ONLINE APPENDIX

"Long-Run Trends in Long-Maturity Real Rates, 1311-2022"
Kenneth Rogoff and Barbara Rossi and Paul Schmelzing.

1. Introduction.

This online appendix contains two sections. The first section presents major robustness exercises. The second section presents extra results.

2. Main Robustness Results

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: our main results in this alternative setting are broadly 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 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 -2.583; -1.95; -1.619. For Japan, they are -2.594; -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 -3.524; -2.983; -2.693. "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.[46] 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.[47]

.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.

[46] 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.

[47] 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 2.1684 at the 5 percent significance level and 2.9569 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. Additional Robustness Results

.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						
Nominal rate series						
Rates	N. of lags	ADF-GLS test statistic	Critical values 1%	5%	10%	Optimal lag
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	
	2	-2.084	-3.770	-3.190	-2.890	
	1	-2.410	-3.770	-3.190	-2.890	Seq(0), SIC, MAIC
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
	1	-2.677	-3.770	-3.190	-2.890	Seq, 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) as estimated in the regression underlying Table 4, calibrated for the Global GW series: $y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + \epsilon_t$, where ϵ_t is an independent Gaussian random variable with zero mean and variance 0.0004402; $\alpha = 0.7332$; $\gamma = [0.1950; -0.1504; -0.0401]$; $\mu_0 = 0.0325$; and $\mu_1 = -0.000046187$.

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 only 150 observations; (c) the ADF-GLS test statistic for the sub-sample including 100 observations; (d) the ADF-GLS test statistic for the sub-sample including 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 99.98 percent of the times in the full sample, whereas it does so only 90.12 percent of the time in a sample of 150 observations, only 68.84 percent of the time in the sample with 100 observations, 49.36 percent in a sample of 75 observations and 29.8 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 2 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 (μ_1) vary and take the following values:-0.000046187,-0.00005,-0.00006,-0.000065,-0.0007. The first value is the estimated value, and then we let it progressively increase 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 0.82, 0.76, 0.54, 0.42, 0, respectively, as the trend coefficient changes from -0.000046187 to -0.0007. 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 -0.000046187, -0.0001, -0.0002, -0.0003 and -0.0004, the empirical rejection frequency of the ADF-GLS test implemented in a regression including a deterministic trend is around 0.70, while that of the ADF-GLS test implemented in a (mis-specified) regression without a deterministic trend are 0.80, 0.70, 0.36, 0.10 and 0.01, respectively, as the trend coefficient varies from -0.000046187 to -0.0004.

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 A.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). The exercise is a pseudo out-of-sample forecast analysis based on the actual data for the global GW series, not on Monte Carlo simulations. 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 autoregressive model with the same number of lags but with a unit root and no deterministic components,[48] as a function of the forecast horizon. Values smaller than one indicate that the trend model forecasts better than the unit root model with no deterministic trend. The forecasts of the former are better than those of the latter at every horizon (the ratio of the former divided by the latter is 0.9145, 0.8541, 0.7920, 0.7307, 0.5376, 0.5431 at horizons 1, 2, 3, 4, 10, 50, respectively). Results are qualitatively similar when comparing the forecasts of the model underlying Table 4 with a pure random walk; the pure random walk performs worse.

.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.

---

[48] We have compared the forecasts of the same model as in Table 4 (deterministic trend model with serial correlation, as in Steinsson, 2008): $y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \sum_{j=1}^{p} \gamma_j \, \Delta y_{t-j} + \varepsilon_t$ with that of an autoregressive model (AR(p)) with a unit root and no deterministic trend components ($\mu_0 = \mu_1 = 0$): $y_t = y_{t-1} + \sum_{j=1}^{p} \gamma_j \, \Delta y_{t-j} + \varepsilon_t$; p is chosen based on the MAIC criterion.

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 a 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.

Another way to look at the degree of mean reversion is to look at variance ratios. Note that the forecast error variance decreases significantly as we extend the forecast horizon anywhere from 2 to 14 years; at longer horizons, say from 20 years to 30 years, the drop is no longer significant, and there are cases where the mean forecast error variance can rise at very long horizons, but this is very small. Assuming stationarity, these results make perfect sense given the relatively fast rate of convergence we find in table 4 since shocks die out relatively quickly, so that over very long horizons, the forecast error begins to flatten out.

Table A.10: Variance Ratios: Real Interest Rates

h	2	4	8	10	12	14	20	30	40	50
Global GW										
var. ratio	1.0666	0.7218	0.4575	0.3649	0.3212	0.2768	0.1977	0.1124	0.1015	0.0849
test stat.	1.4341	-3.1522	-4.0094	-4.1662	-4.0541	-3.9989	-3.7217	-3.3885	-2.9917	-2.7458
p-value	0.1515	0.0016	0.0001	0.0000	0.0001	0.0001	0.0002	0.0007	0.0028	0.0060
Global AW										
var. ratio	1.0357	0.7225	0.4902	0.4012	0.3582	0.3122	0.2180	0.1270	0.1171	0.0970
test stat.	0.5810	-2.5611	-3.2791	-3.4797	-3.4406	-3.4445	-3.3424	-3.1395	-2.8176	-2.6294
p-value	0.5612	0.0104	0.0010	0.0005	0.0006	0.0006	0.0008	0.0017	0.0048	0.0086
Italy										
var. ratio	1.0060	0.6432	0.3400	0.2681	0.2398	0.2020	0.1330	0.0893	0.0783	0.0670
test stat.	0.1176	-3.5451	-4.3205	-4.2599	-4.0255	-3.9078	-3.5614	-3.0952	-2.7465	-2.5115
p-value	0.9064	0.0004	0.0000	0.0000	0.0001	0.0001	0.0004	0.0020	0.0060	0.0120
U.K.										
var. ratio	0.7875	0.5221	0.3103	0.2457	0.2106	0.1856	0.1252	0.0914	0.0668	0.0635
test stat.	-1.1514	-1.5431	-1.6944	-1.7342	-1.7262	-1.7102	-1.6910	-1.6256	-1.5903	-1.5377
p-value	0.2496	0.1228	0.0902	0.0829	0.0843	0.0872	0.0908	0.1040	0.1118	0.1241
Dutch										
var. ratio	0.8298	0.4710	0.2828	0.2117	0.1634	0.1256	0.1006	0.0633	0.0579	0.0457
test stat.	-0.8487	-1.6549	-1.8057	-1.8880	-1.9291	-1.9546	-1.8838	-1.8454	-1.7881	-1.7587
p-value	0.3960	0.0979	0.0710	0.0590	0.0537	0.0506	0.0596	0.0650	0.0738	0.0786
France										
var. ratio	0.9954	0.7209	0.5157	0.4175	0.3782	0.3320	0.2182	0.1455	0.1366	0.1211
test stat.	-0.0809	-2.7272	-3.1610	-3.3946	-3.3208	-3.3285	-3.3542	-3.1090	-2.7896	-2.5975
p-value	0.9355	0.0064	0.0016	0.0007	0.0009	0.0009	0.0008	0.0019	0.0053	0.0094
Germany										
var. ratio	1.1629	0.8611	0.5097	0.3738	0.3111	0.2833	0.2215	0.1432	0.1040	0.0882
test stat.	1.4581	-0.6760	-1.6988	-2.0135	-2.1001	-2.0975	-2.0941	-2.1058	-2.0579	-1.9850
p-value	0.1448	0.4990	0.0894	0.0441	0.0357	0.0360	0.0362	0.0352	0.0396	0.0471
Spain										
var. ratio	0.9319	0.6125	0.2997	0.2585	0.2431	0.2282	0.1565	0.1281	0.1098	0.0913
test stat.	-0.7211	-1.5269	-1.8800	-1.8384	-1.7701	-1.7215	-1.6813	-1.5461	-1.4691	-1.4224
p-value	0.4708	0.1268	0.0601	0.0660	0.0767	0.0852	0.0927	0.1221	0.1418	0.1549
U.S.										
var. ratio	1.1109	1.0698	0.6773	0.5934	0.5643	0.4790	0.3221	0.1904	0.2086	0.1683
test stat.	1.2345	0.4364	-1.3286	-1.4749	-1.4378	-1.5998	-1.7223	-1.6760	-1.4241	-1.3517
p-value	0.2170	0.6625	0.1840	0.1402	0.1505	0.1097	0.0850	0.0937	0.1544	0.1765
Japan										
var. ratio	1.2621	1.2138	0.8962	0.7833	0.7246	0.6861	0.4421	0.2806	0.3175	0.2597
test stat.	1.8727	0.8446	-0.2744	-0.5137	-0.6011	-0.6413	-0.9944	-1.1152	-0.9610	-0.9501
p-value	0.0611	0.3983	0.7838	0.6075	0.5477	0.5213	0.3200	0.2648	0.3365	0.3420

Note: The table reports variance ratios ('var. ratio') together with their test statistic ('test stat.') and p-values – 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 1877, respectively). In this exercise, we initiate all series in 1877 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	0.049	-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 2.2847 at the 5 percent significance level and 3.1728 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, 1914 and 1981 occasionally emerge as break dates for some countries. Given space, we do not report country results here although a break in 1981 emerges for France, Germany and the U.S. while a break in 1914 emerges for Italy and the US, while the UK, Spain and Japan experience no breaks. Our intuition therefore is 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 critical value of the (absolute value of the) t-statistic is 2.2847 at the 5 percent significance level and 3.1728 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.

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.[49] In direct consequence of the sovereign

[49]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.

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 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).[50] 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.[51]

- **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.

[50] 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.

[51] 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, Full Sample and Sub-samples, continued

	α	Confid. Interv. α	h	Confid. Interv. h
Italy				
1318-2022	0.60	(0.53; 0.67)	1.68	(1.54 ; 1.85)
1750-2022	0.83	(0.76; 0.89)	3.93	(2.72 ; 6.21)
1914-2022	0.85	(0.74; 0.94)	6.34	(4.23 ; 12.16)
France				
1318-2022	0.81	(0.77; 0.85)	2.52	(2.06 ; 3.00)
1750-2022	0.86	(0.81; 0.91)	5.71	(4.27 ; 7.86)
1914-2022	0.87	(0.77; 0.95)	6.95	(4.58 ; 14.24)
Spain				
1318-2022	0.83	(0.79; 0.87)	1.93	(1.74 ; 3.97)
1750-2022	0.83	(0.76; 0.90)	2.23	(1.77 ; 5.24)
1914-2022	0.87	(0.76; 0.96)	5.77	(3.56 ; 17.91)
Japan				
1877-2022	0.82	(0.73; 0.90)	4.18	(2.92 ; 6.86)
1877-2022	0.83	(0.73; 0.90)	4.23	(2.96 ; 7.13)
1914-2022	0.90	(0.79; 0.98)	6.97	(3.77 ; 40.63)

The table reports median unbiased estimates and 90% confidence intervals of α based on Hansen (1999)'s grid-bootstrap as well as median unbiased estimates and 90% confidence intervals of the half life (h) based on Steinsson (2008). The regression is: $y_t = \mu_0 + \mu_1 t + \alpha y_{t-1} + \sum_{j=1}^{p} \gamma_j \Delta y_{t-j} + \varepsilon_t$ where α is the largest root. The row with the country name reports the full sample estimates, while the rows with the sub-samples report the sub-sample estimates. The number of lags is the same as in Table 1 (MAIC). The median unbiased estimates of γ in the full sample are: Global real rate: Italy: 0.1888; -0.0760; -0.0388. France: 0.0706; -0.1152; -0.0604. Spain: -0.0011; -0.2139. Japan: 0.3442; -0.0458.

.7 ADF-GLS Regression: Parameter Estimates

Table A.13: Parameter Estimates of the ADF-GLS Regressions in Table 1

	$\widehat{\delta}_0$	$\widehat{\delta}_1$	$\hat{\beta}+1$	$\hat{\gamma}$	ADFGLS test statistic
Global GW	0.0741	-.000075425	0.8555	0.1239; -0.2111; -0.0893	-6.1661
Global AW	0.1054	-.00013566	0.8346	0.1120; -0.1879; -0.0382	-6.7092
Italy	-0.0076	.000083876	0.8089	0.0644; -0.1782; -0.1186	-6.8925
UK	0.2396	-.00039166	0.8077	-0.1541; -0.1677	-6.8857
Dutch	0.0632	-.000070705	0.6570	-0.0431; -0.1252; -0.1220	-8.2102
France	0.0983	-.00013652	0.8381	0.0557; -0.1282; -0.0709	-6.6084
Germany	0.0946	-.00012255	0.7275	0.3223; -0.1633	-9.8656
Spain	0.1919	-.00028889	0.8730	-0.0229; -0.2331	-5.7428
US	-0.0068	.00016436	0.8804	0.1546; 0.0607; -0.1127	-3.6151
Japan	0.0599	-.00053571	0.8250	0.3439; -0.0442	-3.9586

Note: The table reports the parameter estimates in the ADF-GLS regressions in Table 1. The notation is as in Section 4.1: The ADF-GLS unit root test is implemented in an augmented Dickey-Fuller regression: $\Delta\widetilde{y}_t = \beta\widetilde{y}_{t-1} + \sum_{j=1}^{k}\gamma_j\Delta\widetilde{y}_{t-j} + \varepsilon_t$, where $\widetilde{y}_t$ is the GLS-detrended variable. The number of lags k is determined by the MAIC criterion and $\gamma = (\gamma_1, \gamma_2, ..., \gamma_k)$. The GLS detrended variable is obtained as follows: 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 perform the following regression; $y_t^* = \delta_0 x_t + \delta_1 d_t$ and let $\widetilde{y}_t = y_t - \hat{\delta}_0 - \hat{\delta}_1 t$.