LONG-RUN TRENDS IN LONG-MATURITY REAL RATES 1311-2022*

Kenneth S. Rogoff† Barbara Rossi‡ Paul Schmelzing§

First version: September 2022. Revised version: July 2023.

Abstract

Taking advantage of key recent advances in long-run financial and economic data, this paper analyzes the statistical properties of global long-maturity real interest rates over the past seven centuries. In contrast to existing consensus, which has overwhelmingly concentrated on short samples for short-maturity rates, we find that long-maturity real interest rates across advanced economies are in fact trend stationary, with moderately rapid convergence speeds, and exhibit a persistent downward trend since the Renaissance. We investigate structural breaks in real interest rates over time using multiple statistical approaches, and find that structural breaks are generally rare, with only the periods around the Black Death and the "Trinity default" of 1557 appearing as consistent inflection points – while the evidence on "recent" structural breaks, namely 1914 and 1981, overall appears weaker than existing literature would lead one to expect. We further examine trends in persistence, as well as commonly invoked drivers of global real rates: exploiting significant data advances, we argue that historically, demographic and productivity factors appear to show no promising causal role, and in fact diverge from real interest rates over the long run.

JEL Codes: E4, F3, N20.

*The views expressed in this paper are solely those of the authors. We thank Barry Eichengreen, Helene Hall, Andrew Lilley, Pierfrancesco Mei, Vania Stavrakeva, Jenny Tang, Alessandro Tarozzi, and Pierre Yared, as well audiences at the IMF's 23rd Annual Research Conference 2022 and the 2023 Society for Financial Econometrics Annual Conference, seminar audiences at Catolica Lisbon CUBE, the Bank of Canada, Kansas State University, Harvard Kennedy School, the Hoover Institution Stanford, and the Federal Reserve Board for comments, and Jiaming Huang, Shereein Saraf and Daniel Smith for excellent research assistance. We thank the editor, as well as four anonymous referees for their suggestions.

†Harvard University.

‡ICREA-Universitat Pompeu Fabra, Barcelona School of Economics and CREI.

§Boston College, and Hoover Institution Stanford.

1 Introduction

More than a decade after the Global Financial Crisis (GFC), and following a sharp rebound in economic activity from a worldwide pandemic shock, global real interest rates remain at historically depressed levels. Although the fall in real interest rates has been ongoing since the peaks of the 1980s disinflation, the drop was particularly acute after the GFC, leading to an intense academic and policy discussion over which factors might best explain it. Despite disagreements over specific drivers, many researchers have posited that current real interest rate dynamics can be traced to a key inflection point during the 1970s or 1980s, when the transition from fixed exchange rates coincided with a deregulation spurt in international finance and an acceleration of globalization, culminating in the "great moderation" phase (Rachel and Summers, 2019). Other contributions emphasize structural global factors, including financial flows (Bernanke, 2005), demographics (Goodhart and Pradhan, 2021), trend growth (Holston et al., 2017), or monetary policy (Garcia and Perron, 1996; Bernanke, 2004). Summaries are provided by Blanchard (2023) and Lunsford and West (2019), with the latter finding the strongest positive association between demographic factors and the safe rate between 1890-2016. Overall, however, the literature has generally taken as given that most of the post-GFC plummet in real rates reflects latent secular forces, and is likely to be permanent.

This paper will take advantage of recent advances in long-run financial and economic data to approach the debate from a fundamentally different angle – investigating real interest rate dynamics from a multi-century perspective. To preview our main results, the primary finding of the paper is that long term real interest rates have been on a persistent downward trend since the 1300's (about seven hundred years), and are stationary around this linear trend, with half lives typically ranging from one to three years. This is an important finding. First, it calls into question the view that declines in interest rate are a more recent phenomenon. Because that has been the prevailing view, the explanations for low real rates have centered more around changes in the economy in the last several decades to match the fall in interest rates (secular stagnation, demographics, lower economic growth, etc). This paper basically argues those explanations are very far from the full picture, and their importance for recent declining rates possibly overstated. In fact, for most of history, real interest rates have not just consistently trended down – but also trended counter to output and population growth.

Beyond qualifying the debate about trends and drivers for recent decades, this paper is motivated by a literature that sought to characterize the fundamental statistical properties of (real) interest rates, and their relation to other macroeconomic variables. For both areas of interest, (very) long-run data sets can potentially offer far superior results compared to an existing literature that has typically focused on relatively short sample periods, typically 75 to 150 years at most (typically less). First, it is clear that a multi-century overview encompasses a much richer tail event, monetary, and fiscal variation over time than an exclusive post-War, or post-Bretton Woods assessment – and this longer perspective is the obvious desideratum to assess "secular" forces in the global economy, and the structural impact of "generational" macro shocks like the GFC or

COVID. However, such investigations were until recently impossible given a lack of adequate data.

Second, there are key econometric reasons that recommend a very long-run examination of the statistical properties of real rates: most importantly, it is known that the power of unit root and other statistical tests improves significantly over longer sample periods; the advantages remain even relative to increasing the number of observations in a given (shorter) sample using high frequency data (Perron, 1989). Frankel (1986) in his exposition on purchasing power parity (PPP) demonstrated that the rejection of the random walk for U.S.-U.K. exchange rate time series crucially depended on the time horizon: the statistical power significantly improved once horizons of over 100 years were chosen, finally allowing for a rejection of the unit root.[1] Therefore, even though recent contributions have expanded the time span of statistical real interest rate assessments into the 19th century (Hamilton et al., 2016), ideally longer data sets should be used. Having a much longer data set -– multiple centuries -– can also raise the bar on what constitutes a significant structural break. Fluctuations that appear to represent structural breaks in the context of 50-75 years of data, may no longer appear so in a multi-century data set.

Third, existing analyses of real interest rate dynamics overwhelmingly focus on short-maturity interest rates, with very few exceptions. In fact, virtually all main reference contributions have reached generalizations on interest rate properties on the basis of central bank policy or bill rates. Much, however, suggests that examining *long-maturity* interest rates has considerable advantages. Long-maturity financial instruments – perpetual consols, for instance – have been in existence far longer than either short-term consolidated debt or rate-setting policy institutions, with consolidated sovereign long-maturity debt dating back at least to the first Venetian issuance of negotiable *Monte Vecchio* bonds in 1262 (Pezzolo, 2003; Stasavage, 2011; Eichengreen et al., 2021). Historical studies have repeatedly demonstrated that as early as the 14th century, households and institutions treated sovereign long-maturity assets as a relatively "safe asset", to which they increased financial exposure in times of crisis (de Moor and Zuijderduijn, 2013).[2] Moreover, compared to short-maturity rates, long-maturity rates arguably better reflect structural growth, rare disaster, inflation, demographic, and capital flow trends. For the assessment of "secular", "generational" forces affecting large parts of the advanced international system over long investment horizons, therefore, the focus of the existing literature on short-maturity rates over relatively limited time frames seems potentially problematic, despite arguable advantages of using short-maturity rates when only

[1]For a practical example, see Lothian and Taylor (1996). The authors there show that the existence of a unit root for real exchange rates cannot be rejected over the sub-periods of 1946-1990, and 1974-1990 - but this result disappears when the sample length is extended to 1803-1990, for which the series can clearly reject the unit root. The authors use a range of other random 45-year intervals, for which the rejection of unit root is also impossible: this strongly underscores the role of sample length, rather than sub-sample-specific features involved in the post-1945 episodes. We test our basic conclusions for different sample lengths in the appendix in an analogous way, an exercise which confirms the importance of sample length for real interest rates.

[2]de Moor and Zuijderduijn (2013) recently showed that Dutch households systematically increased exposure to long-maturity annuities in times of heightened pandemic, natural disaster, or war frequencies – at the expense of exposure to livestock, houses (though not land), and cash. Mulho (1971) and Mueller (1997) showed separately that significant shares of Venetian and Florentine debt in the 14th and 15th centuries was foreign-owned, bought by European investors in search for financial safety.

focusing on business cycle horizons.[3]

With recourse to a new multi-century real interest rate data set constructed by Schmelzing (2023) and based on long-maturity market rates, this paper will seek to analyze the dynamics underpinning both the secular evolution, as well as the present features of the global interest rate environment.[4] Importantly, however, we will confirm that our econometric results also hold for alternative independently constructed long-run series, including the shorter (by a few centuries) one-country data set compiled by Dimsdale and Thomas (2016). Following a literature review, section three of this paper will describe the new long-run interest rate data, with section four proceeding with an analysis of the general statistical properties of global real rates since the birth of sovereign debt in 1311.

There, we will present three key statistical results. First, we show that – in contrast to the existing consensus – we are able to reject a unit root at conventional significance levels once one allows for a deterministic trend and long-maturity real interest rate data are employed: for both headline global real rates, as well as for all advanced economy country-level series, this result is shown to hold for long samples, and confirmed for a range of test specifications: including for a variety of inflation expectation approaches, and for sub-samples that exclude the very early years of the sample. Global real rates exhibit a downward trend (averaging about 1.7% every 100 years), and the evidence appears strongly consistent with the hypothesis that they are stationary around this trend. We confirm this result in both global and individual country data, using both ADF-GLS tests and the Zivot and Andrews (1992) test, which is robust to the presence of instabilities. Nevertheless, we can replicate the non-stationarity findings for the much shorter data series available for short-maturity rates even when a trend is included. Our finding that long-term real rates are stationary is important because, while theoretical models typically assume stationary real rates, so far the existing empirical evidence overwhelmingly rejected their stationarity: our empirical results thus resolve this puzzle.

Second, using structural break tests, we investigate whether structural breaks are present in our data, testing both methodologies that allow an open-ended identification of break dates (Bai and Perron, 1998), as well as alternatives that test the validity of particular potential break dates (Chow, 1960), which we chose on the basis of historical priors. The Bai and Perron (1998) test points

[3]These dynamics are underpinned by asset pricing theory: as risk-averse investors move from single-periods to infinite horizons, the riskless asset is no longer a real Treasury bill, but a real perpetuity (Campbell, 2018, chapter 9).

[4]The data carefully isolates sovereign long-maturity interest rates that were based on voluntary debt transactions averaging just over 10-years average maturity over time, with short-maturity, coerced and other involuntary transactions identified separately and excluded. Schmelzing (2020) already provided public data and points out the downward trend in nominal and real rates, but did not investigate econometric properties of the data, including ADF-GLS, half-life, or structural break tests – and also did not test the correlations to growth and demographics, which we undertake in section 5. Importantly, the present paper considerably refines the methodology for forming the long-term inflation expectations measure used to construct long-term real interest rates. In in section 3.1 and appendix section 1, we give an overview of the data construction. An unpublished update of Schmelzing (2020), which slightly refines the empirical basis we utilize here exists in Schmelzing (2023). A future book-length study in Schmelzing (2025) also contains a much more extensive discussion, historical context, and source documentations.

towards a far greater degree of continuity over time, and is generally striking for its absence of (20th century) breaks, across countries and global levels. Meanwhile, the alternative Chow (1960) approach also confirms that among the rare relevant breaks appears to be the Black Death, as well as 1557 (the latter marking the onset of a largely forgotten global financial crisis that "shook Europe to its foundation" (Ehrenberg, 1928, 114) and saw the concurrent default of three of the largest sovereign creditors. For the two key 20th century inflections – 1914 and 1981, the focal point of a large body of recent literature – we find much weaker confirmation across structural break tests, though neither can be dismissed outright.

Third, on the basis of our new data, we also present results on real rate persistence; the finding that real interest rates are trend stationary would be of little practical importance if the half-life of deviations from the trend was 50-100 years. When real rates drop, how long do they take to return to trend? Existing long-run knowledge about the persistence of shocks to financial variables is thus far limited to scattered evidence on foreign exchange rates, and deviations from purchasing power parity. We present the first analysis on real interest rate persistence over the very long-run. Although there is a range of estimates and there are alternative statistical tests available, estimated country half lives are typically on the order of one to three years, not one to five decades. Overall, across global and country-level series, the estimated half life deviations from the long run trend hover around two years over the first six centuries of the sample. After 1914, the estimated half lives appear to increase before moderating after 1980, albeit the orders of magnitude do not change dramatically. Such evidence can contextualize the persistence of contemporary deviations of global real rates from deeper trends, and indicates that while mean-reversion from the sharp falls in global real rates induced by the GFC and the pandemic shock is to be expected, directionally the historical trend does suggests that a downward tendency in global rates will continue: the current expected mean-reversion takes places in the context of a gently but persistently falling trend, rather than in the context of a pre-GFC, or pre-1980s constant value.

Finally, part five assesses the historical and economic context of our statistical results on real interest rate dynamics, in particular evaluating the very long-run relationship with demographics and productivity: we take advantage of the significant advances over recent years in the reconstruction of output growth series by economic historians (e.g. Broadberry et al. (2015); Ridolfi and Nuvolari (2021); Pfister (2022); de la Escosura et al. (2022)) to examine the recent literature on the relationship between real interest rates and both output growth and population growth in a long-run perspective: the evidence leads us to question arguments that posit that the "recent" decline in global real rates can be primarily explained by productivity or demographic trends (as opposed to, for example, increasing bond market liquidity across time, declining risk or outright default, or the way global wealth might affect rates of time preference). It appears that for most of history, real interest rates have in fact trended counter to output and population growth, a finding with important implications not least for standard macroeconomic models (for instance those pertaining to representative agent models).

Taken together, these results suggest far stronger continuities in the present, post-GFC international financial system with (pre-) 20th century dynamics than commonly acknowledged – in other words, the dominant focus on post-1945 dynamics in global real rates and its link to key macroeconomic variables leads to stylized empirical facts and causal associations that are demonstrably very far from the full picture: a long-run view on global real rate dynamics thus calls into question many prevailing theories entertained in recent years. There was indeed a significant downward shift in global real rates after the global financial crisis, and while some of this may represent a continuing unwinding of the 1980s peak, it is also possible that some of it may prove temporary. While previous "low interest rate eras" (low relative to trend) can be identified over time, what we do know is that all of them ended. Based on our evidence, therefore, the leading theory has to be that the post-financial-crisis real long-term interest rate, having shown the secular mean-reversion tendencies after the breakout in the upwards direction during the oil shocks, should be expected to resume its predominant downward trend in the coming years but correct any sharp deviations below trend that the GFC and more recently COVID may have induced.

2 Literature

Repeatedly, the macroeconomic literature has posited that, on purely *a priori* grounds, real interest rates should exhibit mean-reverting, stationary properties. Foundational asset pricing frameworks such as the Black-Scholes relationship or the consumption CAPM assume mean-reversion in real interest rates. Standard finance no arbitrage models of the yield curve also assume that interest rates are stationary (see survey in Neely and Rapach (2008) and discussion in, e.g., Bauer and Rudebusch (2020)).

However, finance and macro literature actually employing statistical tests have had significant difficulties to confirm these a priori assumptions. Instead, they have overwhelmingly been unable to reject non-stationarity in real interest rates.

The earliest contributions focused on nominal rates: Fama (1975) instigated an intense (and broadly dismissive) debate after suggesting U.S. short-maturity bill rates during 1953-1971 were able to predict subsequent inflation rates. Shiller and Siegel (1977) investigated both long-maturity and short-maturity U.K. interest rates' correlation with price level changes – but were not interested in statistical analysis.[5] Afterwards, attempting a falsification of expectations theory based on data over 1890-1979, Mankiw and Miron (1986) narratively defined four interest rate sub-samples and suggested that the founding of the Federal Reserve led to a fundamental change in the statistical properties of short-maturity nominal interest rates. They find that over 1915-1979, short-maturity U.S. nominal interest rates followed a random walk; while they exhibited predictable behavior prior to this inflection date.

[5] Specifically, Shiller and Siegel (1977) use British consol yields between 1727-1973 for long-maturity rates, and they splice a variety of 3-month bill rates together for short-maturity rates, beginning with Overend and Gurney corporate rates over 1824-1844, followed by bank bill rates over 1845-1938. Despite the paper title, the authors do not estimate real rates directly. Siegel (1992) relies on the same nominal basis, adding inflation data to construct ex post measures for the U.S. and U.K.

A key contribution was made afterwards by Rose (1988), who showed that U.S. real interest rates appear to exhibit a unit root: at an annual frequency, Rose used an ADF test for high-grade long-maturity corporate bonds, and for short-term commercial paper rates, for two periods spanning 1892-1970, and 1901-1950.[6] Later, Garcia and Perron (1996) analyzed a regime-based framework of U.S. ex post real bill rates between 1961-1986: the authors conclude that U.S. real rates exhibit an "essentially random" process, with shifting means and variances in the post-war period, though without a clear predictable pattern. Ang and Bekaert (2002) studied U.S., German, and U.K. bill rates and propose that a regime-switching framework best captures out-of-sample observations in post-war real rate data. Later, Rapach and Weber (2004) used an extended unit-root test (Ng and Perron, 2001) to re-examine Rose (1988)'s findings – and confirmed non-stationarity across a range of countries. While the clear majority of existing studies tested short-maturity (real) interest rates, Nelson and Plosser (1982), Perron (1989) and Rapach and Weber (2004) tested U.S. long-maturity bond yields, albeit for a relatively short time period (1900-1973, or in the case of Rapach and Weber (2004) for 12 countries beyond the U.S.).[7]

More recently, Hamilton et al. (2016) confirmed real interest rate non-stationarity, partly using policy and discount rates on the basis of U.S. annual data beginning in the late 19th century, and quarterly data beginning in the 1940s. On an annual level, the authors splice together short-maturity rates for various periods, relying on a mix of commercial paper rates and Fed fund rates for the U.S., for instance. A heteroskedasticity-robust test and KPSS results lead the authors to reject the null hypothesis of stationarity for U.S. short-maturity data at the 5% level: Bai-Perron tests identify 1915 and 1921 as U.S. break dates. The only systematic relationship in the authors' identification relates to a stationary gap between U.S. and world real short-maturity rates.

DelNegro et al. (2019) investigated both global long-maturity and short-maturity interest rates over 1870-2016, finding evidence that real interest rates "fluctuated between 1% and 2% for about a century, rose in the late seventies, and [have] been on a steady decline ever since". The authors rely empirically on the "Jorda-Schularick-Taylor" (JST) database – but do not undertake formal statistical tests, are not interested in linking the contribution to the literature on general interest rate properties, and omit longer periods of time.[8] Finally, Bauer and Rudebusch (2020) propose a time-varying model with shifting endpoints, which the authors suggest is superior in accounting for evidence of the "extreme" interest rate persistence. For ten-year treasury yields on a quarterly

[6]We replicate Rose (1988)'s approach in the appendix (tables A.8.1-A.8.3 and discussion), applying both alternative inflation expectation approaches, for his two periods, and two assets, but also allowing for a trend: we fail to reject a unit root for all 16 variations of the data. Rose sourced historical U.S. inflation and interest rate data primarily from Friedman and Schwartz (1963) and Nelson and Plosser (1982). Non-U.S. data in his analysis begins in 1957 and is tested for quarterly three-month bill rate equivalents, using IMF International Financial Statistics. An earlier paper by Huizinga and Mishkin (1986) used short-maturity U.S. data between 1916-1927, and 1953-1984, and posited a higher degree of (ex post) real rate stability, but did not utilize any ADF, KPSS or related tests.

[7]Perron (1989) relies on Nelson and Plosser (1982)'s data, who fail to reject non-stationarity for U.S. nominal bond yields over 1900-1970.

[8]As we shall discuss in section 3, the JST database only covers long rates from 1870-2016 and, because it draws quite heavily on Homer and Sylla (2005), it inherits many of the its limitations, as discussed in Schmelzing (2023), which uses primary sources, fills in many significant gaps, and covers a much longer time span.

basis between 1971:IV - 2018:I, the authors reject stationarity using ADF and Phillips-Perron (PP) tests.

In sum, despite the strong a-priori presumption that real interest rates should exhibit stationary properties, the existing literature has overwhelmingly failed to reject real interest rates non-stationarity – however, such literature has focused almost exclusively on a combination of spliced short-maturity nominal interest rate data fusing different issuers. And despite multiple studies using timespans of more than one century, their statistical power to examine "secular" properties still remains limited, including the identification of "inflections" that, as we shall see, often wash out at longer term horizons. In sum, substantial extensions of sample spans and a clear focus on long-maturity rates could in fundamental ways increase both the quantitative power and qualitative insights in the context of the extensive debates on real interest rate properties: we turn to precisely such an exercise in the next section.

3 Long-run empirics in global real rates, 1311 - 2022

3.1 A short history of short-maturity debt

To understand existing biases arising from the predominant use of short-maturity data in the literature on the properties of real interest rates, it is useful to revisit the historical background to the evolution of short-maturity and long-maturity debt instruments. We have observed in the literature review above that the vast majority of existing contributions assessing real interest rate properties and trends focus on multi-decade empirics for short-maturity yields, or even outright policy discount rates – often splicing short-term discount and market rates. Perhaps one reason for this existing bias is an implicit assumption that an "expectations hypothesis" (EH) holds – which systematically links all maturities in nominal and real terms across the term structure: under this framework, long-maturity yields are a direct function of continuously rolled-over short-maturity assets, and cannot therefore unduly deviate from the properties of the short end: effectively, expected excess returns on long-maturity bonds over short-maturity bonds are constant over time. In other words, "the risk premium of an n-period bond over a one-period bond (the term premium) is constant" (Campbell et al., 2018, chapter 8.2).[9]

Of course, another major reason for the focus on short-rates is that, in modern times, the (very) short rate is the variable controlled by central banks. This does not, however, seem like a compelling reason to prefer it to longer rates that have a much bigger impact on structural variables in the economy. An additional reason – though rarely stated explicitly – is the exclusive focus of recent "natural" interest rate debates on the short-maturity component, typically referencing Wicksell (1936). For instance, in Laubach and Williams (2003) the natural rate is treated as identical to the real short-term rate: "Since Wicksell (1936), the natural rate of interest—the real short-term

[9]Mainstream affine term structure models equally assume that the expectations hypothesis (EH) holds, see chapter 8.3 in Campbell et al. (2018).

interest rate consistent with output equaling its natural rate and constant inflation—has played a central role in macroeconomic and monetary theory". Later in the same paper (p.1064), the authors justify the focus on short-term rates with the inability to measure long-term inflation expectations adequately. This is of course an important issue, and we devote an appendix section to several alternative approaches offered in existing literature.

Hence, however compelling the choice to focus on short-maturities might at first appear – such a choice crucially restricts the insights that can be derived: as is known among economic historians, *consolidated* sovereign short-maturity assets are historically a decidedly "recent" innovation, only emerging as a regularly-used public debt instrument towards the end of the 19th century. Accordingly, this time is also the point of inception for existing "long-run" investigations on general real rate trends (Hamilton et al., 2016; DelNegro et al., 2019). Short-maturity data points, in other words, even if extended by splicing public and private debt data, are only able to cover a fraction of the full empirical horizon of advanced economy "sovereign debt" trends: hence, the statistical power of short-maturity tests has thus far been dramatically limited. Unfortunately, even key reference works, meanwhile, never reconstructed long-maturity data points on a sufficiently high frequency to allow extensions into longer samples. Only recently, with a new contribution (Schmelzing, 2023) did such an extension become possible for the first time for a broad range of countries – a development we exploit in the subsequent sections (a relevant previous contribution is Dimsdale and Thomas (2016), who constructed annual interest rate data for the U.K. for the years 1703-2016, and we will indeed check all our results using their data and time period, see appendix section 2.4). But before we turn to the new data, we provide the historical context.

On the back of a decree by its Grand Council in 1262 (the *ligatio pecuniae*), the Venetian Republic is typically regarded as the first issuer of perpetual consolidated sovereign debt, following a century of experimentation across Northern Italian city states with creditor consortia, who were given multi-year rights on tax revenues in return for emergency loans. Swiftly, other city states and cities north of the Alps follow suit and consolidate individual *rentes* contracts, so that over the course of the 14th century, a pan-European public debt market in long-maturity assets emerges – see Epstein (2000); Munro (2003); Eichengreen et al. (2021). And the evidence suggests that these markets were remarkably liquid: for instance, by the mid-15th century, the face value of annually-traded long-maturity public debt in hubs such as Venice reaches over 28% of GDP.[10]

By contrast – though they often contracted ad hoc personal loans on short-maturity bases – no sovereign on either side of the Alps regularly issued consolidated short-maturity debt prior to the 18th century. The only outlier may be said to have been the Spanish Empire, which famously experimented with large-scale *Asientos* issuance during the 16th century. The monarchy sharply distinguished this type of issuance from the regular long-maturity *Juros* instruments (on which it never defaulted), and even at the peak of Spanish geopolitical dominance towards the end of the century, the *Asientos* share in total public debt stood at less than 10% (Alvarez-Nogal, 2014).

[10]A figure based on Mueller (1997)'s information that in 1434, 300,000 ducats of *Monte Vecchio* debt in face value traded in Venice. Venetian GDP can be estimated via the data constructed by Malanima (2011).

Regular German, French, Italian, and Dutch consolidated debt issuance up until the 19th century takes place exclusively in long-maturity assets, even though instruments with maturities of more than five years are sometimes anecdotally referred to as "bills".[11] In the U.S., the first Treasury Bill is auctioned no earlier than December 1929 (Garbade, 2008). Prior, several war financing operations see the intermittent issuance of U.S. "certificates of indebtedness", which are documented to have been sold over 1812-5, in 1837, during the Mexican War, in 1857, during the Civil War, and in 1907. For many decades in between these dates, the U.S. government does not issue any negotiable short-maturity debt instruments.[12] The situation is analogous in the other leading advanced economies: the first U.K. Treasury bill issuance is legislated in the 1877 Treasury Bills Act, but only formalized in detail in 1889, and only issued in meaningful volumes from the First World War (BoE, 1963). In dynamics similar to U.S. certificates, between 1707 and 1853, various issuances of U.K. "Exchequer Bills" take place; in effect, revenue stream conversions by the Bank of England, with maturities typically of one year and with peak issuance during wartime (Richards, 1936). Prior to the founding of the Bank of England, "tallies" are regularly issued – intended as short-maturity IOUs funded upon anticipated tax revenues: de facto, the instruments are negotiable, but like Dutch "obligaties" they are in practice assuming effective long-maturity, through automatic roll-overs upon expiry at the prevailing long-maturity rate.

Overall, historically, long-maturity debt thus not only constitutes the "original" sovereign debt – it is the sole historically-consistent basis enabling a multi-century assessment of the time-series properties of sovereign interest rates. In this sense, a consistent and longer long-maturity time series allows much higher statistical power, particularly for example to discriminate between unit root and near unit root hypotheses. This statistical power is what we exploit here in our time series analysis of the new data. Econometric analyses of short-maturity interest rate series are by construction far more restricted (a limitation amplified further for *policy* rates, where time series

[11]For the United Provinces, which "occasionally" issued unfunded short-maturity *obligaties* during the 16th century, which "tended to become long-term" debt since they were continually rolled-over, see Gelderblom and Jonker (2011); for the core Italian States, see Gelabert (1999) and Pezzolo (2008). The Crown of Aragon similarly issues *Prestechs* under Alfons V. and John II., which - while de jure short-term debts - are de facto long-maturity obligations, see Kuechler (1983). For the Holy Roman Empire - whose debt over centuries is synonymous with the Habsburg personal domain ("Erblande") debts, as there is no centralized Treasury – Huber (1893) provides a breakdown, by debt type, for the year 1564: 69.9% of Ferdinand I.'s debt is recorded in the form of long-maturity "pledge" debts, de facto consolidated and transferable instruments, with 30% in "unverwiesene" Schulden – often, but not exclusively short-maturity de jure unconsolidated claims (many of which are treated by all involved sides as de facto long-maturity negotiable claims). Much of sovereign debt activity in the Holy Roman Empire takes place on the level of the sovereign prince-electors: for detailed financial studies, including debt breakdowns, see for instance Schirmer (2006) for Saxony; France intensifies short-maturity issuance under State Minister Jean-Baptiste Colbert and during the 18th century, mainly in the form of *billets* and later by erecting facilities such as the designated *Caisse d'emprunts*: such developments lag regular annuities issuance by two centuries, however, and do not assume comparable volumes, regularity, and credibility: at the eve of the French Revolution, only 7.9% of French public debt outstanding exists in short-maturity form (Legay, 2011, 244). Schmelzing (2025) provides further documentation and an overview of more than 330 short-maturity issuances between 1312-1850, together with nominal interest rates, and more detailed discussion.

[12]For the U.S. history, see in particular Hollander (1919): prior to the first irregular issuance of short-maturity certificates in 1812, the Treasury under Alexander Hamilton actually relies on bank overdrafts and loans to satisfy the short-maturity spectrum.

are even shorter).

3.2 Long-maturity data

We turn to the new data that now allows a critical expansion of interest rate samples, and potentially new insights into statistical interest rate properties. We draw upon Schmelzing (2023), who constructed a comprehensive sample of advanced economy long-maturity liquid interest rates since the "birth of sovereign debt" in the early 14th century.[13] The series covers on average over 80% of advanced economy GDP on an annual level (in no single year less than 65%), across eight countries, and incorporates both marketable (consolidated) and relevant personal (unconsolidated) loan transactions by sovereign creditors. The inception year of 1311 for the series is not accidental: from this date onwards, both voluntary long-maturity nominal data points *and* suitable price data are available, allowing the calculation of inflation-expectations-adjusted, real interest rates.[14]

Using historical data from the 14th century onwards of course raises questions on data quality, and the treatment of potential measurement problems. We stress from the outset that overwhelmingly, such potential sources for measurement errors are concentrated in the early part of our data sample: crucially, however, our results regarding trend stationarity and the relative weakness of structural breaks continue to hold when excluding the first few centuries (for example looking just at post-1703 data as in the Dimsdale and Thomas (2016) U.K. data set – which, like the Schmelzing (2023) U.K. series for the same period contains no interpolations). This applies in particular to the use of linear interpolations: even though the new empirical basis dramatically reduces the necessity for interpolation (the new data set is based on a total of 16,479 nominal long maturity data points, or an average of 4.1 nominal data points per country-year for the period of 1311-1850 alone), linear interpolations remain present in the underlying Schmelzing (2023) and Schmelzing (2025) data. While its overall use is highly minimized, linear interpolations could in principle affect the outcome of our unit root tests: as supported by further tests in the online appendix (online appendix section 3.1), we emphasize that the full exclusion of the first two-hundred years – which are most affected by interpolations in a relative sense – still confirms our stationarity results at the 1% significance levels for all series, as does a test that removes selected country-level sub-periods relatively more affected by interpolations.

A discussion of potential errors remains very instructive, however, and while we elaborate on details in an appendix, it is important to pre-empt key aspects here. First, we note that earlier

[13]A version of key long-term interest rate and inflation data we use was posted and documented by Schmelzing (2020) via the Bank of England in 2020 (https://www.bankofengland.co.uk/working-paper/2020/eight-centuries-of-global-real-interest-rates-r-g-and-the-suprasecular-decline-1311-2018): an update of that data mainly containing refinements is presented in Schmelzing (2023), and earlier long-dated price inflation data based on Allen (2001) and underlying sources can also be found in Reinhart and Rogoff (2009). We elaborate further on data construction details in appendix section 1 below – for an extensive discussion see Schmelzing (2023) and Schmelzing (2025).

[14]Note that given our seven-year lagging approach, excluding the current year t (details below), we utilize new annual data starting in 1311 – most series and tables presented subsequently take 1318 as their starting point, the first observation year incorporating the seven-year lagged inflation value.

periods at times featured significant financial repression dynamics or a variety of distortive attributes such as "in-kind" coupon payments – with transactional nominal sovereign interest rates therefore potentially being far removed from reflecting "voluntary" cash market prices of debt. De jure interest rate ceilings – enacted via "usury laws" – are another feature to take into account, as is the non-negligible role of outright forced loans. To address this issue, the data set systematically researched the context of each debt transaction and excluded all transactions where evidence of severe financial repression, in-kind payments, or other involuntary or non-cash features were uncovered. Specifically, while uncertainty over missing context and sources will naturally always remain, Schmelzing (2023) managed to compile a separate sample of 339 forced sovereign loans for which nominal interest rates can be documented in the eight respective economies, spanning 1311-1944, and spanning different types of "involuntary" transactions – these are separately analyzed but fully excluded from the "voluntary" headline data, as are transactions which feature in-kind payments.[15]

Second, a variety of geographical and political adjustments are taking place over the sample period, but are treated in a standard fashion, generally by following current (2023) political boundaries: while a few relevant regions active in financial markets historically (e.g. the autonomous state of Savoy) can be treated in alternative ways, and within our sample Italy and Germany in particular undergo meaningful centralization dynamics over the sample period, these factors are dealt with in line with established practices, not least in the most recent long-run national accounting literature in economic and financial history.

The effective average maturity of the assets used over time closely approximates modern ten-year long-maturity Treasury bonds (with no relevant time trend underlying this maturity), and draws upon a wide sample of primary, printed primary and secondary sources.[16] It thus represents a

[15]The most frequent forced/involuntary nominal interest rate data point in this separate sample is a transaction at 0%: the sovereign forces a "loan" without even a pseudo-compensation for the creditor risk, though promising to repay the principal. This sample also features transactions of office sales (which are widely considered "loan transactions" in the literature, with the guaranteed revenues linked to the office typically representing the coupon), and loans provided by non-independent governmental institutions, such as early state banks or wealthy government officials: for instance, the information that the Holy Roman Emperor in 1361 "orders" German cities to lend him money at 0% (Nuglisch, 1899, 37f.) is taken as evidence of a forced transaction (and thus excluded from the headline series); but also, for instance, the information that the Bank of England in 1697 lends to the government "under pressure" at 5%, as Acres (1931, 271) reports, serves as reference for such an involuntary data point. Schmelzing (2020) also discusses evidence for in-kind transactions.

[16]No assets with a contracted maturity below two years are included at any point in the data set. The all-time average effective maturity of non-perpetual assets in the sample is stated as 13.2 years: for full details, see in particular the appendix of Schmelzing (2025), available via https://scholar.harvard.edu/pfschmelzing/jmp-dataset. Schmelzing (2020) also attempts to isolate a "safe" interest rate over time, which splices successive country-level long-maturity real rates and traces the ex ante contemporary "dominant" economy, a series which aligns with secondary literature and features no major default event on debt principal over seven centuries. Other "safety" features associated with the interest rate data are tested there as well, including the frequency of debasement operations, military and geopolitical power of the issuing sovereign, and the absolute economic size (output) of the issuing sovereign. While of importance for the eventual identification of positive real rate drivers, neither the existence or the particular level of risk premia over time represent relevant issues for our endeavors in this paper.

significant leap over previous reference works such as Homer and Sylla (2005). Focusing only on the most recent period of 1870-2016, a related undertaking is represented by "JST" (Jorda et al., 2017). As noted earlier, a substantial share of observations in "JST" ultimately traces back to Homer and Sylla (2005), with its associated gaps and inconsistent definitions. Besides covering a much shorter time span than the Schmelzing (2023) data, there are issues of consistency for example where JST/Homer and Sylla fuse yields for both medium- and long-maturity sovereign instruments to create a "long-term bond yield"; moreover data for important countries are not observed directly, but based on estimated risk premia.[17]

On the inflation side, our headline series utilize mainly the reference urban price series constructed by Allen (2001), as well as subsequent country-level improvements (e.g. Alvarez-Nogal and de la Escosura (2013) for Spain; Ridolfi (2019) for France). All data points on the inflation side refer to inflation rates of representative consumption baskets for workers denominated in silver equivalents.

To construct real interest rates, we of course require a proxy for inflation expectations; surveys and inflation-linked bonds are recent constructs, and therefore we follow the approach in Homer and Sylla (2005), who form the real interest rate by deflating the nominal rate in year t with a seven-year lagged inflation measure to approximate inflation expectations (over t-7 to t-1, not including the current year t), using a t-1 weight of 33% and then progressively decreasing the weight by 30% for each year (thus, the t-2 weight is 23%, and for subsequent years 16%, 11%, 8%, 6%, and 3%).[18] To form the first such data points, one requires the annual inflation rates over the period of 1311-1317, and all of our main headline analyses begin with the first such consistently-lagged observations in 1318. We also smooth a few extreme inflation outliers, mainly associated with the 20th century, post-fixed exchange rate system, to avoid the severe measurement issues for these years.[19] Given its relevance, in the appendix we test two alternative approaches to construct inflation expectations, adopting methodologies presented in Eichengreen (2015) and Hamilton et al. (2016). The confirmation of our ADF-GLS results in these variations suggest that the particular inflation expectations construct does not crucially change the properties we posit below. Of course, a virtually unlimited menu of alternative expectations constructions is possible, and the issue merits investigation, but the ones we use are well established in the literature and certainly form a plausible benchmark.

[17] For instance, in Jorda et al. (2017) German, Japanese, and Italian government long-maturity bond yields prior to 1913 are based on an unpublished estimate of risk premia over British consols; later periods for Italy splice together evidence from "medium-term" and "long-term" observations. Similarly, the source for Dutch and German yields between 1880-1913 in "JST" is given as Flandreau and Zumer (2000) – but these authors in turn provide as their source for German and Dutch yields Homer and Sylla (2005) – France over 1870-1947 is also fully based on the latter.

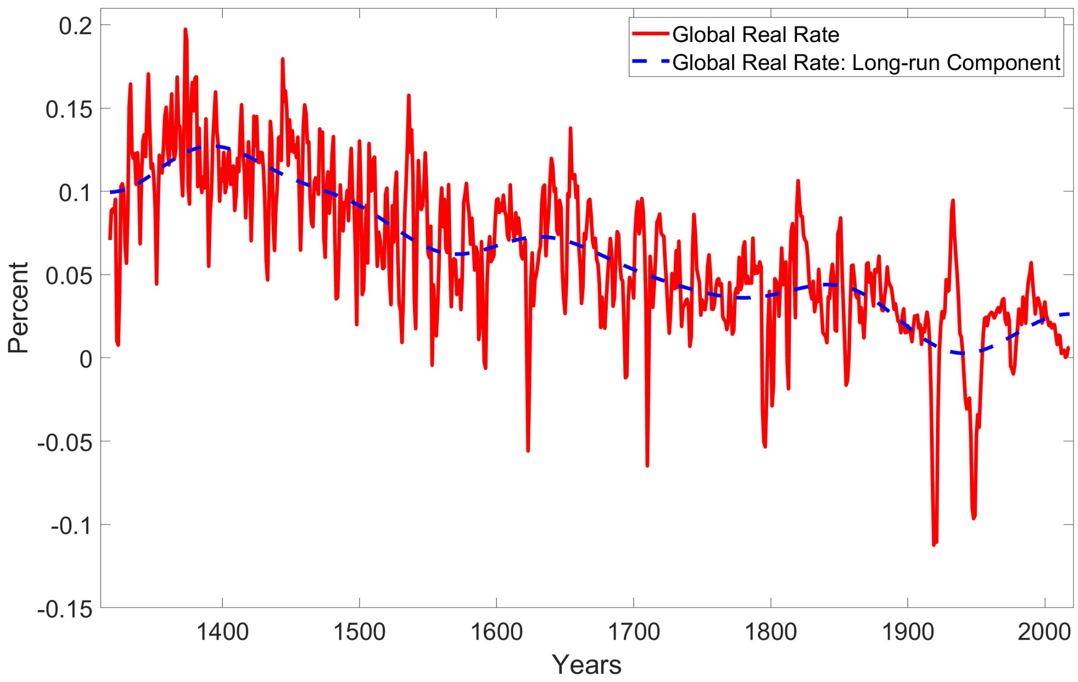
[18] Recall that most previous studies – including Rose (1988) – treat realized past inflation as a proxy for inflation expectations.

[19] In particular, in addition to equal weights, we tested different progressive weightings including the one followed by Homer and Sylla (2005). None of the key results are sensitive to such variations in the initial weight. For the severe inflation shocks, we interpolate the last "pre- shock" CPI figure to the first "post-shock" CPI figure for the hyperinflation years in Germany (1916-1923), Spain (1937-1939), Italy (1943-1945), and Japan (1944-1947).

Taken together, our data therefore traces the only consistent and high frequency definition of sovereign real interest rates over centuries, allowing us to take full advantage of the significantly higher statistical power associated with a multi-century sample length. Figure 1 displays the resulting headline global real rate series on this basis, together with its long-run trend over time. The long-run trend is obtained by filtering the series and retaining only fluctuations above 100 years, based on Müller and Watson (2018). The GDP-weighted basis, weighting each of the eight economies according to its rolling GDP share, is our default basis to report "global" results, unless otherwise noted (this basis is abbreviated "Global GW" in the following). All our empirical tests will also be reported for the eight individual countries in addition to the summary global results.[20]

We note that the dashed structural trend line in Figure 1 already suggests visually that the series appears to exhibit a downward trend over time: apparently, the oft-invoked low interest rate environment of recent decades therefore has deep historical roots. Specifically, the series displays a linear downward trend slope that declines by just over 1.7 basis points (1/100th of one percent – thus amounting to a decline of 1.7% every century) per annum over the observation period (1318-2022). But the next section will proceed beyond "visual" impressions and investigate statistical properties in more detail.

Figure 1: Headline global real rates, and its long-run component, 1318-2022



Notes: Data based on Schmelzing (2023)'s GDP-weighted global real rate data. Long-maturity ex ante basis, deflated by using seven-year progressively-lagged inflation, excluding current year *t*. The dashed line reports the long-run filtered component, isolating fluctuations above 100 years, based on the Müller and Watson (2018) estimator.

[20]The long-run trend is based on Müller and Watson (2018); details on its construction are provided in Section 5.

4 Statistical Properties

We undertake an Augmented Dickey-Fuller test (ADF-GLS), as modified by Elliott-Rothenberg-Stock (Elliott et al., 1996), and also test our series for structural breaks using the most commonly applied approach via Bai and Perron (1998), which tests open-ended break dates, while also testing (via the appendix) a set of historically well-known break dates following Chow (1960). Finally, we conclude this section with an investigation of long-run persistence (half-lives) of global real rates. A range of alternative specifications of the ADF-GLS (including a specification assuming no time trend, and using two alternative inflation expectations constructs from existing literature) and related robustness tests (including results from a Monte-Carlo simulation, and the Chow test) are to be found in the appendix, suggesting a broad confirmation and in important aspects a clear strengthening of our key results.[21] Not least, we present the evidence for short-maturity real rates, using the shorter time period available – we are able to essentially corroborate previous studies using the same approach that we utilize to demonstrate long-maturity trend-stationarity over long horizons.[22]

4.1 An analysis of trend-stationarity properties of real interest rates

We start our analysis by testing the stationarity properties of real interest rates using unit root tests. We find unambiguous evidence in favor of trend stationarity for global real rates when tested via the ADF-GLS test. The ADF-GLS unit root test is implemented in an augmented Dickey-Fuller regression: $\Delta\widetilde{y}_t = \alpha + \beta\widetilde{y}_{t-1} + \sum_{j=1}^{k} \gamma_j \Delta\widetilde{y}_{t-j} + \varepsilon_t$, where $\widetilde{y}_t$ is the GLS-detrended variable.[23] The null hypothesis is that $\beta = 0$. The null hypothesis of a unit root can be rejected at the 5 percent significance level for the headline real series (using GDP-weights in any given year, "Global GW" – this is the default basis on which we report all "global" results unless otherwise noted), for all headline real rate country series, and for an arithmetically-weighted global series ("Global AW"), which gives equal weight to each of the eight constituent countries in any given year, regardless of absolute size of GDP. Specifically, we also note that the evidence in favor of trend stationarity is confirmed for all the individual country series, including for U.S. and U.K. country-level data, with which some researchers might be more familiar: the fact that the results hold up for these variations, and a variety of different inflation and weighting bases, gives us a high degree of confidence in the robustness of the above methodology.

[21] See appendix table A.1 for ADF-GLS without time trend, and section 3.3 (Monte Carlo).

[22] Rose (1988) uses short-maturity U.S. commercial paper rates.

[23] Let the interest rate date be denoted by y_t, $y_t^* = y_t - \alpha^* y_{t-1}$, $t = 2, ..., T$, $y_1^* = y_1$, $\alpha^* = 1 - 13.5/T$; $x_t = 1 - \alpha^*$, $t = 2..., T$, $x_1 = 1$, $d_t = t - \alpha^*(t-1)$, $d_1 = 1$; then $\widetilde{y}_t$ is the residual from a regression of y_t^* on x_t and d_t.

Table 1: ADF-GLS Test with Time Trend			
Real Rate Series, 1318-2022			
Region	Number of lags	ADF-GLS test statistic	Optimal lag
Global GW	3	-6.166	Seq, MAIC
	2	-6.985	SIC
	1	-9.174	
Global AW	3	-6.709	MAIC
	2	-7.234	Seq, SIC
	1	-9.263	
Italy	3	-6.892	Seq, SIC, MAIC
	2	-8.120	
	1	-10.281	
UK	3	-6.742	
	2	-6.875	Seq, SIC, MAIC
	1	-8.561	
Dutch	3	-8.210	Seq, SIC, MAIC
	2	-9.989	
	1	-11.947	
France	3	-6.608	Seq, MAIC
	2	-7.365	SIC
	1	-8.765	
Germany	3	-9.635	
	2	-9.861	Seq, SIC, MAIC
	1	-12.967	
Spain	3	-5.583	
	2	-5.749	Seq, SIC, MAIC
	1	-7.591	
U.S.	3	-3.615	Seq, MAIC
	2	-4.226	
	1	-4.158	SIC
Japan	3	-3.877	
	2	-4.107	MAIC
	1	-4.658	Seq, SIC

Note: The table reports the ADF-GLS test statistic for several choices of the number of lags (with a maximum of 3 lags). The regression includes a constant and a deterministic time trend. The critical values at the 1, 5 and 10 percent significance levels are the following: -3.48 (1%); -2.89 (5%); -2.57 (10%), except for Japan where the levels are -3.52 (1%); -2.98 (5%); -2.69 (10%). "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the Bayesian Information Criterion (SIC), or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value. "Global GW/AW" = GDP-weighted/arithmetically-weighted global real rate basis.

Specifically, the fact that we reject a unit root using both the ADF-GLS test specification without a time trend (table A.1 in the appendix)[24] as well as the specification with a time trend – in addition to extremely similar results when adopting the inflation expectation constructs proposed by Eichengreen (2015) and Hamilton et al. (2016), in the appendix (tables A.2 and A.3) – suggests that these results are not a function of the particular model used: the results, instead, appear to be firmly a function of the extended sample length, the fact that we exclusively focus on long-maturity rates, or a combination of both factors. Additional evidence in favor of trend stationarity is provided by out-of-sample forecast exercises in section 3.3 of the online appendix.

Further, we test a range of alternative asset price series separately for the most recent period in the appendix (section 2.4), including testing private real long-maturity rates, real long sample policy rates, and U.K. real rates. The results broadly confirm our propositions above, and appear to indicate that the ability to reject non-stationarity appears to extend to other fixed income series once long samples are employed. The intuition that our powerful sample length (rather than idiosyncrasies of our data sources, maturities, or geographies) are driving our results is also suggested by Monte Carlo simulations where we generate time series with the same properties (mean, trend, variance and serial correlation) as our global real rate series and, by varying the sample length, we demonstrate that the power of the ADF-GLS test significantly falls when sample sizes shorten. That is, researchers working with samples of progressively smaller sizes are significantly less likely to reject non-stationarity (a more detailed discussion is available in online appendix, section 3.3).

[24] In general, one would expect that under-specifying the trend would result in an inconsistent unit root test: that is, in practice, if there were a deterministic trend and ADF-GLS test did not include a deterministic trend, the test would not reject a unit root (West, 1987). However, in our case, unreported Monte Carlo simulations show that the trend is so small that the misspecification would not affect the power of the test.

Table 2: Zivot-Andrews Test Results	
	Test statistic
Global Real GW	-10.151
Global Real AW	-9.089
Italy Real	-12.624
UK Real	-10.446
Dutch Real	-9.038
France Real	-8.747
Germany Real	-10.397
Spain Real	-8.239
US Real	-6.521
Japan Real	-5.431

Note: The table reports the Zivot and Andrews (1992) unit root test statistic, which allows for a break in both the mean and the trend under the alternative. The critical values at the 1, 5 and 10 percent significance levels are the following for all observations: -5.57 (1%); -5.08 (5%); -4.82 (10%). The trimming parameter is 0.10. The test rejects when the test statistic is negative and larger (in absolute value) than the critical value. When the test rejects the unit root, the column labeled "break date" reports the estimated break date. "Global GW" = GDP-weighted global real rate basis; "Global AW" = arithmetically-weighted global real rate basis.

Next, we report the test by Zivot and Andrews (1992), which tests for non-stationarity while being robust to the presence of a structural break in the mean and the trend. Per table 2, the unit root is rejected once more for both global real rates, as well as for all country-series. While the Zivot-Andrews test is robust to instabilities, it is not a test for structural breaks. Therefore, the next section will discuss more closely the possibility that our series may have undergone structural changes. Finally, we have verified that the stationarity results are robust to using alternative unit root tests, such as variance ratio tests (Cochrane, 1988) and the unit root test by Phillips and Perron (1988). Detailed results are reported in section 3.4 of the online appendix.

4.2 An analysis of instabilities in real interest rates

We present Bai-Perron test results in table 3, which is regularly used in the relevant literature, including in Garcia and Perron (1996) and Hamilton et al. (2016). A key advantage of Bai-Perron is that it does not presuppose specific break dates, which is important given the vast possibility of potential breaks – though for robustness purposes, we also report results for a structural break test that does pre-suppose such break dates, the Chow test following Chow (1960).[25] Bai and

[25] See appendix tables A.4 (unbalanced panel) and A.11.1-2 (balanced panel), and discussions there – we there choose five break dates that are prominently associated with historical inflection points in existing literature, specifically the years 1349, 1557, 1694, 1914, and 1981. The Chow test is implemented in a model that includes a constant and a deterministic trend (as a fraction of the total sample size) using Newey and West (1987)'s heteroskedasticity and autocorrelation consistent (HAC) standard errors with a lag length chosen according to Lazarus et al. (2018). The Chow

Perron's sequential test is applied to the linearly detrended real rate; results are reported at the 5 percent significance level using a trimming parameter of 0.05. Again, as with all other exercises in this main body, we use the progressively-lagged realized inflation measure to construct real rates, excluding the current year *t*, following the methodology in Homer and Sylla (2005).

Per table 3, the Bai-Perron results are notable for the relative scarcity of confirmed breaks in the mean. While a range of 0-4 break dates exists across global and country levels, no less than seven out of the eleven series in fact feature not a single break over a multi-century horizon, including both global levels, the United States, and also the U.K. real rate series via Dimsdale and Thomas (2016). Among the few early modern breaks that are positively detected, several are indeed close to two inflection points that are repeatedly discussed in prior literature (the Black Death in the mid-14th century, and also the series of major financial crises culminating in 1557), and which we discuss further in the context of alternative break tests. Somewhat less prominently, Bai-Perron suggests breaks during the second half of the 15th century (1466, 1477 – points potentially related to the end of the "Bullion Famine", or the general recovery of Britain and France after the Hundred Years' War) and Napoleonic Wars (for the U.K.). But in particular, the rarity of 20th century structural breaks is intriguing – with in fact only a single series (Spain) exhibiting a break over the most recent four decades: even though results from alternative structural tests, including Chow, suggest more caution in positing a definite absence of breaks over the past 200 years, this is a surprising observation in the context of the strong claims in existing literature surrounding especially the 1914 and 1980s inflections.

---

test allows us to test for breaks in mean and trend simultaneously across the five posited historical turning points and, as such, can also be interpreted as a way of capturing the importance of more general types of non-linearities.

Table 3: Bai-Perron Test Results

Series	N. breaks	Estimated break dates
Global Real GW	0	
Global Real AW	0	
Italy Real	1	1356
U.K. Real	4	1477 1552 1814 1922
Dutch Real	0	
France Real	3	1357 1466 1954
Germany Real	0	
Spain Real	2	1916 1982
U.S. Real	0	
Japan Real	0	
Dimsdale U.K. real	0	

Note: the table reports the results of the sequential Bai and Perron's test (Bai and Perron, 1998). The test is implemented in Matlab using the Matlab function 'pbreak' from Pierre Perron's website. The test is applied to the linearly detrended real rate series using a trimming parameter of 5 percent and a HAC variance estimator. The significance level is 5 percent, and the maximum number of allowed break points is 5 for all series. "Dimsdale U.K. real" refers to underlying U.K. data sourced from Dimsdale and Thomas (2016), using HS progressive lags.

Overall, therefore, the evidence across two of the most prominent structural break tests supports a meaningful degree of skepticism about 20th century structural breaks in real interest rates – in any case, the evidence of a "recent" break during the 1980s appears much weaker than existing literature would lead one to expect – with the clear majority of Bai-Perron results rejecting such an idea – even though such a claim cannot be dismissed in its entirety. Where the results from both tests perhaps show the greatest degree of overlap is in suggesting some support for structural breaks during the Black Death, and again in the middle of the 16th century – the latter period coinciding with an exceptional rise in financial volatility, culminating in the combined sovereign default of Spain, France, and the States General in the year 1557.[26]

4.3 Persistence

We measure persistence and half-life next. Because the earlier literature consistently found non-stationarity, the issue of how long shocks take to die out has, not surprisingly, received little attention. To the extent that the existing literature has investigated whether the post-1980s behavior of global real rates is "transitory" – whether rates will eventually return to "a more normal level" (Laubach and Williams, 2016), long-run estimates of such adjustment speeds can provide an

[26]We elaborate on the historical context for this and four other potential inflection points in appendix section 3.5 – with these dates forming the basis of our "pre-supposed" break date tests in Chow. The most detailed account of the tumultuous events during the 16th century remains Ehrenberg (1896).

indication of how quickly global real rates typically return to their structural trend after shocks.[27] As per table 4, which delineates persistence during key historical eras, we observe that the half-life of real interest rate deviations for the entire sample period ranges between 1 and 9 years (most of them 1-3 years) across individual sub-samples.[28] Half-lives stood below three years during the fourteenth and fifteenth centuries, a figure that moderately rises over the remainder of the bullion regime era, after which it sharply increases over the late gold standard and early gold-exchange standard period, only to moderate again towards the end of the 20th century, to levels slightly (but not sharply) above long-run averages seen during the early modern period. This general trend is equally visible if alternative weighting methods are employed ("Global AW"), and if country-level trends are investigated separately. The "safe asset providers" over time – Italy pre-1600, the U.K. after 1557, the U.S. in the 20th century – appear to record somewhat faster adjustment speeds. Though formal break tests for persistence may add nuance to these observations, what we can firmly say is that there at least appears no obvious long-run trend towards faster adjustment speeds in the modern era, albeit there are many factors including the changing composition of, and some degree of measurement error for, the CPI over time (with the share of volatile commodities prices falling) that may be at work. Nevertheless, the fact that adjustment speeds are fairly stable across time and across countries, suggests that the convergence speed is quite plausibly of the order of magnitude that we suggest. It is important to note that our relatively short half lives (around 1-3 years) do not contradict the fact that we need a relatively large sample in order to have sufficient power to reject non-stationarity, as we show in our Monte Carlo simulations in the appendix.[29]

How do our results compare to persistence results for other early modern variables? In terms of levels, our data suggests that early modern persistence in global real rates appears to be broadly comparable to respective persistence in real exchange rates – though after 1914 they appear to be somewhat higher than identified persistence in real exchange rates, for which half-lives of 3-5 years have traditionally been posited then (Rogoff, 1996; Lothian and Taylor, 1996; Itskhoki, 2021).[30] A decomposition of aggregate persistence measures at the country level does not reveal notable cross-country differences: aggregation biases similar to those posited for real exchange rate persistence measurements are therefore not apparent.

[27]We consider both half-lives based on the bootstrap by Kilian (1999) as well as robust measures of half-lives (via lag augmented local projections), robust to complicated dynamics such as local-to-unity roots and high persistence (Olea and Plagborg-Møller, 2021) but computationally less demanding than the grid-bootstrap in Hansen (1999) and Steinsson (2008). In the latter case, the half-lives are obtained as follows: Let the detrended interest rate ($y(t)$) be an AR(p). We regress $y(t+h)$ on $y(t), \ldots, y(t-p)$: the coefficient on $y(t)$ is the impulse response at horizon h; its confidence intervals are obtained via Eicker-Huber-White standard errors. We plot the impulse response and select the first horizon at which the response reaches half of the initial effect. The lags are selected by the Bayes information criterion.

[28]We report all additional country-level results in the online appendix (section 3.6).

[29]In fact, our Monte Carlo simulations, tailored to our estimated global real rate series, show that a researcher would reject non-stationarity only 88 percent of the times in a sample of 75 observations, and only 63 percent of the times in a sample of 50 observations, while he/she would reject it with probability almost equal to one in our sample of roughly 700 observations. See the online appendix section 3.3 for details.

[30]The uncertainty around our estimated half-lives is much lower than the substantial uncertainty surrounding estimates of the half-lives of real exchange rates, due to the high persistence of the latter.

In the context of such findings (some of which are more tangentially relevant in the measures they investigate, but are thus indicative of the limits of existing evidence), while our *levels* of persistence generally fall in line with the higher end of the range of estimates for early modern half lives – we are unable to confirm any evidence of secularly faster adjustment speeds when focusing on long-maturity real rates. If anything, real rate persistence on our basis appears to exhibit a moderate upwards trend over time. A clear increase in half-lives appears to coincide with the surge in war inflation in 1914, followed by a moderation after 1981 (to a present level above the early 20th century). This finding is relevant in the context of existing literature that associated the eve of the First World War – coinciding closely with the founding of the Federal Reserve and the departure from fixed exchange rate regimes – with a decline in financial volatility and seasonality in interest rates (Barsky et al. (1988); Bernstein et al. (2010)): while the management of seasonality, and general predictability in financial markets may both have improved, this did not necessarily imply faster adjustment speeds. Overall, the figures indeed suggest that, at least as of 2022, post-GFC global real rates may have "under-corrected" relative to historical adjustments (given that they still remain visibly below structural trend levels and persistence confidence bands well over a decade after 2008) – however, as of yet this multi-year deviation does not register as a structural break: and it cannot necessarily be expected to do so even in the event of a further downward entrenchment of rates, given the rare occurrence of structural breaks we identified. An ultimate "normalization" to the structural downward trending level would therefore be the default expectation on the basis of our data.

	Period	Half-life	Confidence interval		Robust Half-life	Confidence interval	
Global Real GW	1318-1348	1.45	0.99	2.05	1.28	0.48	2.23
	1349-1556	1.43	1.17	1.71	1.35	0.89	1.77
	1557-1693	2.46	1.80	3.35	2.02	1.15	3.83
	1694-1913	2.06	1.71	2.51	1.93	1.11	2.99
	1914-1980	6.87	4.29	14.72	6.63	1.93	11.64
	1981-2022	1.87	1.02	3.62	1.47	0.70	7.98
Global Real AW	1318-1348	1.49	1.03	2.09	1.37	0.63	2.14
	1349-1556	1.42	1.14	1.75	1.37	0.56	2.21
	1557-1693	2.30	1.73	2.97	1.96	1.17	3.55
	1694-1913	1.91	1.61	2.59	1.82	1.01	3.41
	1914-1980	8.56	4.82	28.39	6.98	1.66	12.05
	1981-2022	3.75	1.84	14.63	2.51	1.03	8.12
U.K. Real	1318-1348	1.44	0.97	2.11	1.39	0.62	2.33
	1349-1556	1.34	1.06	1.67	1.11	0.24	5.57
	1557-1693	0.89	0.73	1.27	0.98	0.60	1.75
	1694-1913	1.59	1.38	1.84	1.55	1.06	1.99
	1914-1980	5.39	3.82	7.91	5.03	1.66	9.28
	1981-2022	4.61	2.40	14.81	4.11	1.44	N.A.
Germany Real	1318-1348	2.22	1.50	3.09	1.99	1.05	2.76
	1349-1556	2.32	1.60	3.23	1.87	1.11	3.86
	1557-1693	2.80	2.26	4.42	2.73	0.64	6.16
	1694-1913	1.79	1.49	2.15	1.72	1.09	2.50
	1914-1980	6.56	2.26	N.A.	3.66	0.76	6.15
	1981-2022	4.86	1.47	N.A.	2.23	0.94	5.50
U.S. Real	1797-1913	2.66	2.06	3.45	2.67	0.95	3.74
	1914-1980	5.29	3.44	8.95	6.16	1.61	12.08
	1981-2022	0.96	0.62	1.40	0.94	0.49	1.54

Table 4: Half-Lives of Real Rates, by Sub-samples

Note: The half-life is estimated as the first horizon at which the impulse response equals one-half of the initial impact effect. The impulse response is estimated using a deterministically detrended linear autoregressive model (with the number of lags chosen by the Bayesian Information Criterion) using Kilian (1999)'s bootstrap; the point estimate is based on the median unbiased response while the confidence interval is based on the 5-th and 95-th quantiles of the bootstrap distribution. The robust half-lives and their confidence intervals are estimated based on lag-augmented local projections on a model with a constant and a deterministic time trend. The half-lives are separately estimated in each sub-sample identified by the statistically significant break dates (according to the Chow test). "Global GW" = GDP-weighted global real rate basis; "Global AW" = arithmetically-weighted global real rate basis. NA denotes situations where the half-life is infinity.

To summarize: this section suggests that testing long horizon samples of real interest rates strongly confirms their stationary properties around a downward deterministic trend – in contrast to existing literature that focused on short-maturity rates, using short samples. Our results are robust to a range of specifications, including alternative inflation expectation constructs and the existence of potential structural breaks, all of which are described in detail in the appendix. Our results on the persistence are novel, too, being the first to assess a key financial series for advanced economies over a multi-century period all the way to the present: here, in the context of existing literature, both the relatively fast adjustment speeds in the early modern period, as well as the absence of striking evidence in favor of secularly faster adjustment speeds are worth noting. Taken together, these results rationalize why the existing consensus on real interest rate properties has arisen: but they equally demonstrate that previous results do not necessarily hold once a higher statistical power is achieved by using longer samples.

5 Real interest rates in the context of growth and demographic variables

The relationship between real interest rates and other macroeconomic variables has been of key interest in the literature for many years, but very long-run empirical investigations were hampered for the same reasons that prevented the study of very long-run interest rate trends: a lack of satisfactory data. Fortunately, over the most recent years in particular, significant advances have been made in the improvement of long-run output and population growth series – two variables of particular interest in the real interest rate literature. These advances allow us to relate our new evidence on real interest rates more granularly to output growth trends, and population growth over centuries. Table 5 reports simple period averages of our three key macro variables over time. All figures are based on an identically-weighted eight country sample. The early phase of the demographic sample still contains a non-negligible share of (decadal- or semi-centennial-level) interpolations – but this does not influence subsequent statements or calculated averages. It is evident from the table already that real interest rates and both population growth on the one hand, and real output growth rate on the other, move in divergent directions secularly. Afterwards, we discuss both variables in detail and undertake a few more formal exercises involving the use of a filter (Müller and Watson, 2018) that removes short- and medium-run fluctuations (shorter than 100 years) from all variables.

Beginning with long-run output growth trends, recent years have seen fundamental updates to the seminal work of Maddison (2010). While "big assumptions" still have to be made in the underlying construction, new series have significantly improved granularity. New multi-century output reconstructions have been undertaken for Northern Italy (Malanima, 2011), France (Ridolfi and Nuvolari, 2021), Spain (de la Escosura et al., 2022), Germany (Pfister, 2022), Holland (van Zanden and van Leeuwen, 2012), and the U.K. (Broadberry et al., 2015). These typically rely on (wage-based) demand-side constructions, or more granular reconstructions of sectoral outputs.[31]

[31] In "demand-side approaches", the GDP series are constructed on the basis of (combinations of urban and rural) real wage data – assuming that they reflect aggregate income trends when combined with estimates of urbanization rates,

The new developments are fortunate as they now allow a nuanced comparison both on the country-level and on the aggregate level of our real interest rate trends, and associated trends in output. To arrive at aggregate real output growth rates, we also created a "GW-weighted" real GDP growth series, by GDP-weighting the country constituent output growth rates based on the aggregate total GDP share for the benchmark years: Figure 2 displays the resulting series over the period 1318-2021, together with our respective Global GW real rate series. We report the long-run trend in these variables, extracted using Müller and Watson's filtering approach tailored to retain only fluctuations at frequencies longer than 100 years. We observe modest output dynamics over the first half of the series, with aggregate growth driven mainly by (moderate) levels of population growth prior to the 18th century, though with some more vigorous dynamics on the regional level: individual growth spurts do take place in the early modern period, with Dutch real per capita incomes rising by 82% between 1500-1600, for instance, and a more general 29% real per capita output increase for our entire sample over 1500-1700. A sharp acceleration is detectable across all country levels during the 18th and 19th centuries, giving rise to the familiar Industrial Revolution dynamics. During the second half of the 20th century, a deceleration is notable, with current (post-2008) rates of growth meaningfully below trend lines: yet, absolute rates of growth in the early 21st century remain highly elevated compared to very long-run averages.

Table 5: Macro Variables: Averages by Period (in percent per annum)

	1318-2021	1318-1500	1500-1800	1800-1914	1914-2021
Real aggregate output growth rate	1.04	0.14	0.48	2.08	2.99
Aggregate population growth rate	0.37	-0.17	0.29	1.03	0.81
Real interest rate	6.16	11.17	5.87	3.67	1.14

Note: The table reports the average value for three variables real aggregate output growth rate; aggregate population growth rate; real interest rate. All three variables are based on the GW-weighted sample of our eight advanced economies over time. For full underlying output and population sources, see text and notes to figures 2 and 3.

Crucially – and as is clearly visible from the chart itself – the output growth series overall displays a monotonic *upwards* trend.[32] The positive correlation between output growth and real rates appears to be a rather unique phenomenon of the very most recent part of the data series, starting in the second half of the 20th century. Generally speaking, therefore, real interest rates and aggregate output growth rates exhibit trend dynamics in directly *opposite* directions over the secular term.[33]

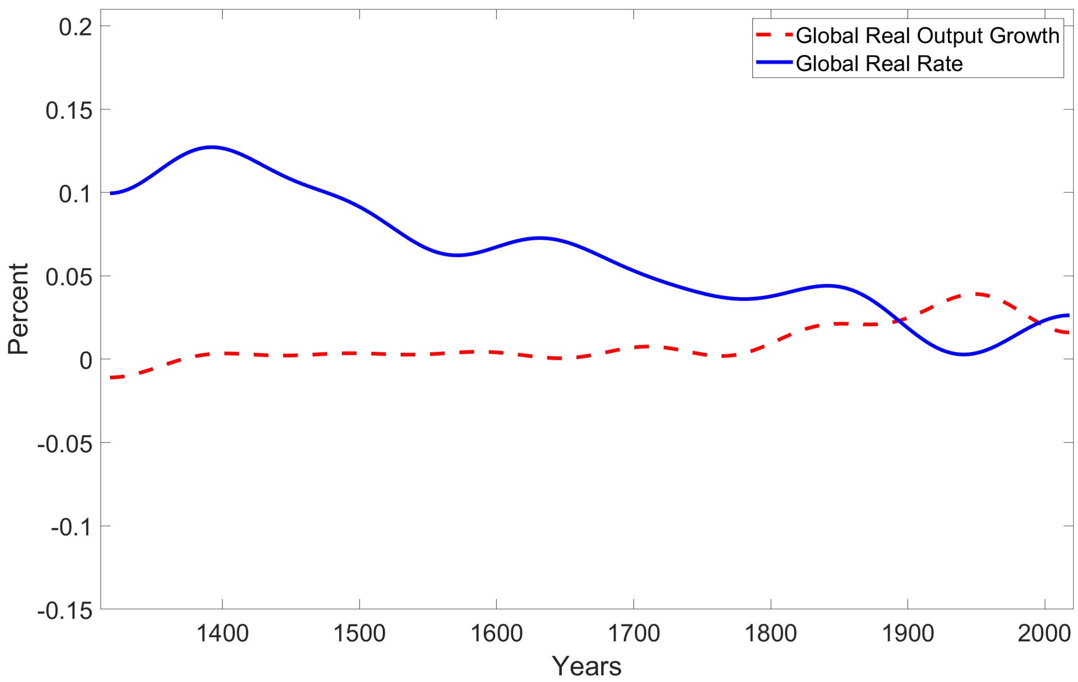
and of trends in days worked: the series for France (Ridolfi and Nuvolari, 2021), Germany (Pfister, 2022), and Northern Italy Malanima (2011) fall into this group. English and Dutch data (van Zanden and van Leeuwen, 2012; Broadberry et al., 2015), meanwhile, rely on sectoral output reconstructions. For a recent alternative approach using wealth data to construct Italian GDP benchmarks, see van Zanden and Felice (2022).

[32]This upwards trend is readily confirmed visually, but can also be obtained via standard tests such as Mann-Kendall (Mann, 1945; Kendall, 1955).

[33]Of course, historically, countries that have accumulated very high debt will often experience extended periods of slower growth and higher real interest rates, as suggested by Reinhart et al. (2012), and now the central conclusion of a large literature as surveyed in Abbas et al. (2019). This would not, however, explain the long-term opposing trends.

Taken together, these observations imply that output growth trends and real interest rate trends do not have a correlation in the direction commonly assumed. In fact, over extended periods of time over the long-run, both measures in fact go in opposite directions.

Figure 2: A comparison of Müller-Watson-filtered real aggregate output growth rates and real interest rates, 1318-2021.



Notes: The aggregate output growth rate also covers the eight country sample, with country constituent series also weighted according to GDP-weights (GW). Sources include Malanima (2011) for Northern Italy; Pfister (2022) for Germany, starting in 1501; de la Escosura et al. (2022) for Spain; Ridolfi and Nuvolari (2021) for France; van Zanden and van Leeuwen (2012) for Holland, starting in 1349; Broadberry et al. (2015) for the U.K. For modern data covering the most recent period, we rely on IMF International Financial Statistics. The dashed line displays the (filtered) long-run trend of the real aggregate output growth rate, also showing the GW weights. The solid line displays the (filtered) long-run trend of the global real rate. The long-run trends are obtained via the Müller and Watson (2018) filtering approach, retaining fluctuations with periodicity over 100 years.

Demographic factors have been among the most frequently-invoked factors in the assessment of real interest rate dynamics in recent decades. Lunsford and West (2019) for instance consider demographic variables as the most plausible explanatory variables compared to growth and productivity variables – in particular, the authors find statistically significant roles for the share of 40-64-year olds in the population, as well as the dependency ratio, over the period 1890-2016, relying on U.S. data.[34] Gagnon et al. (2021) find that "demographic factors account for much, if not all of the actual permanent decline in the equilibrium real rate [since the 1980s]" – highlighting in particular falling fertility rates and mortality in the 1960s and 1970s as causal variables. But what does the (very) long-run data suggest for demographic context? A first-order reconstruction

[34] In Lunsford and West (2019), the dependency ratio is positively correlated with real rates, while the share of 40-64-year olds is negatively correlated; the authors find no statistically significant role for life expectancy.

of population growth rates since the Renaissance can equally utilize Maddison (2010)'s data. But advances in output growth empirics since then have been mirrored by advances in demographic data for advanced economies, often undertaken by the same authors. New population estimates – from which we derive aggregate population growth rates over the long-run – are equally provided by Malanima (2011), van Zanden and van Leeuwen (2012), Broadberry et al. (2015), Ridolfi (2019, building on Dupaquier (1988)), Pfister (2022), and de la Escosura et al. (2022).

Figure 3 now displays long-run fluctuations in the aggregate population growth rate; as before, we extract fluctuations of 100 years and more using the Müller and Watson (2018) filtering approach. The series covers the sample 1318-2021 also applying GW weights to the population growth data that we use for the global real rate series. As is visible from the chart, and can be confirmed by the statistical exercise below, aggregate population growth rates and real interest rates have structurally trended in *opposite* directions, except for the most recent decades. During major population growth shocks, the two series are positively correlated, such as during the Black Death of 1348-1351, the most dramatic demographic "rare disaster" striking advanced economies since the Renaissance – a period coinciding with a sharp plunge in global real rates: but these episodes resulting in positive correlations remain a clear exception over the long-run.[35]

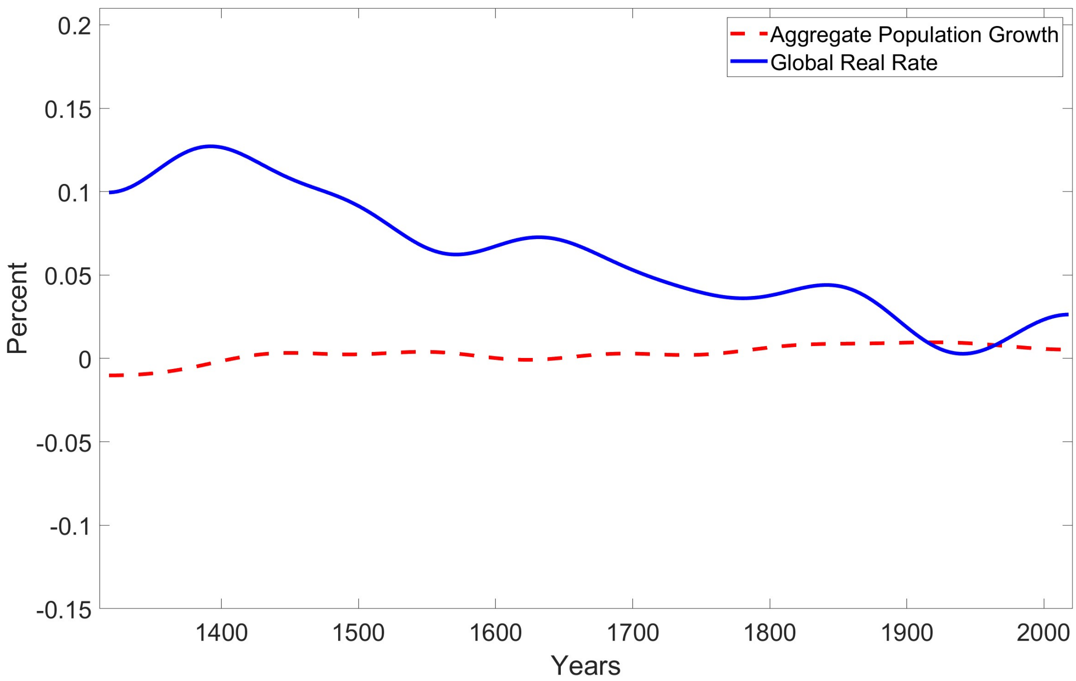
The finding that a secular upwards trend in population growth coincides with a downwards trend in real rates is not necessarily inconsistent with some recent literature – even though most contributions assume a positive correlation between the two variables (Rachel and Smith, 2016). But Carvalho et al. (2016) stress that population growth impacts (short-maturity) real rates through two divergent channels: a "supply effect" implying a rising (falling) population growth rate lowers (raises) capital per-worker, thus putting upwards (downwards) pressure on real rates; on the opposite, a (lagged) "demand effect" in which a rise (fall) in population growth should eventually lead to a lower (higher) dependency ratio, which generates downwards (upwards) pressure on real rates. Both channels are of course posited in the context of modern welfare states and the existence of substantial retirement periods: it is not obvious ex ante whether and to what extent prior to the introduction of modern welfare states and prolonged life expectancies, the balance between the two channels changes significantly.[36] However, note that even Carvalho et al. (2016) posit that the forces generating a positive correlation between population growth and real rates dominate: therefore, similar to the evidence from output dynamics, the long-run evidence from our new data appears to run counter to recent causal demographic propositions of real rate dynamics.

Table 6 now reports the estimated correlation between the long-run filtered component of global real rates versus those of aggregate output and population growth. The table confirms a strong

[35]Of course, we are studying here global sovereign borrowing rates – this does not necessarily imply that the rate of return on, say, land moved in the same way.

[36]Empirically, economy-wide life expectancy at birth for advanced economies accelerates sharply from the mid-19th century, but shows no decisive trend prior to this (thus also displaying an apparent disconnect with real interest rate dynamics). For a comprehensive recent empirical contribution to elite lifespan dynamics see Cummins (2017): though his sample over-weights English and Welsh data, his finding that European "elite" longevity increased sharply around two particular points, around 1400 and again around 1650, seems somewhat more consistent with our real rate evidence than data from either life expectancy at birth or population.

Figure 3: A comparison of Müller-Watson-filtered aggregate population growth rates and real interest rates, 1318-2021.



Notes: The aggregate population growth rate covers identical eight country sample, with country constituent series weighted according to GDP-weights of the sum of country outputs. Sources include Malanima (2011) for Northern Italy; Pfister (2022) for Germany, starting in 1501; de la Escosura et al. (2022) for Spain ("compromise estimate"); Dupaquier (1988) and Ridolfi (2019) for France; van Zanden and van Leeuwen (2012) for Holland; Broadberry et al. (2015) for the U.K. For modern data covering the most recent period, we rely on IMF International Financial Statistics. The dashed line displays the (filtered) long-run trend of the aggregate population growth rate. The solid line displays the (filtered) long-run trend of the global real rate. The long-run trends are obtained via the Müller and Watson (2018) filtering approach, retaining fluctuations with periodicity over 100 years.

negative and statistically significant correlation.

Table 6: Global Real Rates Vs. Aggregate Real Output and Population Growth

	Correlation	67% CI		90% CI	
Müller-Watson-filtered Long-run Component of Aggregate Population Growth	-0.447	(-0.800	-0.133)	(-0.850	0.028)
Müller-Watson-filtered Long-run Component of Aggregate Real Output Growth	-0.448	(-0.800	-0.106)	(-0.850	0.099)

Note: The table reports the correlation of each series in the first column with the long-run component of the global real rate (all series with GW weights), based on Müller and Watson (2018). The long-run is defined as lasting more than 100 periods. The table reports the median of the posterior distribution as well as the 67% and 90% confidence intervals.

In sum, the long-run evidence puts doubts on both productivity-centered and demographics-centered explanations of the post-1980s trend fall in real interest rates, as represented for instance by Gordon (2016), Holston et al. (2017), or Goodhart and Pradhan (2021). Several contributions have already voiced skepticism about productivity channels, including Hamilton et al. (2016) and Lunsford and West (2019) – we can confirm these from a much longer-term perspective. But importantly, we can qualitatively go beyond these existing facts, and find strong evidence in favor a long-run *negative* prediction of aggregate population growth rates and aggregate real output growth rates for real rates. The very most recent decades – where for both variables the correlation with real rates is positive – in fact appear to be a unique exception in a multi-century perspective.

Since demographics and productivity therefore cannot explain the 700-year trend decline in long-term real interest rates, is there a compelling alternative explanation? This important question is, unfortunately, beyond the scope of this paper: we can only offer a couple of suggestions for future research. Stefanski and Trew (2022), using the evidence in Schmelzing (2020) as motivation, offer a theoretical explanation based on trend rising social patience using a dynamic heterogeneous agent model of fertility, in which more patient economic and social groups have more progeny. Changes in war are another possibility. Post-war Germany and Japan both (partially) defaulted outright on their pre-Second World War debt, Germany inflated away its First World War debt. Over time, related evidence from debasement frequencies (e.g. via Karaman et al. (2020)) suggests that sovereign default risk may have secularly trended downwards, even if outright inflation risks are higher post-1914. Ongoing reconstructions of long-run inequality trends in advanced economies may allow for a future refinement of theories focusing on the respective channels – recent evidence points towards plausible links between rising inequality over ca. 1450-1914 in the relevant economies, though the association appears more inconsistent before and after this period (Alfani, 2022).[37] Relatedly, although fluctuations are massive, it is almost surely the case that trend liquidity in major government bond markets has been positive over the centuries and continues to be.[38] Whatever the answer(s), we can only say that it needs to be consistent with the evidence here on the fact that long-term real interest rates series appear to be trend stationary: and regardless of the underlying cause, the analysis does suggest that when real interest rates rise far above or fall far below the (declining) long-term trend, there is a high chance of mean reversion unless of

[37] Alternatively, the gradual change in the *relative* financial risks between the private and public sectors may have induced a faster decline of sovereign risk premia, a hypothesis that would be consistent with secularly rising measures of bank stress in our eight economies since at least the 17th century (Metrick and Schmelzing, 2021).

[38] One would of course expect substantial increases in market liquidity over our sample length. But to what extent this development is driven by financial broadening rather than deepening is not established. Interestingly, in a number of polities at at the relative forefront of innovation and market depth, market turnover in the early modern period is far from negligible: as expounded above, Mueller (1997) reports a turnover figure for Venetian *Monte Vecchio* sovereign debt in the year 1434 equivalent to over 28% of Venetian nominal GDP for the year. Hammel (1988), who assembled one of the most nuanced housing market data sets for multiple centuries (1284-1700) in a leading Imperial city found turnover rates of 3.4% - 6.8% of the housing stock in a given year, overwhelmingly debt-financed (the U.S. turnover equivalent in 2003-2018 stood at 5.4% for single family homes). At least for the U.K., evidence also suggests that liquidity factors may have been a more relevant factor for equity markets, rather than for bond markets (Campbell et al., 2018). As of now, we lack high-frequency bid-ask spread data which would be one variable able to provide particularly nuanced insights.

course there has been a rare structural break.

6 Conclusion

The sharp drop in the real interest rate in the 21st century, particularly in the years after the global financial crisis, has been arguably the most important macroeconomic development in modern times. The main goal of our paper is to investigate to what extent the fall should be regarded as permanent (or even intensifying), and to what extent can it be regarded as temporary, with expected eventual reversion to a longer-term trend. This paper has offered a novel approach to this problem by taking advantage of recent dramatic advances in quantitative financial history to get a better handle on the econometric properties of real interest rates than has previously been possible. Importantly, we focus on long-maturity sovereign debt (comparable to modern 10-year Treasury bonds) over eight centuries, from 1311-2022. This is in contrast to earlier research that has mainly focused on short rates, often splicing government and private bonds, and uniformly with much shorter time periods. Importantly, prior to this paper, virtually all of the literature has been unable to reject the hypothesis of a unit root in real interest rates. The results here, based on much longer data sets, and on eight countries, are markedly different – and besides strongly rejecting the unit root hypothesis, firmly establish the existence of a structural downward trend in global real interest rates over centuries.

Overall, our findings regarding structural breaks is that they occur less frequently over the past centuries that one would expect, and not necessarily at the points in time where existing literature would lead one to expect them: these are surprising results given the severity and frequency of wars and changes in monetary regimes, such as 1914. In particular, although 1981 does show up in multiple variations of individual series (but exclusively when using the set-up of the Chow test), there is far from consistent evidence that a break in global real rates have in fact taken place over "recent" decades. In fact, using the most common structural break test in the literature, we find that several advanced economies may not have seen structural breaks since the Renaissance – not least, this appears to concern modern "safe" long-dated assets, including U.S. Treasury bonds (from 1787) and U.K. sovereign debt. This, in turn, suggests that if short-term real interest rates are indeed non-stationary – and we are not convinced they are – it is a premium for short-term debt – potentially the particular safety and liquidity premia inherent in short term debt (Krishnamurthy and Vissing-Jorgensen, 2012) – rather than the real interest rate itself, that is non-stationary: at least this is an issue to be explored. It is unlikely that the reason our results are so different from the literature has to do with some anomalous feature of long-term sovereign bonds; in fact, the appendix shows that we can equally reject the unit root hypothesis for long-maturity mortgage and corporate real rates, even though these series are much shorter than our headline sovereign rate data.

We are aware of the debate in the econometrics literature on the difficulties in distinguishing trend breaks from unit roots. However, the glaring downward trend in our 700-year real interest rate series is so consistent and clear it is hard to imagine not allowing for a trend, and in fact, and even for the series where we detect a statistically significant trend break, the stationarity results still hold without allowing for one. Of course, further research is needed, though we are doubtful alternative approaches would radically change our novel finding that long-term real interest rates are stationary. We are also keenly aware that extrapolating our negative trend line another two centuries even at the pace of around two basis points per annum, would imply that the global trend real interest rate proceeds linearly, and ultimately reaches levels well below zero. A more likely possibility, perhaps, is that it will asymptote to a lower bound (positive or negative), though we do not see evidence of that yet and leave this issue to further study.

Finally, the recent literature has given a great deal of attention to demographics and productivity as drivers of today's real interest rates. Our research in no way rules this out, since it is possible there has been a modern structural break that has not yet lasted long enough to be statistically significant in the 700-year time series. Nevertheless, again making use of substantial recent advances in quantitative economic history, we can employ much-improved long-term demographic and real output data, which over most of our sample have been rising, not falling, yet the sovereign long-term real interest rate has trended down consistently. Specific alternative explanations for the very slow trend decline include decreasing default risk and increasing liquidity, both of which are likely to be important contributing variables and merit exploration in future research, though we do not presume an answer here. Hence, although we absolutely cannot conclude that the post-2008 epoch of seemingly very low real rates will eventually unwind back to a (declining) trend, we would argue that our data gives a valuable new perspective that should have important implications for understanding a broad range of modern macroeconomics and asset pricing issues.

References

Abbas, Ali S., Alex Pienkowski, and Kenneth S. Rogoff. 2019. *Sovereign Debt: A Guide for Economists and Practitioners*. Oxford University Press.

Acres, William Marston. 1931. *The Bank of England from Within 1694-1900. Volume 1.* Oxford University Press.

Alfani, Guido. 2022. Epidemics, Inequality, and Poverty in Preindustrial and Early Industrial Times. *Journal of Economic Literature* 60(1): 3–40.

Allen, Robert. 2001. The Great Divergence in European Wages and Prices from the Middle Ages to the First World War. *Explorations in Economic History* 38(4).

Alvarez-Nogal, Carlos. 2014. Monetary and Financial Innovation in the Spanish Empire: Lights and Shadows. In *Explaining Monetary and Financial Innovation*, edited by Bernholz, Peter, and Roland Vaubel, 301–321. Springer.

Alvarez-Nogal, Carlos, and Prados de la Escosura. 2013. The Rise and Fall of Spain (1270-1850). *Economic History Review* 66(1): 1–37.

Ang, Andrew, and Geert Bekaert. 2002. Regime Switches in Interest Rates. *Journal of Business and Economic Statistics* 20(2): 163–182.

Bai, Jushan, and Pierre Perron. 1998. Estimating and Testing Linear Models with Multiple Structural Breaks. *Econometrica* 66(1): 47–78.

Barbadoro, Bernardino. 1929. *Le Finanze della Repubblica Fiorentina. Imposta diretta e debito pubblico fino all'istituzione del Monte*. Leo S. Olschki.

Barsky, Robert, N. Gregory Mankiw, Jeffrey Miron, and David Weill. 1988. The Worldwide Change in the Behavior of Interest Rates and Prices in 1914. *European Economic Review* 32: 1123–1154.

Bauer, Michael D., and Glenn D. Rudebusch. 2020. Interest Rates under Falling Stars. *American Economic Review* 110(5): 1316–1354.

Bernanke, Ben. 2004. The Great Moderation. *Speech at the Eastern Economics Association* .

Bernanke, Ben. 2005. The Global Savings Glut and the US current account deficit. *Sandridge Lecture, Virginia Association of Economists* .

Bernstein, Asaf, Eric Hughson, and Marc Weidenmier. 2010. Identifying the effects of a lender of last resort on financial markets: lessons from the founding of the Fed. *Journal of Financial Economics* 98(1): 40–53.

Blanchard, Olivier. 2023. *Fiscal Policy under Low Interest Rates*. MIT Press.

BoE. 1963. The Treasury Bill. In *Bank of England Quarterly Bulletin, Q3-1963*, 186–193.

Bolt, Jutta, and Jan Luiten van Zanden. 2020. Maddison style estimates of the evolution of the world economy. A new 2020 update .

Broadberry, Stephen, Bruce Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen. 2015. *British Economic Growth 1270-1870*. Cambridge University Press.

Campbell, Gareth, John D. Turner, and Qing Ye. 2018. The liquidity of the London capital markets, 1825-1870. *Journal of Economic History* 71(3): 823–853.

Campbell, John Y. 2018. *Financial Decisions and Markets. A Course in Asset Pricing.* Princeton University Press.

Carvalho, Carlos, Andrea Ferrero, and Fernanda Nechio. 2016. Demographics and real interest rates: Inspecting the mechanism. *European Economic Review* 88: 208–226.

Chilosi, David, Max-Stephan Schulze, and Oliver Volckart. 2018. Benefits of Empire? Capital Market Integration North and South of the Alps, 1350–1800. *Journal of Economic History* 78(3): 637–672.

Chow, Gregory C. 1960. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. *Econometrica* 28(3): 591–605.

Clark, Gregory. 2016. Microbes and Markets: Was the Black Death an Economic Revolution? *Journal of Demographic Economics* 82: 139–165.

Cochrane, John. 1988. How Big Is the Random Walk in GNP? *Journal of Political Economy* 96: 893–920.

Cummins, Neil. 2017. Lifespans of the European Elite, 800-1800. *Journal of Economic History* 77(2): 406–439.

de la Escosura, Prados, Carlos Alvarez-Nogal, and Carlos Santiago-Caballero. 2022. Growth recurring in preindustrial Spain? *Cliometrica* 16(2): 215–241.

de Luca, Giuseppe. 2015. Finanza Pubblica, State Building, ed economia reale nella lombardia Spagnola. In *Antico Regime e finanza pubblica: gli Stati italiani preunitari*, edited by Cini, Marco, 33–66. Edizioni ETS.

de Moor, Tine, and Jaco Zuijderduijn. 2013. Spending, saving, or investing? Risk management in sixteenth-century Dutch households. *Economic History Review* 66(1): 38–56.

DelNegro, Marco, Domenico Giannone, Marc Giannoni, and Andrea Tambalotti. 2019. Global trends in interest rates. *Journal of International Economics* 118: 248–262.

Delumeau, Jean. 1959. *Vie économique et sociale de Rome dans la seconde moitié du XVIe siècle. Tome second*. Bibliothèque des Ecoles françaises d'Athènes et de Rome.

Denzel, Markus. 2010. *Handbook of World Exchange Rates, 1590–1914*. Routledge.

Dimsdale, Nicholas, and Ryland Thomas. 2016. A Millennium of Macroeconomic Data, Version 3.1. *Bank of England Data Repository* .

Dupaquier, Jacques. 1988. *Histoire de la population française, 4 vols.* Presses Universitaires de France.

Ehrenberg, Richard. 1896. *Das Zeitalter der Fugger: Geldkapital und Creditverkehr im 16. Jahrhundert (2 Vols.)*. Gustav Fischer, Jena.

Ehrenberg, Richard. 1928. *Capital and Finance in the Age of the Renaissance. A Study of the Fuggers and their Connections.* Harcourt, Brace, and Row.

Eichengreen, Barry. 2015. Secular stagnation, the long view. *American Economic Review: Papers and Proceedings* 105(5): 66–70.

Eichengreen, Barry, Asmaa El-Ganainy, Rui Esteves, and Kris Mitchener. 2021. *In Defense of Public Debt*. Oxford University Press.

Elliott, Graham, Thomas J. Rothenberg, and James H. Stock. 1996. Efficient Tests for an Autoregressive Unit Root . *Econometrica* 64(4): 813–836.

Epstein, Stephen. 2000. *Freedom and Growth: The Rise of states and markets in Europe, 1300-1750*. Routledge.

Fama, Eugene. 1975. Short-Term Interest Rates as Predictors of Inflation. *American Economic Review* 65(3): 269–282.

Federico, Giovanni, Max-Stephan Schulze, and Oliver Volckart. 2021. European Goods Market Integration in the Very Long Run: From the Black Death to the First World War. *Journal of Economic History* 81(1): 276–308.

Flandreau, Marc, and Frederic Zumer. 2000. *The Making of Global Finance, 1880-1913*. OECD Publishing.

Frankel, Jeffrey A. 1986. International Capital Mobility and Crowding-Out in the U.S. Economy: Imperfect Integration of Financial Markets or of Goods Markets. In *How open is the U.S. economy?*, edited by Hafer, Rik W., 33–67.

Friedman, Milton, and Anna J. Schwartz. 1963. *A Monetary History of the United States, 1867-1960*. Princeton University Press.

Fritschy, Wantje. 2017. *Public Finance of the Dutch Republic in Comparative Perspective: The Viability of an Early Modern Federal State (1570s-1795)*. Brill.

Gagnon, Etienne, David Johannsen, and David Lopez-Salido. 2021. Understanding the New Normal: The Role of Demographics. *IMF Economic Review* 69(2): 357–390.

Garbade, Kenneth D. 2008. Why the U.S. Treasury Began Auctioning Treasury Bills in 1929. *Federal Reserve Bank of New York Review* 14(1): 1–17.

Garcia, Rene, and Pierre Perron. 1996. An Analysis of the Real Interest Rate Under Regime Shifts. *Review of Economics and Statistics* 78(1): 111–125.

Gelabert, Juan. 1999. The Italian States in the Early Modern Period. In *The Rise of the Fiscal State in Europe, c.1200-1815*, edited by Bonney, Richard, 417–442. Oxford University Press.

Gelderblom, Oscar, and Joost Jonker. 2011. Public Finance and Economic Growth: The Case of Holland in the Seventeenth Century. *Journal of Economic History* 71(1): 1–39.

Gelderblom, Oscar, and Joost Jonker. 2020. Province of Holland: market prices 1596-1794 communicated by Oscar Gelderblom and Joost Jonker .

Goodhart, Charles, and Manoj Pradhan. 2021. *The Great Demographic Reversal*. Springer.

Gordon, Robert. 2016. *The Rise and Fall of American Growth*. Princeton University Press.

Hamilton, James D., Jan Hatzius, Ethan Harris, and Kenneth D. West. 2016. The Equilibrium Real Rate: Past, Present, and Future. *IMF Economic Review* 64(4): 66–707.

Hammel, Rolf. 1988. Häusermarkt und wirtschaftliche Wechsellagen in Lübeck von 1284 bis 1700. *Hansische Geschichtsblätter* 106: 41–107.

Hansen, Bruce E. 1999. The Grid Bootstrap And The Autoregressive Model. *The Review of Economics and Statistics* 81(4): 594–607.

Hauser, Henri. 1930. The European Financial Crisis of 1559. *Journal of Economic and Business History* (2): 241–255.

Hildebrandt, Reinhard. 1996. *Quellen und Regesten zu den Augsburger Handelshaeusern Paler und Rehlinger 1539-1642. 2 Vols.* Franz Steiner.

Hollander, Jacob H. 1919. *War Borrowing. A Study of Treasury Certificates of Indebtedness of the United States*. Macmillan Company.

Holston, Kathryn, Thomas Laubach, and John Williams. 2017. Measuring the Natural Rate of Interest: international trends and determinants. *Journal of International Economics* 108(S1): 59–75.

Homer, Sidney, and Richard Sylla. 2005. *A History of Interest Rates*. University of Rutgers Press.

Huang, Angela, David Chilosi, and Alexandra Sapoznik. 2019. A source collection on urban annuities, 14th-18th centuries: an introduction to the data. *Vierteljahresschrift fur Sozial und Wirtschaftsgeschichte* 106(1): 67–80.

Huber, Alfons. 1893. Studien über die finanziellen Verhältnisse Oesterreichs unter Ferdinand I. *Mitteilungen des Instituts fuer Oesterreichische Geschichtsforschung* 181–247.

Huizinga, John, and Frederic Mishkin. 1986. Monetary Policy Regime Shifts and the Unusual Behavior of Real Interest Rates. *Carnegie-Rochester Conference on Public Policy, April 1985* .

Itskhoki, Oleg. 2021. The Story of the Real Exchange Rate. *Annual Review of Economics* 13: 423–455.

Jorda, Oscar, Moritz Schularick, and Alan M. Taylor. 2017. Macrofinancial History and the New Business Cycle Facts. In *NBER Macroeconomics Annual 2016*, edited by Eichenbaum, Martin, and Jonathan A. Parker. University of Chicago Press.

Karaman, K. Kinvanc, Sevket Pamuk, and Secil Yildirim-Karaman. 2020. Money and monetary stability in Europe, 1300–1914. *Journal of Monetary Economics* 115(11): 279–300.

Kendall, M.G. 1955. Rank Correlation Methods .

Kilian, Lutz. 1999. Exchange rates and monetary fundamentals: what do we learn from long-horizon regressions? *Journal of Applied Econometrics* 14(5): 491–510.

Kindleberger, Charles P. 1998. Economic and Financial Crises and Transformations in Sixteenth-Century Europe. *Princeton Essays in International Economics* (208).

Kreimes, Wilhelm, Heinrich Otto, Fritz Vigener, and Ernst Vogt. 1913. *Regesten der Erzbischöfe von Mainz 1289–1396. 5 Vols. (1913-1958).* Veit.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2012. The Aggregate Demand for Treasury Debt. *Journal of Political Economy* 120(2): 233–267.

Kuechler, Winfried. 1983. *Die Finanzen der Krone Aragon waehrend des 15. Jahrhunderts (Alfons V. und Johann II.)*. Aschendorffsche Verlagsbuchhandlung.

Lapeyre, Henri. 1955. *Une famille de marchands - les Ruiz.* Librairie Armand Colin.

Laubach, Thomas, and John Williams. 2003. Measuring the Natural Rate of Interest. *Review of Economics and Statistics* 85(4): 1063–1070.

Laubach, Thomas, and John C. Williams. 2016. Measuring the Natural Rate of Interest Redux. *Business Economics* 51(2): 57–67.

Lazarus, Eben, Daniel J. Lewis, James H. Stock, and Mark W. Watson. 2018. HAR Inference: Recommendations for Practice. *Journal of Business and Economic Statistics* 36(4): 541–574.

Legay, Marie-Laure. 2011. *La banquerote de l'Etat royal: la gestion des finances publiques de Colbert a la Revolution*. Editions EHESS.

Longnon, Auguste. 1878. *Paris pendant la domination anglaise (1420-1436). Documents: extraits des registres de la chancellerie de France*. H. Champion.

Lothian, James R., and Mark P. Taylor. 1996. Real Exchange Rate Behavior: The Recent Float from the Perspective of the Past Two Centuries. *Journal of Political Economy* 105(3): 488–509.

Lunsford, Kurt G., and Kenneth West. 2019. Some Evidence on Secular Drivers of the Real Rate. *American Economic Review: Macroeconomics* 11(4): 113–139.

Maddison, Angus. 2010. *The World Economy - A Millennial Perspective*. OECD.

Malanima, Paolo. 2011. The long decline of a leading economy: GDP in central and northern Italy, 1300-1913. *European Review of Economic History* 15(2): 169–219.

Mankiw, N. Gregory, and Jeffrey Miron. 1986. The Changing Behavior of the Term Structure of Interest Rates. *Quarterly Journal of Economics* 101(2): 211–228.

Mann, H.B. 1945. Nonparametric tests against trend. *Econometrica* 13: 245–259.

Metrick, Andrew, and Paul Schmelzing. 2021. Banking-Crisis Interventions, 1257-2019. *NBER Working Paper* .

Mian, Atif, Amir Sufi, and Ludwig Straub. 2021. What Explains the Decline in r*? Rising Income Inequality Versus Demographic Shifts. *Chicago BFI Working Paper* 1–48.

Mueller, Reinhold. 1997. *The Venetian money market: banks, panics, and the public debt, 1200-1500*. Johns Hopkins University Press.

Mulho, Anthony. 1971. *Florentine public finances in the Early Renaissance, 1400-1433*. Harvard University Press.

Müller, Ulrich, and Mark Watson. 2018. Long-run Covariability. *Econometrica* 86(3): 775–804.

Munro, John H. 2003. The Monetary Origins of the 'Price Revolution' : South German Silver Mining, Merchant-Banking, and Venetian Commerce, 1470-1540' 1–60.

Neely, Christopher J., and David Rapach. 2008. Real Interest Rate Persistence: Evidence and Implications. *Federal Reserve Bank of St. Louis Review* 90(6): 609–642.

Nelson, Charles R., and Charles I. Plosser. 1982. Trends and Random Walks in Macroeconomic Time Series. *Journal of Monetary Economics* 10: 139–162.

Newey, Whitney K., and Kenneth D. West. 1987. A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55(3): 703–708.

Ng, Serena, and Pierre Perron. 2001. LAG Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Journal of Econometrics* 69(6): 1519–1554.

North, Douglass C., and Barry R. Weingast. 1989. Constitutions and commitment: the evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History* 49(4): 803–832.

Nuglisch, Adolf. 1899. *Das Finanzwesen des Deutschen Reiches unter Kaiser Karl IV.* Schlesier and Schweikhardt.

Olea, José Luis Montiel, and Mikkel Plagborg-Møller. 2021. Local Projection Inference is Simpler and More Robust Than You Think. *Econometrica* 89(4): 1789–1823.

Pamuk, Sevket. 2007. The Black Death and the origins of the 'Great Divergence' across Europe, 1300–1600. *European Review of Economic History* 11(3): 289–317.

Perron, Pierre. 1989. The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis. *Econometrica* 57(6): 1361–1401.

Pezzolo, Luciano. 2003. The Venetian Government Debt, 1350-1650. In *Urban public debts, urban government and the market for annuities in Western Europe (14th-18th centuries)*, edited by Boone, Marc, Karel Davids, and Paul Janssens, 61–74.

Pezzolo, Luciano. 2008. Government Debts and Credit Markets in Renaissance Italy. In *Government Debts and Financial Markets in Europe*, edited by Caselli, Fausto Piola, 17–32. Pickering and Chatto.

Pfister, Ulrich. 2022. Economic growth in Germany, 1500–1850. *Journal of Economic History* 62(4): 1071–1107.

Phillips, P.C.B., and P. Perron. 1988. Testing for a Unit Root in Time Series Regression. *Biometrika* 75(2): 335–346.

Rachel, Lukasz, and Thomas D. Smith. 2016. Secular drivers of the global real interest rate. *Bank of England Staff Working Paper* (571).

Rachel, Lukasz, and Lawrence S. Summers. 2019. On Falling Neutral Real Rates, Fiscal Policy, and the Risk of Secular Stagnation. *Brookings Papers on Economic Activity* (1): 1–54.

Rapach, David E., and Christian E. Weber. 2004. Are real interest rates really nonstationary? New evidence from tests with good size and power. *Journal of Macroeconomics* 26(3): 409–430.

Reinhart, Carmen, and Kenneth Rogoff. 2009. *This Time is Different: Eight Centuries of Financial Folly*. Princeton University Press.

Reinhart, Carmen M., Vincent Reinhart, and Kenneth S. Rogoff. 2012. Public Debt Overhangs: Advanced-Economy Episodes since 1800. *Journal of Economic Perspectives* 26(3): 69–86.

Richards, R.D. 1936. The Exchequer Bill in the History of English Government Finance. *Economic Journal* 3(11): 193–211.

Ridolfi, Leonardo. 2019. Six Centuries of Real Wages in France from Louis IX to Napoleon III: 1250–1860. *Journal of Economic History* 79(3): 589–627.

Ridolfi, Leonardo, and Alessandro Nuvolari. 2021. L'histoire immobile? A reappraisal of French economic growth using the demand-side approach, 1280–1850. *Explorations in Economic History* 25(3): 402–428.

Rogoff, Kenneth S. 1996. The Purchasing Power Parity Puzzle. *Journal of Economic Literature* 34(2): 647–668.

Rose, Andrew K. 1988. Is the Real Interest Rate Stable? *Journal of Finance* 43(5): 1095–1112.

Schirmer, Uwe. 2006. *Kursachsische Staatsfinanzen (1456-1656) : Strukturen, Verfassung, Funktionseliten*. Franz Steiner.

Schmelzing, Paul. 2020. Eight Centuries of Global Real Rates, R-G, and the 'suprasecular' decline, 1311-2018. *Bank of England Staff Working Paper* (845).

Schmelzing, Paul. 2023. Eight centuries of global real rates, R-G, and the 'suprasecular' decline, 1311-2021. *mimeo – update of Schmelzing (2020), with underlying data publicly available* .

Schmelzing, Paul. 2025. *Eight Centuries of Global Real Rates, 1311-2024*. Yale University Press (under contract).

Schulz, Julian. 2017. Regesta Imperii online. *RIDE. A journal for digital editions and resources* 6: 1–29.

Shiller, Robert J. 2015. *Irrational Exuberance. Revised and expanded Third edition*. Princeton University Press.

Shiller, Robert J., and Jeremy J. Siegel. 1977. The Gibson Paradox and Historical Movements in Real Interest Rates. *Journal of Political Economy* 85(5): 891–908.

Siegel, Jeremy. 1992. The Real Rate of Interest from 1800 to 1990. *Journal of Monetary Economics* 29: 227–252.

Stasavage, David. 2011. *States of Credit: Size, Power, and the Development of European Polities*. Princeton University Pres.

Stefanski, Radek, and Alex Trew. 2022. Selection, Patience, and the Interest Rate. *University of Glasgow Working Paper* .

Steinsson, Jón. 2008. The Dynamic Behavior of the Real Exchange Rate in Sticky Price Models. *American Economic Review* 98: 519–33.

Sussman, Nathan. 2022. Financial Developments in London in the Seventeenth Century: The Financial Revolution Revisited. *Journal of Economic History* 82(2): 480–515.

Sussman, Nathan, and Yishay Yafeh. 2006. Institutional reforms, financial development and sovereign debt: Britain 1690–1790. *Journal of Economic History* 66(4): 906–935.

Tracy, James D. 1985. *A Financial Revolution in the Habsburg Netherlands. Renten and Rentiers in the County of Holland, 1515-1565*. University of California Press.

van den Sprenkel, J.W. Berkelback. 1937. *Regesten van Oorkonden betreffende De Bisschoppen van Utrecht uit de Jaren 1301-1340*. Broekhoff.

van Zanden, Jan Luiten, and Emanuele Felice. 2022. Benchmarking the Middle Ages: fifteenth century Tuscany in European perspective. *European Review of Economic History* .

van Zanden, Jan Luiten, and Bas van Leeuwen. 2012. Persistent but not Consistent: the growth of national income in Holland, 1347-1807. *Explorations in Economic History* 49(2): 119–130.

Vuehrer, A. 1886. *Histoire de la Dette Publique en France*. Berger-Levrault et Cie.

West, Kenneth D. 1987. A note on the power of Least Squares Tests for a Unit Root. *Journal of Business and Economic Statistics* 24: 249–252.

Wicksell, Knut. 1936. *Interest and Prices (trans. of 1898 ed.)*.

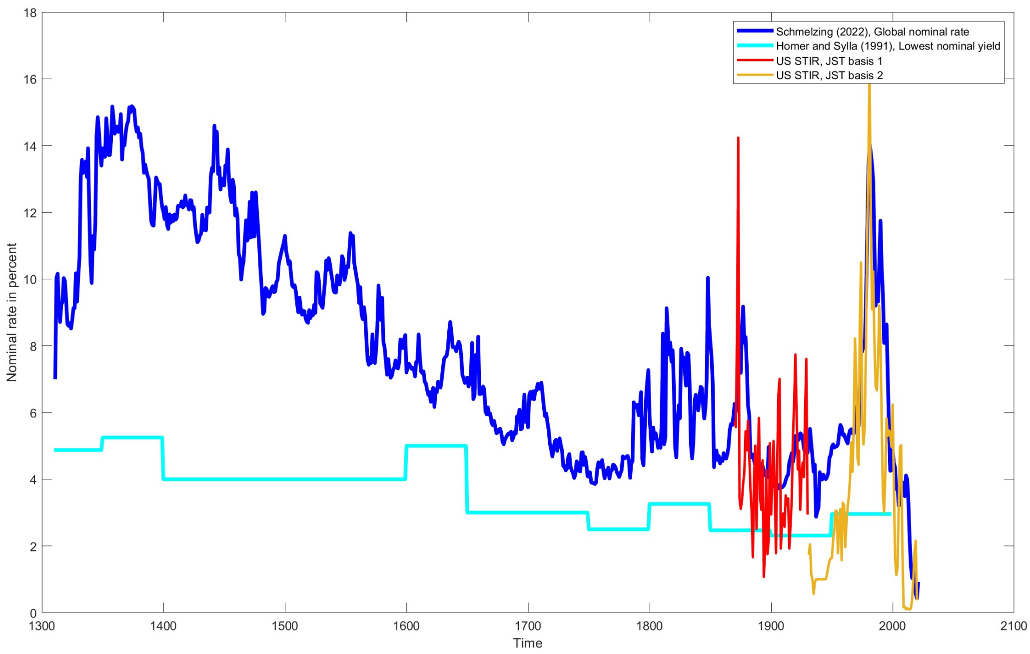
Zivot, Eric, and Donald W.K. Andrews. 1992. Great Crash, the Oil-Price Shock, and the Unit-Root Hypothesis. *Journal of Business Economic Statistics* 10(3): 251–270.

APPENDIX

1. Data construction - real rates

Here we elaborate on the construction of the "global real rate" time series in more detail: for an extensive discussion, readers should refer to the discussions in Schmelzing (2023) and Schmelzing (2025). Figure A.1 displays the comparative coverage of the new data set, relative to the reference data in Homer and Sylla (2005) and "JST" (Jorda et al., 2017), the same underlying data that is also the basis for DelNegro et al. (2019), among others.

Figure A.1: Headline global nominal series in comparison: Schmelzing (2023), Homer and Sylla (2005), and "JST", 1311-2022.



Notes: Charts displays the new global nominal interest rate series constructed in Schmelzing (2023), in comparison to the benchmark nominal Homer and Sylla (1991) series, as well as the short-maturity "JST" series (Jorda et al., 2017), as used by DelNegro et al. (2019), among others.

The country-level data focuses on a collection of long-maturity nominal sovereign interest rate data points. A total of 16,479 long-maturity nominal data points are obtained – the vast majority of which concern the pre-1850 era. "Long-maturity" is defined as sovereign debt transactions with either a de jure (contracted) or de facto contract length of at least two years (e.g. perpetual instruments in practice redeemed well after two years, such as life or heritable annuities). A detailed data overview of both contracted and de facto maturities of hundreds of contracts included in the series yields an average maturity of the assets used over the period 1311-1830 of 13.2 years. The lowest average maturity for a 50-year sub-period is 6.1 years (1510-1560), the highest average maturity for any sub-period is 24 years (1409-1459). While some variation is present within this range, therefore, there is no apparent strong trend in the average maturity over time.

An innovative feature of the reconstructions in Schmelzing (2023) is that it builds separate time series of "unconsolidated" and "consolidated" long-maturity debt transactions for the eight economies over centuries. "Unconsolidated" transactions are all transactions that were not formally structured on the basis of standardized, subscription-based issuance that involved a multitude of creditors. In the early modern period, sovereigns for a variety of purposes chose to engage in "unconsolidated" debt transactions: these transactions – to the extent they are included in the series – were fully voluntary, and secured, but de facto created an individualized "over-the-counter" contract between a creditor (often a wealthy merchant, a member of the nobility, a religious order) and the sovereign. Such transactions were quick to conclude when funds were needed swiftly, and they allowed tailored contract terms – but often they were conducted in parallel to standardized, consolidated issuance of debt. 935 such loans have been identified over the period of 1311-1866 across the eight economies. The majority of these data points are based on new primary and printed primary sources: among primary sources, the extensive Habsburg collections in the Austrian State Archives (OeStA) have been used, as well as the Prussian (GStaPK) and Dutch (StAAM) counterparts; examples in the printed primary space are the German Imperial Registries ("Regesta Imperii"), a source collection covering more than 145,000 medieval and early modern documents from the court and administration of the Holy Roman Emperors;[39] In the British realm, the equivalent *Calendar of State Papers* spanning the reign of the Tudors to the Stuarts (1509-1714) were analyzed. Dozens of complementary registry collections contain data points on transactions involving ducal, regional, or merchant creditor counterparties: here, the duchy (and archdiocese) of Mainz collection via Kreimes et al. (1913) is representative of an important German region; for Paris in the 15th century, Longnon (1878) has assembled a typical document collection; a Dutch example – covering the region of Utrecht during the 14th century – is the collection of van den Sprenkel (1937). And on the merchant side, comprehensive contract collections have been utilized, for instance, for the influential houses of Paler and Rehlinger (Nuremberg-based merchant families and European-wide sovereign lenders), published via Hildebrandt (1996).

All these collections also contain relevant new data on *consolidated* debt transactions, for which new primary material was thus also sourced: for instance, 19th century data points are based on daily newspaper and stock exchange reports (so-called *Börsenkursblätter*, held from 1832 in the Bethmann-Bank archive in Frankfurt); from 1866 for Spain, Italy, Germany, Holland, and the U.S. in the *Frankfurter Zeitung*). But for these latter assets, recent years have also seen advances in the secondary literature, and such recent contributions, including de Luca (2015, for new rates of the Duchy of Milan), Huang et al. (2019, for data specifically covering urban annuities), Gelderblom and Jonker (2020, for new Dutch market yields over 1596-1794), and Sussman (2022, for new 17th century British rates) have equally been integrated. In addition, separate samples of "short-maturity", intra-governmental, and "involuntary" debt transactions have been identified in the data collection process, but these are not part of any samples in this paper and are analyzed fully in Schmelzing (2025).[40]

Methodologically, separate country series are then first constructed on the basis of these unconsolidated interest rates, with equal weights given to individual regions (see details below). Next, equivalent country series are constructed for consolidated interest rates – rates, hence, which refer exclusively to funded, standardized debt issuance transactions, as long as they are equally in the long-maturity spectrum. Finally, the respective unconsolidated and consolidated country series are fused for each country, by averaging the

[39] The current stage of this long-running project is set to conclude in the year 2033 under the auspices of the "Deutsche Kommission für die Bearbeitung der Regesta Imperii e. V.", for an introduction see Schulz (2017).

[40] In this category, we find also yields on "office sales", loans from quasi central banks (such as the Venetian *Banco Giro* or the *Bank of Amsterdam*), loans from higher officials employed in the sovereign courts, "withholding taxes" (disguised forced loans), and war indemnities (also often disguised as loans).

two observations for each year (one unconsolidated country average, one consolidated country average): this fusion is undertaken irrespective of exact issuance volumes for unconsolidated vs consolidated debt until the year 1800, when issuance of unconsolidated debt has largely ceased. Prior to 1800, Schmelzing (2020, see Sheet VIII., "Pers. sov. loans, 1312-") and Schmelzing (2023) report the issuance volume for a meaningful share of unconsolidated debt – when compared with consolidated debt stock data, it is reasonable to conclude that equal-weights do not critically distort aggregations for individual years.[41] These final steps yield the eight benchmark country series which then respectively incorporate the full range of unconsolidated as well as consolidated debt operations passing the specified filters. Finally, we note that the early modern data basis overwhelmingly refers to coupon rates (primary yields) for sovereign debt, as determined at the issuance date – this is in line with conventions, and in the case of the unconsolidated debt sample represents the only basis to record such rates (since these individualized contracts were not traded with high frequency, though they were typically transferable). Existing studies have found a high degree of coherence between market rates and coupon data (Chilosi et al., 2018): in the case of consolidated debt issuance, new bonds were of course priced not necessarily at par, but below par in the case of weak prevailing market conditions – such information is generally well documented and always taken into account in the data.[42]

Interpolations. As mentioned in the main body of the paper, while dramatically reducing the need for interpolations compared to earlier long-term data, the Schmelzing (2023) data still features linear interpolations. In particular, interpolations remain non-negligible in nominal interest rate data for the United Kingdom prior to 1650, for France between 1375-1450, for the Netherlands between 1330-1470, and for Spain between 1670-1785. Meanwhile, in select long-run secondary sources that we could have used for the country-level inflation series, interpolations are also present – specifically for "Northern Italy" over the years 1620-1700 in Allen (2001): in all of the cases on the inflation side, we were able to choose otherwise-equivalent appropriate time series that did not feature interpolations – in the case of Northern Italy, we therefore relied on the silver-price CPI construction of Malanima (2011). As further examined in the interpolation exercise in appendix section 3.1, all of our main results, including on trend stationarity at the 1% significance level in country-level (Italy, U.K., France) and global level (GW and AW) continue to hold when fully excluding those sub-periods affected by interpolations.

Geographical bases. The data set follows "best practice" approaches on geographical definitions familiar from standards in long-run economic history literature, including national accounting practices established over recent years. While six out of our eight economies undergo relatively limited geographical changes, two of them display significant centralization tendencies over time, as well as political variation: the Holy Roman Empire/Germany, and (Northern) Italy. We follow reference resources such as the "Maddison Database" (Bolt and van Zanden, 2020), which treat data points from "Northern Italy" – primarily composing the Republics of Venice, Siena, Florence (later Duchy of Tuscany), and Genoa – as proxies for "Italy" prior to the Napoleonic Wars of the early 1800s, and followed by the intermittent polities (from 1815 also including

[41]On both the central government and non-central government levels, our information on total debt levels is reasonably good, even though such information has to be pieced together from a large variety of individual studies, and does not exist in summarized form: e.g. Mueller (1997) reports debt stocks for Venice in the 14th and 15th centuries; Tracy (1985) and Fritschy (2017) for Holland from the 16th century; Vuehrer (1886) remains the resource for the French central government, and so on.

[42]Wherever secondary yields in the early modern period are available, these have been incorporated, for instance Florentine data points in Barbadoro (1929), Venetians in Mueller (1997) etc. Chilosi et al. (2018) show that differences in primary and secondary yields for early modern geographies with both data points available averaged 28 basis points, for cases where absolute nominal long-maturity interest rates averaged 5 per cent: in other words, the prevailing early modern gap between primary and secondary yields averaged around five per cent of the par value.

the "Kingdom of the Two Sicilies", a region treated as a Spanish component prior to 1800) that eventually united into the modern state of Italy in 1860. Similarly, for "Germany", we also follow both older and most recent work, for instance Bolt and van Zanden (2020) or Pfister (2022), and incorporate data points from the individual German states which were bound together institutionally by a common administrative level under the Holy Roman Empire, but undertook their own autonomous financial and economic policies, with no centralized debt issuance. In particular, the "Habsburg" lands and rulers, which over ca. 1400-1800 assumed a de facto hereditary claim to the Imperial Crown but run their own territorial region within the Empire, are treated as one among other relevant dukes and princes engaged in sovereign debt issuance in their territories. While not one of the polities undergoing "significant" geographical changes, we follow standard geographical practices in long-run economics for the U.K., too, when considering data points prior to 1715 from England, Scotland, and Wales. When aggregating data points on this sub-national level, in all cases we follow an arithmetic weighting, irrespective of the population, financial asset stock, or output sizes of the regional component: for example, the "Paris" and "Gascogne" regions are equal-weighted in cases where each reports a long-maturity unconsolidated sovereign debt transaction in a respective year, to constitute the "France" data point: if an additional transaction by the monarch himself exists for the same year, for instance, each data point would be weighted with a 1/3 share.

Currency denominations. Overwhelmingly, the debt transactions involved are denominated in gold currency: the most frequently used currencies found in early sovereign debt contracts include the Venetian ducat, the Florentine florin, the German Rheinflorin, the French écu, and the Dutch guilder. With the weights and finesse of these international currencies increasingly being standardized in the early modern period, it was relatively difficult for the sovereign to engage in quasi-default on gold-denominated liabilities by way of a debasement operation: in practice, moreover, a significant share of debt contracts under consideration contained explicit clauses about the exact *weight* to be used in the coins for the repayment.[43] Hence, while debasement represented an easy - and frequently-used - recourse to diminish silver-denominated *domestic liabilities*, such as wage payments to mercenaries and other court employees – we should not assume the systematic distortion of interest rate data points via premia related to debasement risk.[44] Similarly, most long-run inflation series are denominated in silver-price equivalents: the alternative is to base series on local-currency price changes, which tend to be more volatile in the short-run. Key reference works such as Allen (2001) report both bases: this allows a comparison of long-run trends, which have been conducted: while temporary divergences are notable over short time-spans, no meaningful long-run differences are observed for resulting real rate series, which is intuitive given that arbitrage dynamics and sufficient commodity market integration existed: whatever gaps between local currency and silver currency prices opened up tended to converge relatively swiftly, certainly over multi-year horizons (Federico et al., 2021).

The **Global GW basis**. When reference is made to "global real rates" in Schmelzing (2023), Schmelzing (2025), or in this paper, we refer – unless explicitly stated otherwise – to the series obtained by weighting the eight individual country series according to their respective rolling GDP shares over time. These GDP shares are obtained using the consistent definitions of population and per capita real GDP figures in Maddison (2010): the sum of the eight country-level aggregate GDP figures (population x per capita GDP, for each country, linearly interpolating between Maddison's benchmark years) represents the "global GDP"

[43] Schmelzing (2023) details typical clauses in these contracts. Representative is the stipulation regularly found in German contracts that the loan is to be repaid in "Gut Gulden und Gewicht" ("Proper Gulden coins and weight"), for instance found in the contracts by the Margrave of Brandenburg in the 1460s, with the original contracts in the *Geheimes Staatsarchiv Preussischer Kulturbesits* (GStA PK), I. HA Rep. 78, Nr. 10, fols. 96-131.

[44] For recent summaries and advances in this lengthy literature, see the introduction in Denzel (2010, LVIII-CX), and Karaman et al. (2020).

figure, with the U.S. and Japan entering the sample in 1790 and 1870, respectively. Alternative weighting approaches are of course possible, and while very minor, the recent advances in country-level accounting (as discussed above) have also slightly shifted the *relative* GDP shares over Maddison's figures – hence, our addition of the "Global AW" weighting approach in all exercises, which simply weights all eight countries with equal weights regardless of aggregate output. The fact that our results hold across these different approaches (as well as for the country series themselves, as detailed in the respective result tables) suggests that the particular weighting methodology is ultimately not decisive.

2. Robustness

Here, we present robustness checks and a variety of additional results following from our benchmark results in the main body of the text. We begin by testing the data in using the ADF-GLS test in a specification *without* a time-trend, via table A.1.

Tables A.2 and A.3 in section 2.1 then proceed with variations on the inflation expectations construct used to calculate real interest rates: recall that our exercises in the main text follow the methodology in Homer and Sylla (2005), and use a progressively lagged seven-year realized inflation basis, not including the current year, in order to infer inflation expectations. Table A.2 alternatively follows the approach in Hamilton et al. (2016), and estimates inflation expectations via an autoregressive model: per table A.2 we can confirm all our ADF-GLS results in this approach, too. Second, we follow Eichengreen (2015), who uses an equal-weighted seven-year realized inflation basis, including the current year, to adjust U.S. nominal yields in order to obtain real rates. For all except one series (Global GW), we can equally confirm the ADF-GLS results from the main body. Our stationarity results also hold for the approach used in Schmelzing (2023), where current time t inflation is used to proxy for inflation expectations.

In section 2.2, we then present the results from the alternative structural break test, following Chow (1960), and undertake an analogous variation of the inflation expectations approach, applying an equal-lagged inflation expectation approach (again in the spirit of Eichengreen (2015)) to five break dates, which are selected based on strong claims in existing literature: all our main results in this alternative setting are consistent with our existing discussion (table A.4). Afterwards, in section 2.4, we test our basic ADF-GLS approach for a variety of alternative long-sample fixed income series, to assess whether there are indications that peculiarities inherent in our long-sample sovereign debt series may be driving results – importantly the independently constructed long-run real rate data for the U.K. in Dimsdale and Thomas (2016): overall, as reported via tables A.6.1 – and A.6.3, we find that stationarity can be confirmed for a number of alternative long-run data sets which mainly focus on British variables starting in the 18th and 19th centuries.

In the subsequent online appendix sections (3.1 - 3.6), we first show that our results are robust to interpolation (Tables A.7.1 and A.7.2), and go on to show that we can fully replicate the failure to reject unit roots in real rates when we apply two alternative inflation expectations approaches for Rose (1988)'s shorter sample data (tables A.8.1-A.8.3). We confirm this failure to reject a unit root uniformly at the 1% level for all 16 possible real rate variations, including for his long-maturity corporate bond rate basis, as well as for his short-maturity commercial paper rate basis. This underscores that our headline results do not appear to be a function of how we combine the individual nominal rate and inflation series. Part 3.3 reports our Monte Carlo simulation and out-of-sample forecasting results. Part 3.4 discusses results based on Phillips and Perron's and Cochrane's (1988) variance ratio tests. Section 3.5 elaborates on the rationale that informed our choice of the five pre-supposed structural break years – which we applied to delineate the persistence eras, and also use in the Chow test; finally section 3.6 reports the remaining country-level half-life results.

Table A.1 now reports stationarity test results based on the ADF-GLS test without including a time trend. The results confirm our main results in the body of the paper. Again, as noted in the text, the Monte Carlo results confirm that given the volatility of the real rate series, this is quite plausible.

Table A.1: ADF-GLS Test without Time Trend

Real Rate Series			
Region	Number of lags	ADF-GLS test	Optimal lag
Global GW	3	-4.758	Seq, SIC, MAIC
	2	-5.475	
	1	-7.339	
Global AW	3	-3.400	Seq, MAIC
	2	-3.788	SIC
	1	-5.080	
Italy	3	-5.817	Seq, SIC, MAIC
	2	-6.913	
	1	-8.851	
UK	3	-2.638	MAIC
	2	-2.777	Seq, SIC
	1	-3.635	
Dutch	3	-7.363	Seq, SIC, MAIC
	2	-9.064	
	1	-10.983	
France	3	-4.952	Seq, SIC, MAIC
	2	-5.606	
	1	-6.795	
Germany	3	-5.469	MAIC
	2	-5.847	Seq, SIC
	1	-8.095	
Spain	3	-3.002	MAIC
	2	-3.152	Seq, SIC
	1	-4.320	
U.S.	3	-2.852	Seq, MAIC
	2	-3.367	
	1	-3.361	SIC
Japan	3	-3.025	
	2	-3.268	MAIC
	1	-3.784	Seq, SIC

Note: the table reports the test statistic for several number of lags (for a maximum of three lags). The regression includes a constant. For all series except for U.S. and Japan, the critical values at 1, 5, 10 percent significance levels are the following for all observations: -2.58 (1%); -1.95 (5%); -1.62 (10%). For the U.S., the critical values are -2.582; -1.95; -1.619. For Japan, they are -2.593; -1.95; -1.613. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

.1 Variations of inflation expectations

Hamilton et al. (2016) use an autoregressive inflation expectations approach. We replicate their approach for our multi-century data in table A.2 below. We calculate one-period-ahead inflation expectations following the same procedure as in Hamilton et al. (2016). That is, the evolution of inflation is recursively estimated in rolling windows using an autoregressive model with one lag; then, the estimated model is used to predict one-year-ahead inflation in an out-of-sample fashion. The rolling window estimation procedure guards against instabilities by using only past observations in the latest window of data. We use a window size equal to 30 years. As shown, we are able to reject a unit root with this inflation basis, too, for both the global (GW and AW), and all country series.

Second, Eichengreen (2015) uses a seven-year equal-lagged realized inflation expectation construct to adjust nominal yields, and includes current-year inflation (year t). We replicate his approach in table A.3 below. As shown, we are equally able to reject a unit root with this inflation basis, for all series, including Global AW – with the only exception of the Global GW series.

Table A.2: ADF-GLS Test for the Hamilton et al. (2016) inflation basis			
Real Rate Series			
Region	Number of lags	ADF-GLS test	Optimal lag
Global GW	3	-9.249	Seq, MAIC
	2	-10.680	SIC
	1	-13.267	
Global AW	3	-8.515	MAIC
	2	-9.417	Seq, SIC
	1	-11.733	
Italy	3	-7.848	MAIC
	2	-8.704	Seq, SIC
	1	-11.865	
UK	3	-11.676	MAIC
	2	-13.636	
	1	-16.982	Seq (0), SIC
Dutch	3	-3.665	Seq, MAIC
	2	-3.914	SIC
	1	-4.498	
France	3	-8.013	
	2	-8.305	Seq, SIC, MAIC
	1	-10.038	
Germany	3	-10.310	MAIC
	2	-11.617	Seq, SIC
	1	-17.402	
Spain	3	-5.341	Seq, MAIC
	2	-5.862	SIC
	1	-8.241	
U.S.	3	-4.180	SIC, MAIC, Seq
	2	-5.934	
	1	-7.015	
Japan	3	-3.649	
	2	-3.697	Seq, MAIC
	1	-4.861	SIC

Note: the table reports the test statistic for several number of lags (for a maximum of three lags). The regression includes a constant and a time trend. For all series except for Japan, the critical values at 1, 5, 10 percent significance levels are the following for all observations: -3.48 (1%); -2.89 (5%); -2.57 (10%). For Japan, the critical values are -3.552; -3.007; -2.717. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

Table A.3: ADF-GLS Test for the Eichengreen (2015) inflation basis

Real Rate Series			
Series	Number of lags	ADF-GLS test	Optimal lag
Global GW	3	-6.711	
	2	-6.717	Seq, SIC, MAIC
	1	-7.871	
Global AW	3	-6.549	
	2	-6.367	Seq, SIC, MAIC
	1	-7.378	
Italy	3	-8.020	MAIC
	2	-8.809	
	1	-9.931	Seq, SIC
UK	3	-5.155	Seq, MAIC
	2	-5.646	
	1	-5.994	SIC
Dutch	3	-6.640	Seq, MAIC
	2	-7.522	
	1	-8.688	SIC
France	3	-5.202	
	2	-5.344	Seq, MAIC
	1	-5.830	SIC
Germany	3	-8.113	
	2	-8.254	Seq, SIC, MAIC
	1	-9.727	
Spain	3	-4.400	
	2	-4.522	Seq, SIC, MAIC
	1	-5.883	
U.S.	3	-3.973	
	2	-4.240	Seq, SIC
	1	-3.622	MAIC
Japan	3	-4.176	Seq
	2	-3.714	MAIC
	1	-4.038	SIC

Note: the table reports the test statistic for several number of lags (for a maximum of three lags). The regression includes a constant and a time trend. For all series except for Japan, the critical values at 1, 5, 10 percent significance levels are the following for all observations: -3.48 (1%); -2.89 (5%); -2.57 (10%). For Japan, the critical values are -3.518 ; -2.978 ; -2.688. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

.2 Chow test results

Table A.4: Chow Test Results - Global Real Rate GW – progressively-lagged inflation

	Coefficient	Std error	t-statistic
trendbreak 1349	2.087	.159	13.12
trendbreak 1557	-.197	.075	-2.63
trendbreak 1694	-.036	.074	0.48
trendbreak 1914	-.258	.213	-1.21
trendbreak 1981	.966	.224	4.31
meanbreak 1349	-.085	.008	-10.85
meanbreak 1557	.071	.029	2.49
meanbreak 1694	-.001	.033	-0.02
meanbreak 1914	.254	.191	1.33
meanbreak 1981	-.938	.207	-4.54
trend	-.760	.118	-6.42
constant	.759	.115	6.57

Note: The model includes a constant and a deterministic trend (as a fraction of the total sample size). The standard errors are based on Newey and West's (1987) Heteroskedasticity and Autocorrelation consistent (HAC) procedure where the lag length is chosen according to Lazarus et al. (2018). The critical value of the (absolute value of the) t-statistic is 2.0969 at the 5 percent significance level and 2.8266 at the 1 percent. The coefficient associated with "trendbreak 1349" denotes the difference of the estimated trend coefficients before and after a break in 1349 (allowing for breaks in all the other potential break dates); hence, the associated t-statistic is the Chow test for the absence of a structural break in 1349. Similarly, "meanbreak 1349" refers to a break in the mean in 1349 .

Per table A.4, the Chow test results indicate that for the GDP-weighted global real rate ("Global GW") we find empirical evidence in favor of breaks in the trend and the mean in 1349, 1557, and 1981. For the equal-weighted global real rate ("Global AW", not separately shown here), we also find empirical evidence of breaks in either the trend or the mean in the years 1349, 1557, and 1981.[45] It is important to note that, in the implementation of the Chow test, at any given break date we simultaneously allow for breaks at all the other potential break dates.[46]

.3 Inflation expectation variation for Chow test

We now present the Chow test applying an equal-lagged inflation basis – analogous to the Eichengreen (2015) approach (which includes current year inflation) discussed for table A.3. We observe that all the main results from the progressively-lagged approach are confirmed, in particular the relevance of the 1349 and 1914 break points.

[45]On the country level, the empirical evidence points towards the existence of a break at the following times: for the Italian real rate in 1349, 1694, 1914, and 1981 in either the trend or the mean; for the UK real rate in 1349 and 1557 in either the trend or the mean; for the French real rate in 1349, 1557, and 1981 in either the trend or the mean; for the German real rate in 1349, and 1981 in either the trend or the mean; for the Spanish real rate in 1694 and 1914 in either the trend or the mean; for the Dutch real rate in 1981 in either the trend or the mean; for the US real rate in 1981 in either the trend or the mean as well. There is no evidence of breaks in the Japanese real rate.

[46]That is, we do not tests for breaks one-at-a-time. It is also worth stressing that we test for breaks in the mean and the slope of the deterministic trend.

Table A.5: Chow Test Results - Global Real Rate GW - equal-lagged inflation			
	Coefficient	Std error	t-statistic
trendbreak 1349	2.047834	0.15912	12.87
trendbreak 1557	-0.2015759	0.0748545	-2.69
trendbreak 1694	.0439645	0.0732705	0.60
trendbreak 1914	-0.2568229	0.2036364	-1.26
trendbreak 1981	0.9735082	0.213969	4.55
meanbreak 1349	-.0835765	0.007186	-11.63
meanbreak 1557	.0747736	0.0285402	2.62
meanbreak 1694	-.004816	.0323252	-0.15
meanbreak 1914	.2522121	0.1824518	1.38
meanbreak 1981	-.9453108	.1973209	-4.79
trend	-.7694233	.1132473	-6.79
constant	.7675294	.1103859	6.95

Note: The Chow test is implemented using Newey and West (1987) standard errors where the lag length is chosen according to Lazarus et al. (2018). The critical value of the (absolute value of the) t-statistic is 2.0969 at the 5 percent significance level and 2.8266 at the 1 percent. The coefficient associated with "trendbreak 1981" denotes the difference of the estimated trend coefficients before and after a break in 1981 (and allowing for breaks in all the other potential break dates); hence, the associated t-statistic is the Chow test for the absence of a structural break in trend in 1981. Similarly, "meanbreak 1981" refers to a break in the mean in 1981.

.4 Testing alternative long-run rates series, including independently constructed U.K. data set of Dimsdale and Thomas (2016).

In this subsection, we test a number of alternative long-sample fixed income datasets offered over recent years, to determine more precisely if our stationarity results could possibly be driven by particularities in our multi-century dataset, or whether a confirmation of such results for other series suggests that an extended sample length could result in common statistical properties for additional assets and geographies. For one, Dimsdale and Thomas (2016) constructed a series of long-maturity U.K. real interest rates on an annual basis over the period of 1703-2016. The authors construct a measure of inflation expectations to adjust nominal yields, describing their approach as follows: "Long term inflation expectations are based on a simple HP filter with lambda of 100 starting in the year 1600". We take the underlying Dimsdale and Thomas (2016) U.K. nominal interest rate, and inflation data, applying the Homer and Sylla progressive weights. To our knowledge, no existing literature has tested the statistical properties of this particular series, but to further investigate whether our conclusions on stationarity are robust to a variety of methodological approaches, we have conducted an equivalent ADF-GLS test, with time trend, which strongly confirm stationarity at the 5% and and 10% levels. We take these results as further indication that our conclusions are robust across a variety of inflation and data construction methodologies. Also, this exercise gives an additional example of how our results hold even when excluding the very early centuries in the data sample. Table A.6.1 contains details, followed by the Chow test results for the same series (complementing the results for the Bai-Perron variation reported in the main body of the paper).

Next, table A.6.3 tests the statistical properties for an additional set of alternative fixed income series over the period of 1695-2021, 1853/4-2016, and 1913-2021, respectively. Our intention is to assess whether our result of trend stationarity for global sovereign real rates, as discussed in the main text above, extends to other asset classes. At least for selected countries and asset classes, this question can be addressed.

Table A.6.1: U.K. real rates 1703-2016

ADF-GLS with trend

Rates	N. of lags	ADF-GLS test statistic	Critical values 1%	5%	10%	Optimal lag
1703-2016						
U.K. Long-maturity rates	3	-3.736	-3.480	-2.890	-2.570	Seq, MAIC
	2	-4.244	-3.480	-2.890	-2.570	SIC
	1	-6.089	-3.480	-2.890	-2.570	

Note: Using U.K. real rate series constructed by Dimsdale and Thomas (2016). The table reports the ADF-GLS test statistic for several number of lags (for a maximum of three lags). The regression includes a constant and a time trend. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

Table A.6.2: Chow Test Results - Dimsdale U.K. series

	Coefficient	Std error	t-statistic
trendbreak 1914	-13.54	13.95	-0.97
trendbreak 1981	41.81	21.63	1.93
meanbreak 1914	12.23	11.29	1.08
meanbreak 1981	-39.53	19.34	-2.04
trend	-30.24	16.02	-1.89
constant	31.22	15.21	2.05

Note: The Chow test is implemented using Newey and West (1987) standard errors where the lag length is chosen according to Lazarus et al. (2018). The critical value of the (absolute value of the) t-statistic is 4.7018 at the 5 percent significance level and 8.7433 at the 1 percent. The coefficient associated with "trendbreak 1981" denotes the difference of the estimated trend coefficients before and after a break in 1981 (and allowing for breaks in all the other potential break dates); hence, the associated t-statistic is the Chow test for the absence of a structural break in trend in 1981. Similarly, "meanbreak 1981" refers to a break in the mean in 1981.

We find that, indeed, trend stationarity is confirmed for a number of additional asset bases, using equivalent real ex post rates, including British corporate bond real rates (long-maturity), and U.K. real mortgage rates (long-maturity) at the 1% critical level. Trend stationarity is equally confirmed for Bank of England real policy rates (short-maturity). The shorter U.S. real interest rate sample as collected by Shiller (2015), covering long-maturity government bonds over 1872-2015, (which we deflate here with current year year-on-year CPI change provided in the same database) on the other hand, only fails to reject non-stationarity at the 10% level. This last result is in line with our analogy to the literature on real exchange rates (e.g. Frankel (1986)), which has shown analytically that even a hundred years of data may not be enough to reject non-stationarity, given volatility of the series even with a half life of three years.

Overall, these results suggest that stationarity properties of real asset yields and returns are not restricted to sovereign assets, at least in the post-1695 period. This casts further doubt on the assertion that interest rates should be modeled as non-stationary – that they, essentially, follow a random walk over the long run.

Table A.6.3: ADF-GLS - Other Rate Variations

Series	lag	Test statistic	Critical values		
			1%	5%	10%
1 - Shiller US Real Rate, 1879-2015	MAIC	-2.379	-3.54	-2.99	-2.70
		-2.609	-3.54	-2.99	-2.70
	SIC	-2.848	-3.54	-2.99	-2.70
2 - BOE Bank Real Series, 1694-2016	MAIC	-8.612	-3.48	-2.89	-2.57
		-10.450	-3.48	-2.89	-2.57
	Seq, SIC	-12.919	-3.48	-2.89	-2.57
3 - UK Mortgage Real Series, 1853-2016		-4.298	-3.50	-2.97	-2.68
	Seq, MAIC	-4.271	-3.50	-2.97	-2.68
	SIC	-5.257	-3.50	-2.97	-2.68
4 - UK Corporate Bond Real Series, 1854-2016		-4.304	-3.50	-2.97	-2.68
	MAIC	-4.515	-3.50	-2.97	-2.68
	SIC	-5.097	-3.50	-2.97	-2.68

Notes: The three rows for each variable correspond to each of the following number of lags: 3, 2, 1. [1], Shiller real rate, uses series "U.S. long government bond rate" data in Shiller (2015), deflated by seven-year progressively lagged change in "consumer prince index" series, both over 1913-2015, and both available from the author's homepage (http://www.econ.yale.edu/ shiller/data.htm); [2-4] use British interest rate series in the BoE's Millennium dataset, via Dimsdale and Thomas (2016). Deflation is throughout undertaken via seven-year progressively lagged headline U.K. consumer price index change, via the same source. The corporate bond series combines industrial and railway debentures and is long-maturity. The BoE bank rate is defined as the minimum lending rate between 1695-1972, the minimum band 1 dealing rate for 1981-1997, the repo rate for 1997-2006, and the bank rate for 2006-2016. The estimated model is an autoregression with a constant and a deterministic trend, and a maximum number of three lags. The test rejects a unit root when the t-statistic is negative and, in absolute value, larger than the critical value.

3. Online Appendix

.1 Interpolation exercise

We proceed with an interpolation robustness exercise, which removes specific global and country-level sub-periods from our key tests which are disproportionately affected by linear interpolations. Specifically, we present two additional ADF-GLS exercises. First, removing the first two centuries entirely from our data for all countries and global levels (table A.7.1). Second, removing the following specific country-level episodes and forming the global rates on the more limited country basis: Netherlands between 1367-1470, France between 1375-1450, the U.K. pre-1650, Spain between 1670-1785 (table A.7.2). The global GDP weights are adjusted accordingly when these country-level sub-periods are removed (so that the remaining constituents are assigned proportionately greater GDP shares, based on the same Maddison calculations).

Table A.7.1: ADF-GLS Test Excluding the First Two Centuries			
Real Rate Series			
Region	Number of lags	ADF-GLS test	Optimal lag
Global GW	3	-7.659	
	2	-7.919	Seq, SIC, MAIC
	1	-9.752	
Global AW	3	-7.012	
	2	-6.890	Seq, MAIC
	1	-8.210	SIC
Italy	3	-7.321	MAIC
	2	-7.964	Seq
	1	-9.663	SIC
UK	3	-6.680	
	2	-6.537	Seq, SIC, MAIC
	1	-8.015	
Dutch	3	-5.491	Seq, MAIC
	2	-6.250	SIC
	1	-7.509	
France	3	-6.520	
	2	-6.812	MAIC
	1	-7.640	Seq, SIC
Germany	3	-8.553	
	2	-8.654	Seq, SIC, MAIC
	1	-11.346	
Spain	3	-5.768	
	2	-5.787	Seq, SIC, MAIC
	1	-7.580	
U.S.	3	-3.615	Seq, MAIC
	2	-4.226	
	1	-4.158	SIC
Japan	3	-3.877	
	2	-4.107	MAIC
	1	-4.658	Seq, SIC

Note: the table reports the test statistic for several number of lags (for a maximum of three lags). The regression includes a constant. For all series except for Japan, the critical values at 1, 5, 10 percent significance levels are the following for all observations: -3.48 (1%); -2.89 (5%); -2.57 (10%). For Japan, they are -3.525; -2.984; -2.694. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

Table A.7.2: ADF-GLS Test Excluding Selected Sub-periods			
Real Rate Series			
Region	Number of lags	ADF-GLS test	Optimal lag
Global GW	3	-6.661	Seq, SIC, MAIC
	2	-7.786	
	1	-10.246	
Global AW	3	-6.298	Seq, MAIC
	2	-7.050	SIC
	1	-9.265	

Note: the table reports the test statistic for several number of lags (for a maximum of three lags). The regression includes a constant. For both series, the critical values at 1, 5, 10 percent significance levels are the following for all observations: -3.48 (1%); -2.89 (5%); -2.57 (10%). "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

.2 Replicating Rose (1988) with two inflation expectation variations

Tables A.8.1-A.8.3 replicate Rose (1988)'s results for U.S. shorter sample unit root tests, including long-maturity corporate bond yields (CB) and short-maturity commercial paper rates (CP), using his two alternative observation periods, 1892-1970 and 1901-1950 , and using annual nominal (table A.8.1) and real rate (tables A.8.2 and A.8.3) data. We use the same data for both nominal rates and inflation, sourcing from Friedman and Schwartz (1963) and Nelson and Plosser (1982): the latter data is taken from Peter C.B. Phillips data page (http://korora.econ.yale.edu/phillips/data/npenp.dat). Table A.8.2 uses the progressively-lagged inflation expectation approach as in Homer and Sylla (2005); table A.8.3 uses the equal-lagged approach as in Eichengreen (2015)).

We observe that once we apply our two alternative inflation expectation approaches to his underlying data - which for the underlying price data is based on a GNP deflator series, and alternatively a CPI index, both of which we test - and construct equivalent real rate samples we are unable to reject a unit root for all progressively-lagged real rate variations (table A.8.2), as well as for equal-lagged variations (table A.8.3), with no exceptions. In particular, we emphasize that we replicate Rose's results when allowing for a deterministic trend in ADF-GLS.

Table A.8.1: Rose (1988) replication

Rates	N. of lags	ADF-GLS test statistic	Critical values 1%	5%	10%	Optimal lag
Nominal rate series						
1892-1970						
Corporate bond yields	3	-0.793	-3.660	-3.097	-2.803	
	2	-0.523	-3.660	-3.097	-2.803	
	1	-0.246	-3.660	-3.097	-2.803	Seq, SIC, MAIC
Commercial paper rates	3	-1.429	-3.660	-3.097	-2.803	Seq, SIC, MAIC
	2	-0.785	-3.660	-3.097	-2.803	
	1	-1.763	-3.660	-3.097	-2.803	
1901-1950						
Corporate bond yields	3	-1.494	-3.770	-3.190	-2.890	
	2	-1.212	-3.770	-3.190	-2.890	
	1	-1.355	-3.770	-3.190	-2.890	Seq, SIC, MAIC
Corporate bond yields	3	-2.164	-3.770	-3.190	-2.890	
	2	-1.766	-3.770	-3.190	-2.890	MAIC
	1	-2.405	-3.770	-3.190	-2.890	Seq(0), SIC

Note: The table reports the ADF-GLS test statistic for several number of lags (for a maximum of three lags). The regression includes a constant and a time trend. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). Seq(0) denotes cases where the sequential lag length procedure selects zero lags. The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

Table A.8.2: Rose (1988) replication

Real rate series - progressively lagged inflation

Rates	N. of lags	ADF-GLS test statistic	Critical values 1%	5%	10%	Optimal lag
GNP Deflator						
1892-1970						
Corporate bond yields	3	-2.669	-3.660	-3.097	-2.803	
	2	-2.757	-3.660	-3.097	-2.803	
	1	-3.026	-3.660	-3.097	-2.803	Seq, SIC, MAIC
Commercial paper rates	3	-2.452	-3.660	-3.097	-2.803	Seq
	2	-2.048	-3.660	-3.097	-2.803	MAIC
	1	-2.570	-3.660	-3.097	-2.803	SIC
1901-1950						
Corporate bond yields	3	-2.257	-3.770	-3.190	-2.890	
	2	-2.340	-3.770	-3.190	-2.890	
	1	-2.614	-3.770	-3.190	-2.890	Seq, SIC, MAIC
Commercial paper rates	3	-2.261	-3.770	-3.190	-2.890	Seq, SIC, MAIC
	2	-2.084	-3.770	-3.190	-2.890	
	1	-2.410	-3.770	-3.190	-2.890	
CPI Index						
1892-1970						
Corporate bond yields	3	-2.751	-3.660	-3.097	-2.803	
	2	-2.655	-3.660	-3.097	-2.803	Seq, MAIC
	1	-3.533	-3.660	-3.097	-2.803	SIC
Commercial paper rates	3	-2.571	-3.660	-3.097	-2.803	Seq, SIC
	2	-2.054	-3.660	-3.097	-2.803	MAIC
	1	-2.873	-3.660	-3.097	-2.803	
1901-1950						
Corporate bond yields	3	-2.196	-3.770	-3.190	-2.890	
	2	-2.189	-3.770	-3.190	-2.890	MAIC
	1	-3.005	-3.770	-3.190	-2.890	Seq, SIC
Commercial paper rates	3	-2.246	-3.770	-3.190	-2.890	
	2	-2.029	-3.770	-3.190	-2.890	MAIC
	3	-2.677	-3.770	-3.190	-2.890	Seq(0), SIC

Note: The inflation approach follows Homer and Sylla (2005), using seven-year progressively lagged inflation (t-7 to t-1, excluding the current-year inflation t). The GNP deflator is from Nelson and Plosser (1982); the CPI index is from Nelson and Plosser (1982). The table reports the ADF-GLS test statistic for several number of lags (for a maximum of three lags). The regression includes a constant and a time trend. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). Seq(0) denotes cases where the sequential procedure selects zero lags. The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

Table A.8.3: Rose (1988) replication						
Real rate series - equal lagged inflation						
Rates	N. of lags	ADF-GLS test statistic	Critical values 1%	5%	10%	Optimal lag
GNP Deflator						
1892-1970						
Corporate bond yields	3	-2.303	-3.660	-3.097	-2.803	
	2	-2.220	-3.660	-3.097	-2.803	
	1	-2.078	-3.660	-3.097	-2.803	Seq, SIC, MAIC
Commercial paper rates	3	-2.080	-3.660	-3.097	-2.803	Seq
	2	-1.428	-3.660	-3.097	-2.803	
	1	-1.659	-3.660	-3.097	-2.803	SIC, MAIC
1901-1950						
Corporate bond yields	3	-2.214	-3.770	-3.190	-2.890	
	2	-2.130	-3.770	-3.190	-2.890	
	1	-1.945	-3.770	-3.190	-2.890	Seq, SIC, MAIC
Commercial paper rates	3	-2.135	-3.770	-3.190	-2.890	
	2	-1.886	-3.770	-3.190	-2.890	
	1	-1.969	-3.770	-3.190	-2.890	Seq, SIC, MAIC
CPI Index						
1892-1970						
Corporate bond yields	3	-2.657	-3.660	-3.097	-2.803	
	2	-2.438	-3.660	-3.097	-2.803	MAIC
	1	-2.793	-3.660	-3.097	-2.803	Seq, SIC
Commercial paper rates	3	-2.857	-3.660	-3.097	-2.803	Seq, SIC, MAIC
	2	-1.653	-3.660	-3.097	-2.803	
	1	-2.080	-3.660	-3.097	-2.803	
1901-1950						
Corporate bond yields	3	-2.196	-3.770	-3.190	-2.890	
	2	-2.137	-3.770	-3.190	-2.890	
	1	-2.339	-3.770	-3.190	-2.890	Seq, SIC, MAIC
Commercial paper rates	3	-2.486	-3.770	-3.190	-2.890	Seq
	2	-1.870	-3.770	-3.190	-2.890	
	3	-2.152	-3.770	-3.190	-2.890	SIC, MAIC

Note: The inflation approach follows Eichengreen (2015), using seven-year equal lagged inflation (*t-6* to *t*, thus including the current-year inflation *t*. The GNP deflator is from Nelson and Plosser (1982); the CPI index is from Nelson and Plosser (1982). The table reports the ADF-GLS test statistic for several number of lags (for a maximum of three lags). The regression includes a constant and a time trend. "Optimal lag" indicates the optimal number of lags according to the sequential procedure ("Seq"), the SIC, or the Modified Information Criterion (MAIC). The test rejects when the test statistic is negative and larger (in absolute value) than the critical value.

.3 Monte Carlo simulations on the power of the ADF-GLS test for the global real rate and out-of-sample forecast analysis

To investigate the use of long versus small sample sizes in testing for unit roots in our data, we performed a small Monte Carlo simulation exercise. We generated 5000 time series that have the same features as our (stationary) global real rate (such as the mean, the deterministic trend and the standard error): $y_t = 0.7289392y_{t-1} + 0.0331001 - 0.0000467t + 0.02179\varepsilon_t$, where ε_t is an independent, zero mean and unit variance Gaussian random variable. For each generated series, we calculated: (a) the ADF-GLS test statistic for a sample of 705 observations (the same size as our full sample); (b) the ADF-GLS test statistic for the sub-sample including the first 100 observations; (c) the ADF-GLS test statistic for the sub-sample including the first 50 observations. We then report the number of times the ADF-GLS test rejects a unit root at the 5 percent significance level across Monte Carlo simulations.

The simulation results indicate that the ADF-GLS test statistic rejects a unit root 100 percent of the times in the full sample, whereas it does so only 88 percent of the time in the sample with 75 observations and 63 percent of the time in the sample with 50 observations. This Monte Carlo simulation confirms the lack of power of the ADF-GLS test for time series such as the global real rate when the span of the data is short.

The average estimated half-life across Monte Carlo simulations is about 5 periods (as the random series are generated to share the same features as the observed data). Thus, our estimated half-lives are not incompatible with the fact that the test would need a relatively large sample (such as our full sample of 705 observations) to have large power.

In the same Monte Carlo exercise, we also compare the empirical rejection frequencies of the ADF-GLS test implemented with and without a deterministic trend included in the regression. We consider the same data generating process as above, except that we let the magnitude of the trend coefficient vary and take the following values: -0.00005, -0.0001, -0.00012, -0.00014, -0.00016, -0.00018, -0.0002. The first value is close to the estimated value, and then we let it progressively increases in absolute value. In a sample of 705 observations, like the one in our dataset, the empirical rejection frequency of the ADF-GLS test implemented in a regression including a deterministic trend is 1, while that of the ADF-GLS test implemented in a (mis-specified) regression without a deterministic trend are 1, 0.36, 0.05, 0.001, 0, 0, 0, respectively, as the trend coefficient changes from -0.00005 to -0.0002. In other words, the (mis-specified) ADF-GLS test without a trend becomes progressively inconsistent as the trend magnitude increases (in absolute value); however, in the case of a small trend like the one in our data, the rejection frequencies are still close to the nominal value.

Results are qualitatively similar in a small sample of 100 observations. In that case, as the magnitude of the trend coefficient takes on the values -0.00005,-0.0001,-0.0002,-0.0003,-0.0004,-0.0005,-0.001, the empirical rejection frequency of the ADF-GLS test implemented in a regression including a deterministic trend is 0.97, while that of the ADF-GLS test implemented in a (mis-specified) regression without a deterministic trend are 0.97, 0.95, 0.78, 0.43, 0.12, 0.001 and 0, respectively, as the trend coefficient varies from -0.00005 to -0.001.

On the other hand, however, if there were no trend in the data, the ADF-GLS test without a trend would be the most efficient one, which motivates our interest in reporting results for that test in Table 1.

Finally, we show that the trend-stationary model provides better forecasts than the random walk model at various forecast horizons. This evidence provides additional support to the trend-stationary specification considered in our paper. We perform an out-of-sample forecast exercise, as follows: we split the sample in two equal parts, and progressively re-estimate the trend in real time using a rolling window of past observations (of size equal to half of the total sample size). We calculate the ratio of the out-of-sample mean squared forecast error of the trend model proposed in this paper versus that of a random walk, as a function of the forecast horizon. Values smaller than one indicate that the trend model forecasts better than the random walk model. At the one-step-ahead horizon, the ratio is 3.14, while at two-step-ahead it lowers to 1.36, at three-step-ahead it is 1.03; at any longer horizons, the ratio is always less than one: at horizon four, it is 0.88; at horizon ten, it is 0.59; at horizon fifty it is 0.57. Overall, the trend model forecasts better than the random walk at horizons bigger than three years. Results are similar when comparing the trend model with an autoregressive model.

.4 Phillips-Perron's Test and Cochrane's Variance Ratios

To verify the robustness of the unit root tests to the presence of possible moving average components, we report results based on Phillips and Perron's (1988) test as well as Cochrane's (1988) variance ratios test. Both confirm our main finding. They are reported in Tables A.9 and A.10, respectively. The Phillips-Perron's test rejects the presence of a unit root and Cochrane's (1988) variance ratios highlight mean reversion.

Table A.9: Phillips Perron Test with Time Trend

Real Rate Series					
		Test Statistic	Critical Value		
			1%	5%	10%
Global GW	Z_ρ	-158.413	-29.50	-21.80	-18.30
	Z_τ	-9.751	-3.96	-3.41	-3.12
Global AW	Z_ρ	-174.584	-29.50	-21.80	-18.30
	Z_τ	-9.867	-3.96	-3.41	-3.12
Italy	Z_ρ	-229.871	-29.50	-21.80	-18.30
	Z_τ	-12.060	-3.96	-3.41	-3.12
UK	Z_ρ	-403.407	-29.50	-21.80	-18.30
	Z_τ	-14.665	-3.96	-3.41	-3.12
Dutch	Z_ρ	-357.032	-29.50	-21.80	-18.30
	Z_τ	-14.572	-3.96	-3.41	-3.12
France	Z_ρ	-162.886	-29.50	-21.80	-18.30
	Z_τ	-9.468	-3.96	-3.41	-3.12
Germany	Z_ρ	-127.458	-29.50	-21.80	-18.30
	Z_τ	-8.735	-3.96	-3.41	-3.12
Spain	Z_ρ	-275.605	-29.50	-21.80	-18.30
	Z_τ	-12.098	-3.96	-3.41	-3.12
U.S.	Z_ρ	-34.235	-28.23	-21.20	-17.92
	Z_τ	-4.615	-4.00	-3.43	-3.13
Japan	Z_ρ	-19.213	-27.70	-20.88	-17.65
	Z_τ	-3.263	-4.03	-3.44	-3.14

Note: The table reports the Phillips-Perron Z_ρ and Z_τ test statistics and their critical values at the 1, 5 and 10 percent significance levels. The regression includes a constant and a deterministic time trend. The standard errors are obtained by Newey–West's HAC estimator. The test rejects when the test statistic is negative and larger (in absolute value) than the critical value. "Global GW" = GDP-weighted global real rate basis; "Global AW" = arithmetically-weighted global real rate basis, taking equal weights for each of the eight countries.

Table A.10: Variance Ratios: Real Interest Rates

Horizon	5	10	15	20	25	30	35	40	45	50
Global GW	0.815	0.723	0.673	0.642	0.620	0.607	0.598	0.592	0.585	0.577
Global AW	0.861	0.784	0.738	0.706	0.684	0.668	0.656	0.646	0.634	0.621
Italy	0.550	0.375	0.299	0.259	0.236	0.225	0.220	0.214	0.208	0.201
UK	0.872	0.813	0.786	0.770	0.758	0.747	0.736	0.727	0.717	0.706
Dutch	0.557	0.410	0.338	0.294	0.266	0.250	0.239	0.229	0.219	0.209
France	0.812	0.701	0.634	0.593	0.568	0.552	0.542	0.532	0.523	0.514
Germany	0.666	0.472	0.391	0.341	0.304	0.277	0.256	0.241	0.229	0.219
Spain	0.734	0.627	0.583	0.544	0.514	0.488	0.465	0.444	0.422	0.401
US	0.707	0.489	0.360	0.283	0.239	0.212	0.197	0.188	0.181	0.169
Japan	0.721	0.488	0.327	0.208	0.135	0.099	0.082	0.064	0.050	0.043

Note: The table reports variance ratios calculated as the variance of cumulative returns over h-periods horizons (where the horizon is indicated in the top line) divided by the variance of the one-period return times the horizon - see Cochrane (1988).

.5 Discussion of Chow break dates, historical context, and balanced panel exercise

From the discussion of statistical properties, we now proceed to an attempt at broader interpretations of historical patterns and possible economic drivers. How do the statistical results accord with economic and historical evidence, and to what extent do structural breaks and general trends fit the empirics of other macroeconomic variables?

We concentrate here on a discussion of the much-discussed significance of "recent" trend breaks in global real rates, particularly in 1914 and the 1980s. We recall that in our (unbalanced) panel above, we found some evidence of 1914 as a break – while 1981 was confirmed in multiple time series. Meanwhile, 1349 or 1557 show a clear constancy across all global and country-levels, and there is sufficient ex-post data coverage making these breaks robust.

In tables A.11.1 and A.11.2. we now test for structural breaks in a balanced panel, to take account of the fact that the U.S. and Japanese series begin notably later than the European observations (in 1790 and 1870, respectively). In this exercise, we initiate all series in 1870 and end in 2016: below, the results are shown for the GDP-weighted global real series ("Global GW"), and for the arithmetically-weighted global real series ("Global AW"). We test the 20th century break points (1914 and 1981) once more in this set-up via the Chow test.

Table A.11.1: Chow Test Results - balanced panel Global GW - progressively-lagged inflation

Regressor	Coefficient	Std error	t-statistic
trendbreak 1914	-0.181	0.056	-3.21
trendbreak 1981	0.181	0.069	2.62
meanbreak 1914	0.064	0.036	1.77
meanbreak 1981	-0.164	0.048	-3.37
trend	-0.140	0.034	-4.11
constant	0.147	0.031	4.77

Note: The model includes a constant and a deterministic trend (as a fraction of the total sample size). The standard errors are based on Newey and West's (1987) HAC estimator where the lag length is chosen according to Lazarus et al. (2018). The critical value of the (absolute value of the) t-statistic is 5.2197 at the 5 percent significance level and 10.0664 at the 1 percent. The coefficient associated with "trendbreak 1981" denotes the difference of the estimated trend coefficients before and after a break in 1981 (allowing for breaks in all the other potential break dates); hence, the associated t-statistic is the Chow test for the absence of a structural break in trend in 1981. Similarly, "meanbreak 1981" refers to a break in the mean in 1981.

Once we focus on this balanced panel, with its much shorter time period, neither 1914 nor 1981 emerge as break dates. Given space, we do not report country results here although neither 1914 nor 1981 emerge as break dates in any of the countries. Our intuition therefore remains that narratives of a "recent", a 1980s inflection point in global real rates, are partly a function of the too short sample length that existing research was able to utilize (overall, we interpret our results to be a function of three factors: our approach of allowing for a trend in the data, of our much longer sample length, and of the fact that we focus on long-maturity rates). Simply put, inflections that appear significant in the relatively short "balanced panel" modern time period, may appear as large – but ultimately transitory – shocks when multi-century data is used.

Table A.11.2: Chow Test Results - balanced panel Global AW - progressively-lagged inflation			
Regressor	Coefficient	Std error	t-statistic
trendbreak 1914	-0.187	0.062	-3.01
trendbreak 1981	0.159	0.078	2.04
meanbreak 1914	0.074	0.038	1.93
meanbreak 1981	-0.158	0.058	-2.73
trend	-0.129	0.047	-2.72
constant	0.139	0.043	3.22

Note: The model includes a constant and a deterministic trend (as a fraction of the total sample size). The standard errors are based on Newey and West's (1987) HAC estimator where the lag length is chosen according to Lazarus et al. (2018). The coefficient associated with "trendbreak 1981" denotes the difference of the estimated trend coefficients before and after a break in 1981 (allowing for breaks in all the other potential break dates); hence, the associated t-statistic is the Chow test for the absence of a structural break in trend in 1981. Similarly, "meanbreak 1981" refers to a break in the mean in 1981.

While we refrain from sweeping generalization about the actual underlying factors behind the confirmed break dates across structural break tests (a question that deserves a separate detailed investigation), we provide some intuition here behind our choice of the five break dates that we "imputed" into the Chow tests, which test the validity of specific *known* potential break dates. Our choices were grounded in long-standing economic and historical literature, and strong claims surrounding the particular episodes and exact years. The following list discusses the historical context and existing literature that motivated the selection of our five specific Chow break dates: what is the particular logic of selecting these dates? And what do our results suggest for the economic-financial context of the five break years?

- **1349.** This year was chosen against the backdrop of a long literature positing a major financial trend break associated with the Black Death (Epstein, 2000; Pamuk, 2007; Clark, 2016). The death of one-third to one-half of the population is said to have created a boost in capital-per-capita, and a substantial increase in real wages resulting from labor scarcity. As we detail below, we confirm the epochal role of the Black Death on credit markets: however, the directional context pre- and post-1349 appears more idiosyncratic than recognized thus far.
- **1557.** While little progress on our understanding on the subject has been made since an authoritative study appeared in German in 1896 (Ehrenberg (1896), only partially translated via Ehrenberg (1928)) – which regarded the initial wave of shocks over 1557-62 as events "that shook the finance and trade of Europe to its foundation" – a variety of case-study and piecemeal literature appearing since confirmed that an epochal crisis afflicted the international economy during the second half of the 16th century. Indications are that the "triple sovereign default of 1557-8" (France, Spain, and the States General) could have been either a consequence or the actual inception point of a very deep-seated reversal in financial markets lasting at least to the early 17th century.[47] In direct consequence of the sovereign defaults, by far the largest bank of its day – perhaps the most significant "systemically important financial institution" ever to have existed – the Fuggers, slid to the brink of default. Narrowly escaping

[47]The Netherlands are not a formally recognized independent nation state until the Peace of Westphalia 1648: de facto, however, the States are characterized by a high degree of autonomy, with the Receiver General of the States General, and the individual provinces, issuing their own debt secured by land and other taxes, the *bede* Fritschy (2017, chapter 2). Ehrenberg (1928, 113f.) puts the Fugger exposure alone at 600,000fl., which the Receiver General refused to honor in 1557: it is not specified in the source whether the debt was denominated fully in Dutch guilders or Ducats. Hauser (1930) dates the French default to September 1, 1557, with other sources dating the default to the year 1558.

themselves, a seemingly unending string of international merchant and merchant-bank defaults can afterwards be traced in the respective sources. Recurring chaos at the largest financial fairs over decades – from Seville over Lyon, to Rome – is reported in equal measure (Ehrenberg, 1928; Lapeyre, 1955; Delumeau, 1959; Kindleberger, 1998).[48] We chose this date mainly to test for a major early modern standalone sovereign default-cum-private financial crisis event: our affirmative results on a break date here indicate that secularly, though financial turmoil on its own may still not be a sufficient driver of structural inflections – at least when accompanied by significant political volatility (the French Wars of Religion; the near-uninterrupted Habsburg military campaigns in the 16th century) these combined forces can culminate into deep structural change.

- **1694.** This year was chosen against the backdrop of the prominent "North Weingast" thesis, which associated the Glorious Revolution with a revolution in "credible commitments" (North and Weingast, 1989). It appears particularly relevant that the authors used (nominal) interest rate data to prove their seminal thesis. Yet, this event is widely rejected in our Chow tests, on both the global levels and all country levels. Of course, the original statement of the North-Weingast thesis received extensive criticism over the years (Sussman and Yafeh, 2006) – but has not been contextualized in the setting of a comprehensive (real) interest rate data set covering centuries. The failure to identify 1694 with a structural break is part of more comprehensive evidence that fails to associate all other central bank inceptions with structural breaks, too.[49]

- **1914.** This year was chosen against the backdrop of the strong narrative of an institutional inflection arising from the founding of the Federal Reserve (Barsky et al., 1988; Bernstein et al., 2010), coupled with the major monetary inflection related to the (de facto) departure from the centuries-old bullion standard. In addition, the economies deal with the geopolitical shock of the First World War – the first industrial war far outstripping the human and financial costs of earlier inter-state conflicts. And indeed, across our series, the years in and around the First World War constitute strong (though not fully consistent) breaks across global and country-levels: ironically, however, the U.S. are not one of these. As in 1694, we therefore fail to confirm any obvious monetary channel influencing real rates.

- **1981.** This year was chosen against the backdrop of prevalent narratives – intensifying after 2008 – of a key inflection point in advanced economies and financial markets during the early 1980s, though there is no clear consensus on a single driver of this alleged inflection (Rachel and Summers, 2019; Mian et al., 2021; Goodhart and Pradhan, 2021). Antedating the 2008 crisis, a sizable literature found evidence of structural changes in inflation and interest rates during the late 1970s or early 1980s (Garcia and Perron, 1996; Ang and Bekaert, 2002; Neely and Rapach, 2008) – coupled with sharply declining general macroeconomic volatility (Bernanke, 2004) – invoking monetary policy dynamics as the driving causal force.

[48] Ehrenberg (1928, 114ff.) details the uncertainty over the Fugger's survival in 1557-8, and also presents balance sheets of the bank: he indicates that the Fuggers had formally written off more than 620,000fl from the Dutch and Spanish defaults by 1563, against a balance sheet size of 5.6M fl. Ehrenberg regards 1557 as the beginning of the end of the Fugger empire. Significant parts of his work remain untranslated; the contemporary account of Hauser (1930) relies mainly on him.

[49] The same failure to associate central bank inceptions with structural breaks is observed when undertaking Bai-Perron tests, see table 3 and discussion.

.6 Half-lives, continued country-level results

Table A.12: Half-Lives of Real Rates, by Sub-samples, continued

	Period	Half-life	Confidence interval		Robust Half-life	Confidence interval	
Italy Real	1318-1348	1.29	1.01	1.59	1.07	0.35	1.82
	1349-1556	1.26	1.00	1.53	1.21	0.76	1.67
	1557-1693	1.45	1.19	1.73	1.44	0.81	1.96
	1694-1913	1.85	1.56	2.23	1.75	1.17	2.50
	1914-1980	10.23	2.82	N.A.	6.01	0.84	9.15
	1981-2022	4.53	2.33	16.75	2.90	0.86	7.35
Netherlands Real	1374-1556	1.01	0.86	1.43	1.09	0.29	2.61
	1557-1693	1.02	0.79	1.50	1.05	0.61	1.91
	1694-1913	2.22	1.70	2.83	1.97	0.71	4.17
	1914-1980	7.06	3.19	31.93	6.34	1.14	10.94
	1981-2022	3.73	2.31	8.18	2.74	1.04	4.68
France Real	1318-1348	1.18	0.63	3.32	0.95	0.45	2.52
	1349-1556	1.39	1.14	1.65	1.36	0.79	1.93
	1557-1694	1.84	1.24	2.81	1.50	0.71	5.54
	1694-1913	2.16	1.77	2.65	2.22	0.73	3.77
	1914-1980	7.80	3.44	39.64	6.36	0.94	8.99
	1981-2022	6.10	1.69	N.A.	2.44	0.68	N.A.
Spain Real	1311-1348	2.08	0.60	N.A.	1.34	0.48	5.57
	1349-1556	1.15	0.88	1.63	1.22	0.63	3.00
	1557-1693	1.94	1.32	2.94	1.41	0.63	2.70
	1694-1913	1.91	1.62	2.83	1.84	0.54	4.94
	1914-1980	4.23	2.54	8.42	4.27	1.47	7.23
	1981-2022	3.64	1.80	12.08	2.95	0.87	7.34
Japan Real	1870-1913	2.39	0.91	9.86	1.89	0.70	3.42
	1914-1980	6.65	3.61	17.73	6.14	2.27	10.70
	1981-2022	2.51	1.52	4.41	2.11	0.66	3.81

Note: The half-life is estimated as the first horizon at which the impulse response equals one-half of the initial impact effect. The impulse response is estimated using a deterministically detrended linear autoregressive model (with the number of lags chosen by the Bayesian Information Criterion) using Kilian (1999)'s bootstrap; the point estimate is based on the median unbiased response while the confidence interval is based on the 5-th and 95-th quantiles of the bootstrap distribution. The robust half-lives and their confidence intervals are estimated based on lag-augmented local projections on a model with a constant and a deterministic time trend. The half-lives are separately estimated in each sub-sample identified by the statistically significant break dates (according to the Chow test). NA denotes situations where the half-life is infinity.