takes into account the timing and magnitude of income streams and net purchases.[13] In our setting, computing IRRs requires four components of cross-border portfolios: the market value of initial positions (the initial outlay), period-by-period flows (interim outlays), period-by-period income streams (coupon payments for bonds, dividend payments for equities), and the market value of terminal positions (the final payout). Specifically, IRR computed over T periods is:

$$C_0 + \frac{C_1}{(1 + \mathrm{IRR})^1} + \frac{C_2}{(1 + \mathrm{IRR})^2} + \cdots + \frac{C_T}{(1 + \mathrm{IRR})^T} = 0$$

Alternatively:

$$\sum_{t=0}^{T} \frac{\mathrm{C}_t}{(1 + \mathrm{IRR})^t} = 0$$

Wwere:

[13] On IRRs, see Fabozzi (2014) for a textbook exposition, Dichev (2007) and Tabova and Warnock (2026) for applications to equities and Treasuries, and Bessembinder et al. (2025) for a discussion of other measures.

- C_t: net cash flow at time t
  - C_0: initial investment, equal to the negative of the market value of holdings at initial period 0 (2003)
  - C_t: intermediate distributions and contributions; for each period after the initial investment and before the final payout, this cash flow equals net sales (i.e., the negative of net purchases) plus coupon or dividend payments.
  - C_T: final payout, equal to the market value of holdings at the end of the sample (2022) plus and the coupon or dividend payments and net sales from period T-1 to T.

Using security-level data, we can observe or readily compute all four components. For all bonds (U.S. Treasury bonds, U.S. agency bonds, U.S. corporate bonds, and foreign bonds) we use security-level data on each bond's market value, face value, and coupon yield to compute net purchases (the change in the face value of holdings) and coupon payments (face value times the coupon yield). For U.S. and foreign equities, security-level net purchases are calculated as the change in the number of shares times price, while dividend payments are calculated as the dividend, reported as dollars per share, times the number of shares. With aggregate BEA data, positions come from the IIP, while income streams and net purchases come from the BOP. For foreigners' U.S. portfolios we construct aggregate IRRs using June-to-June data.[14] See OA for details.

3.2 Realized International Returns Differentials

We discuss estimates of realized RDs in two steps, beginning with a comparison of AMs and GMs. While neither is our preferred measure of realized returns, both are widely used.

[14]For quarterly series, we compute annual series measured as of June, as discussed around equations (1) and (2). The only other requirement is an initial (June 2003) position; we use the survey amount.

GM-based RDs are 0.16% with aggregate data and -0.05% with security-level data (Table 1, bottom row, columns (8) and (7), respectively). These are 140 bps and 98 bps, respectively, below the corresponding AM differentials. The reason for the sizable gap between AM and GM differentials is straightforward: Claims returns, which include exchange rate effects, are more volatile than liabilities returns (variances of 2.34 vs. 0.46, in italics in Table 1), and GM is approximately equal to AM minus half the variance of the underlying returns series Fisher (1966). Because GM declines more as volatility rises, the GM for claims falls more than the GM for liabilities; consequently, the GM-based RD is much lower than the AM-based RD. At the asset-class level, equities are more volatile than bonds (e.g., the variances of liabilities are 2.58 for equities and 0.16 for bonds), so GMs decrease substantially for equities and very little for bonds.[15] The same relationships—a lower overall GM-based RD because claims returns fall more than liabilities returns, similar AM and GM values for bonds, and a larger gap for equities—are evident in the aggregate data (cols. 4 and 8).

Our preferred measure of realized returns, security-level IRRs (Table 2, column 1), also produces a negative RD. U.S. investors' realized annual returns on their foreign portfolios were 5.14%, while foreigners earned 5.60% on their U.S. portfolios, for an overall U.S. realized RD of *negative* 0.45%. The composition and timing effects that create this negative differential are discussed Section 3.2.2. For now, two features of within-asset-class security-level IRRs are notable: equity returns exceed bond returns, and U.S. investors earned more than foreign investors on cross-border bond portfolios (3.41% per year for bond claims vs. 2.71% for bond liabilities) but far less in their cross-border equity portfolios (5.79% on claims vs. 9.90% on liabilities).

IRRs computed using aggregate data (col. 2) provide a very different result: an overall differential of 1.41%. Aggregate and security-level *liabilities* IRRs are similar (5.51% and 5.60%, respectively), but aggregate *claims* IRRs are much larger than security-level IRRs

[15]Income returns have small variances for both equities and bonds.

(2 percentage points higher for equities and almost 1 percentage point higher for bonds). In the next subsection, we explore the sources of this discrepancy between aggregate and security-level IRRs.

3.2.1 Differences between Aggregate and Security-Level IRRs: Income Streams and Flows

Recall the components of the cash flows in an IRR calculation: the change in positions (ending minus initial), interim net purchases (which are outlays) and interim income streams (from dividends or coupons). IRRs increase with some combination of a larger change in positions (from initial to final), fewer interim net purchases (i.e., more interim net sales), and larger interim income streams. Table 2 (cols. 3-6) shows a decomposition of the difference in the cash-flow components.

The large difference in security-level and aggregate IRRs is not due to differences in changes in positions. Over our sample period, BEA based its IIP positions on aggregates of the security-level data. The small differences arise only from securities omitted because of insufficient information and therefore have little effect on IRRs (Table 2 column 4; see time series in OA). Rather, the difference in IRRs is due to differences in income streams and flows.

For income streams (Figure 2 and column 6 of Table 2), BEA income streams consistently exceed our security-level estimates. The differences are sizeable and material. Over the sample period, cumulative BEA-reported income streams exceed our security-level estimates by $573 billion for liabilities and $1016 billion for claims; in 2022, each was $125 billion higher. These are line items in the primary income section of the U.S. current account. The source of the discrepancy is clear: BEA computes income streams by applying index yields—for example, the S&P 500 dividend yield for U.S. equities and yields from various ICE indices for U.S. bonds—to the dollar amount of cross-border positions. This technique

is accurate when the within-asset-class composition of cross-border positions resembles that of the selected index. However, many securities in cross-border portfolios generate income streams that differ substantially from those implied by index yields.

Consider equity liabilities as an example. The discrepancy arises largely from differences in the treatment of holdings with zero or missing dividend data. This does not affect common and preferred stocks (Types 1 and 2 in the survey), which are about 80% of the sample (by holdings; see OA Figure 3) and have a 99% match rate with external data (CRSP) that provide dividends data as well price data that we use to check, verify, and correct implied prices available in the surveys. For common and preferred stocks, the SP500 yield is an appropriate approximation. The question is whether SP500 yields are appropriate for the other 20% of holdings: security Types 3 and 4 (fund shares and other equities), groups that include private equity, hedge funds, mutual funds, and other vehicles. We are able to match the majority of these and conclude that although an occasional entity pays a dividend, zero is the best estimate. We therefore treat both missing and reported-zero dividends in the survey as true zeros.[16]

There are similar issues with equity claims (with Types 3 and 4 equities) and zero-coupon bonds; details are in OA. Overall, we are comfortable with our income series and believe they provide a straightforward improvement for official BOP data.

For flows, we compute flows from security-level information described above (change in face value for bonds; for equities, change in the number of shares times price). Before 2013, BEA transactions data were primarily based on TIC S flows, which were so inaccurate that they were discontinued in 2023. Since 2013, BEA imputes transactions from reported TIC positions, which requires estimating valuation changes based on aggregate indexes.[17]

[16]Details on the composition of Types 3 and 4 and the composition of missing dividends are in OA Tables 6 and 7. Income streams applying SP500 dividend yields to missing values, shown in Fig. 2 (top-right panel), are close to BEA's reported series for most of our sample.

[17]For details on the relationship between BEA and TIC S flows before 2013, see Bureau of Economic Analysis (2019). For details on how BEA imputes flows since 2013, see Bureau of Economic Analysis (2023).

Figure 3 shows that, relative to security-level flows, aggregate flows are very similar for bond liabilities, materially higher for equity liabilities, and much lower for assets—especially equities. These differences translate into aggregate IRRs that are slightly lower for liabilities and much higher for claims (Table 2, column 5). While we are confident in our security-level data, we estimate our equity flows, and aggregate flows are no longer collected.[18]

On net, smaller net purchases and larger income streams in the aggregate data produce a much larger U.S. RD.

3.2.2 Reasons the Realized Returns Differential is Negative: Composition and Timing

Realized returns depend on composition—both the securities held within each asset class and the allocation across asset classes—and on timing.

The within-asset-class composition effect is evident in GM returns (Table 1, col. 7, first four rows), which are "securities" returns, i.e., the buy-and-hold returns the securities would provide. GM returns are much higher on equity liabilities (9.91% per year) than on equity claims (6.46%); this gap more than offsets the difference between bond returns for claims (3.69%) and liabilities (3.12%).[19] Within asset classes, therefore, RDs reflect very strong U.S. equity returns that more than offset the modestly higher returns on foreign bonds relative to U.S. bonds.

The asset allocation aspect of the composition effect is determined by the relative weights of lower-return bonds and higher-return equities in claims and liabilities. U.S. international assets have been predominately in equities—about 70% over the past two decades—while

[18]The biggest difference in flows is for equity claims: aggregate flows average $58 billion a year less than security-level flows. As a check, we estimated the IRR for equity claims using what we call BTBJ flows—the spliced Bertaut and Tryon (2007) and Bertaut and Judson (2014) flow series—which we show as a third line in Figure 3 (bottom right graph). The IRR is very similar to our reported result.

[19]GM returns also help explain why security-level and aggregate AMs for equity liabilities differ in Table 1 and OA Table 1. Aggregate "securities" returns for equity liabilities, scaled by average or lagged positions, are 7.66% and 9.22%, respectively, compared with 9.91% in the security-level data and 9.73% for the broader CRSP index.

foreigners' U.S. debt-equity mix has changed dramatically (OA Figure 4). Foreigners' U.S. securities holdings had long been skewed toward bonds, with equities comprising about 30% from 1980 through 2011, but more recently foreigners' U.S. equity positions have surged to match their U.S. bond holdings. While this surge in the share of equities in foreigners' U.S. portfolios is striking, during our sample period the changing debt-equity ratios in liabilities and claims have had a modest effect on the RD. The "fixed-weight" RDs in Table 1, formed by fixing the bond-equity ratios at their initial (2003) values, are only about 25 bps higher than RDs that use actual weights.

The effect of the timing of net purchases is measured by IRR (realized return) minus GM (buy-and-hold return); if investors timed purchases well (e.g., by buying just before strong returns), IRR would exceed GM.[20] For security-level returns, comparing the IRRs in column (1) of Table 2 with the GMs in column (7) of Table 1 indicates that U.S. investors' timing reduced their equity returns by 67 bps per year (5.79% IRR minus 6.46% GM). In contrast, timing eroded foreigners' U.S. equity returns by only 1 bp. The discrepancy in U.S. and foreign investors' timing reduces the U.S. RD by 40 bps and accounts for the difference between the near-zero GM differential (-0.05% in Table 1) and the realized differential of -0.45% (the IRR in Table 2).

4 Implications for the Exorbitant Privilege

Our results suggest that the U.S. international RD is much lower than previously thought. Over 2003-2022, the realized-returns differential (based on IRRs) was negative 0.45%. This result alone should prompt a reassessment of our understanding of the broad notion of exorbitant privilege.

Our results do not speak directly to the narrow notion of privilege, which concerns

[20] See, for example, Morningstar's annual Mind the Gap series (https://www.morningstar.com/business/insights/research/mind-the-gap).

the ability to issue debt at low rates rather than cross-border returns. On that, note that government borrowing costs are not measured by AM or GM; they are measured by a market IRR. Tabova and Warnock (2026) estimate that the Treasury market's IRR is 113 bps lower than its GM. Researchers might find a similar relationship in foreign bond markets. For now, index returns for 1986-2023 suggest a modest 40 bps gap between the GMs of global ex-U.S. bonds (6.49%) and Treasuries (6.09%).

Our results can also inform IIP projections. Atkeson et al. (2025) introduced the term ex post privilege, but most studies of exorbitant privilege focus on expected returns. Our analysis provides one such estimate. An unbiased estimate of future buy-and-hold returns is the weighted average of AM and GM (Table 1, columns 1 and 7), where the weight on GM equals the ratio of the investment horizon to the sample period (Jacquier et al. 2003; Bessembinder et al. 2025). Although our 19-year sample is short, this method produces an estimated U.S. international RD of 0.44% over the next 10 years, much lower than the aggregate AM (1.56%).

This, in turn, has implications for the future path of the IIP, and therefore bears on the sustainability of the large U.S. IIP deficit. Assuming 4.5% annual nominal GDP growth, aggregate data and the literature's historical focus on AM imply that the portfolio portion of the U.S. NIIP deficit would change little as a share of GDP over 10 years, widening slightly from 38% in 2022 to 39.8%. In contrast, using the security-level data and the weighted average of AM and GM to estimate future expected returns, the portfolio portion of the deficit would increase substantially to 55.5% of GDP.

5 Conclusion

The belief that U.S. investors earn substantially more on foreign securities than foreign investors earn on U.S. securities has been central to the notion of exorbitant privilege. Ag-

gregate data give the appearance of a positive and sizable U.S. RD. But high-quality security-level data show that over the past two decades foreign investors earned higher returns on U.S. portfolios than U.S. investors earned on foreign portfolios. Accordingly, researchers should not assume a large positive international RD on portfolio debt and equity securities.

References

Atkeson, Andrew, Jonathan Heathcote, and Fabrizio Perri (2025). The end of privilege: A reexamination of the net foreign asset position of the United States. American Economic Review 115(7): 2151-2206.

Bertaut, Carol C., Stephanie Curcuru, Ester Faia, Pierre Olivier Gourinchas (2024). New Evidence on the US Excess Return on Foreign Portfolios. International Finance Discussion Paper 1398.

Bertaut, Carol C., and Ruth A. Judson (2014). Estimating U.S. Cross-Border Securities Positions: New Data and New Methods. International Finance Discussion Papers 1113. Board of Governors of the Federal Reserve System (U.S.).

Bertaut, Carol C., and Ralph W. Tryon (2007). Monthly Estimates of U.S. Cross-Border Securities Positions. International Finance Discussion Papers 910. Board of Governors of the Federal Reserve System (U.S.).

Bessembinder, H., Chen, T-F, Choi G., Wei, K.C. (2025). How Should Investors' Long-Term Returns be Measured? Financial Analysts Journal.

Bureau of Economic Analysis (2019). How BEA Aligns and Augments Source Data from the U.S. Treasury Department for Inclusion in the International Transactions Accounts. Survey of Current Business 99(7), https://apps.bea.gov/scb/2019/07-july/0719-treasury-international-transactions.htm

Bureau of Economic Analysis (2023). Technical Note for U.S. International Transactions, 2nd Quarter 2023. https://www.bea.gov/sites/default/files/2023-09/trans223-tech.pdf

Bureau of Economic Analysis (2024). U.S. International Economic Accounts: Concepts and Methods. https://www.bea.gov/resources/methodologies/us-international-economic-accounts-concepts-methods

Curcuru, Stephanie, Tomas Dvorak, and Francis E. Warnock (2008). Cross-border Returns Differentials. Quarterly Journal of Economics 123(4): 1495 - 1530.

Curcuru, Stephanie, Tomas Dvorak, and Francis E. Warnock (2010). Decomposing the U.S. external returns differential. Journal of International Economics 80: 22 - 32.

Curcuru, S., C. Thomas, and F. Warnock (2013). On Returns Differentials. Journal of International Money and Finance 36: 1-25.

Dichev, Ilia (2007). What Are Stock Investors' Actual Historical Returns? Evidence from Dollar-Weighted Returns. American Economic Review 97(1): 386-401.

Eichengreen, Barry (2011). Exorbitant Privilege: The Rise and Fall of the Dollar and the Future of the International Monetary System. Oxford University Press.

Fabozzi, Frank (2014). Bond Markets, Analysis, and Strategies. 9th ed. Pearson.

Fisher, L. (1966). Some new stock market indices. Journal of Business 34: 191-224.

Gohrband, C., Howell, K (2015). U.S. international financial flows and the U.S. net investment position: new perspectives arising from new international standards. In: Hulten, C., Reinsdorf, M. (Eds.), Measuring Wealth and Financial Intermediation and Their Links to the Real Economy. Studies in Income and Wealth 73 (University of Chicago Press).

Gopinath, G., Stein, J. (2021). Banking, Trade, and the Making of a Dominant Currency. Quarterly Journal of Economics 136(2): 783–830.

Gourinchas, P., Rey, H. (2007). From world banker to world venture capitalist: the U.S. external adjustment and the exorbitant privilege. In: Clarida, R. (Ed.), G7 Current Account Imbalances: Sustainability and Adjustment. University of Chicago Press, Chicago, pp. 11–55.

Gourinchas, P., Rey, H. (2022). Exorbitant privilege and exorbitant duty. CEPR Discussion Paper No. DP16944.

Griever, William, Gary Lee, and Francis E. Warnock (2001). The U.S. System for Measuring Cross-Border Investment in Securities: A Primer with a Discussion of Recent Developments, Federal Reserve Bulletin 87: 633-650.

Itskhoki, Oleg, Dmitry Mukhin (2025). The Optimal Macro Tariff. NBER Working Paper 33839.

Jacquier, E., A. Kane, and A. Marcus (2003). Geometric or Arithmetic Mean: A Reconsideration. Financial Analysts Journal 59: 46-53.

Kekre, Rohan, and Moritz Lenel (2024). The flight to safety and international risk sharing. American Economic Review 114(6): 1650-1691.

Lane, P., Milesi-Ferretti, G. (2007). A Global Perspective on External Positions. In G-7 Current Account Imbalances: Sustainability and Adjustment, edited by Richard Clarida, Chicago, University of Chicago Press for NBER.

Lane, P., Milesi-Ferretti, G. (2009). Where Did All The Borrowing Go? A Forensic Analysis of the U.S. External Position. Journal of the Japanese and International Economies 23 (2): 177-99.

Maggiori, M. (2017). Financial Intermediation, International Risk Sharing, and Reserve Currencies. American Economic Review 107(10): 3038-71.

Meissner, C. M., Taylor, A. (2008). Losing our Marbles in the New Century? The Great Rebalancing in Historical Perspective. In Global Imbalances and the Evolving World Economy, edited by J. S. Little. Boston, Mass.: Federal Reserve Bank of Boston.

Milesi-Ferretti, G. (2021). The US Is Increasingly a Net Debtor Nation. Should We Worry? Up Front, Brookings Institution.

Obstfeld, M., Rogoff, K. (2005). Global current account imbalances and exchange rate adjustments. Brookings Papers on Economic Activity 1: 67–123.

Rogoff,. K, Tashiro, T. (2015). Japan's Exoribitant Privilege. Journal of the Japanese and International Economies 35: 43–61.

Tabova, Alexandra, and Francis E. Warnock (2026). Preferred Habitats and Timing in the World's Safe Asset. Journal of International Economics.