

ECONOMICS WORKING PAPERS

# The Anatomy of Capital Returns over the Very Long Run: Germany, 1465–2025*

Paul Schmelzing†

Economics Working Paper 26128

HOOVER INSTITUTION
434 GALVEZ MALL
STANFORD UNIVERSITY
STANFORD, CA 94305-6010

November 2026

This paper for the first time reconstructs the weighted return on capital over centuries, on the basis of new archival data covering a substantial share (>70%) of the capital stock. The approach focuses on a single country for which the most nuanced decomposition can be obtained – including what arguably constitutes the most comprehensive real estate data, as well as the first-ever bond total return series for such a horizon. Across a range of data variations, the return on capital is moderately upwards sloping in both real and nominal terms (whether or not including cash components), typically fluctuated in a nominal range of 3-10%, and exhibits trend stationarity. The return on capital used to be strongly correlated with the macroeconomy – but especially since the 1800s, capital returns are gradually detaching from the business cycle – with the trend being most pronounced for sovereign assets.

JEL Codes: G12, G19, N01

The Hoover Institution Economics Working Paper Series allows authors to distribute research for discussion and comment among other researchers. Working papers reflect the views of the author and not the views of the Hoover Institution.

---

* A previous version of this paper circulated under the title "Global Housing Returns, Discount Rates, and the Emergence of the Safe Asset, 1465–2025". I thank participants at the 2025 SFS Cavalcade, the 2026 Western Finance Association (WFA) Conference, the 2026 NY Fed Monetary and Financial History Conference, the 2026 Rotman Asset Pricing and Investments Workshop, as well as seminar audiences at the Hoover Institution's Economic Policy Working Group, and Boston College Finance for comments – in addition to helpful comments from Kenneth Rogoff, John Cochrane, Andrey Malenko, Peter Koudijs (discussant), Cameron LaPoint (discussant), Allison Shertzer (discussant), Christophe Spaenjers, Barry Eichengreen, Vincent Bogousslavsky, and Paul Goldsmith-Pinkham (discussant) for comments. I thank Sébastien Pouget, William Goetzmann and David Le Bris for sharing discount rate data, Allison Shertzer for sharing U.S. data in advance, Cornel Zwierlein for sharing German housing statistics, Thomas Wenderoth for Nuremberg rental data, and Walter Bauernfeind for discussing details of the TOPO N database project, and I thank Dan Smith, Guoqi Chen and Ruiyang Wang for outstanding research assistance.

† Boston College and Hoover Institution, Stanford. Contact: schmelzi@bc.edu, schmelzing@stanford.edu.

# 1 Introduction

Understanding and measuring the aggregate return on the capital (wealth) stock has become an inquiry that now spans critical parts of macroeconomics, asset pricing, inequality and related research areas. The relationship of such aggregate returns relative to the rate of output growth, as well as its decomposition and trend have fundamentally informed stylized facts on the economic environment of the early 21st century, such as "secular stagnation", risk premia, or the proposition of rising inequality (Piketty, 2014).

Nuanced empirical reconstructions are now covering a period of up to 150 years, for leading Western economies, and include housing, bond, bill, and equity components. Based on these efforts, the leading view today posits a long-run downward bias in "safe" returns, a generally "smooth" risky real return in a 5-8% range, and no clear trend in either the risk premium or aggregate capital returns since the 19th century (Jorda et al., 2019). Similarly, Piketty (2014), mainly on the basis of British and French wealth data extending back into the late 1700s, famously concluded that "from the eighteenth century to the twenty- first, the pure return on capital has oscillated around a central value of 4– 5 percent a year, or more generally in an interval from 3– 6 percent a year. There has been no pronounced long-term trend either upward or downward."

The purpose of this paper is two-fold. First, it seeks to reconstruct the longest conceptually consistent approximation of the "aggregate return on capital" – covering over 70% of the value-weighted capital stock since the late Renaissance. Given the substantial measurement challenges such an undertaking presents, this paper focuses on one single country for which by far the most refined annual data has been obtained from primary archival sources: Germany. By basing the new data overwhelmingly on actual archival (primary) data, the contribution offers what is arguably the most comprehensive multi-century total return dataset for housing (using modern repeat sales methodologies), as well as the first-ever multi-century time series on sovereign bond *total returns*. In addition, this paper also presents an extension that incorporates weighted cash returns.

While it is conjecturally known that real estate (including land) and sovereign debt historically represented two of the largest investable components of capital stocks, neither their relative importance, nor the dynamic evolution of their shares in the capital stock are known in any nuanced detail. Another contribution of this paper is to fundamentally improve on this situation and present such dynamic capital stock shares – which are the prerequisite for a *weighted* reconstruction of the aggregate return on capital, based again on actual historical household portfolio compositions.

Importantly, existing research has arguably been biased towards positing "trendless" patterns in financial returns. In this context, apart from suggesting revisions of key stylized empirical facts, a multi-century empirical basis offers another key advantage. Specifically, even recent ambitious 100-year data sets arguably suffer from limited econometric and statistical power. As Rogoff et al. (2024) showed recently for the example of government bonds, data over such a horizon remains prone to mis-identifications in key econometric attributes and long-run trends – including the identification of "inflection" points and unit root properties. Multi-century data, on the other

hand, almost fully eliminates such inadequacies.[1] This paper argues that conceptually consistent multi century data – even if limited to a single country – can offer far superior statistical and econometric results. In particular, this concerns the identification of stationarity, structural break, and mean-reversion elements.

As such, the German environment offers a new panoramic analysis of return trends. Key among the findings is the identification of a moderate upwards-sloping trend of aggregate capital returns since the Renaissance, across most plausibly definitions of capital returns – contrary to existing propositions of a "stagnant" or "range-bound" long-run behavior of returns. This upwards slope is observed in the two key capital components individually – government debt total returns, and real estate total returns – and it is mirrored not least in the identification of a multi-century trend in housing excess returns. Not least – almost as a "byproduct" of the new data – it becomes apparent that the widespread identification of a "low rate environment" in the early 21st century is misleading: government bond total returns are *rising* over time. Across asset classes, capital *yields* are falling over time, while capital *gains* are rising.

While not fully universal across variations, these stylized facts are robust to a range of plausible modifications in the return assumptions, cost components, and inflation adjustments. And importantly, the trends in aggregate returns, excess returns, and its components, can now be formally confirmed through leveraging the full statistical power of multi-century data sets: robust findings of trend stationarity, and structural breaks do formally reject the prevailing notions about "trendless" properties in the returns on wealth. Together, these results suggest the existence of long-horizon trends in key economic variables, and they qualify stylized narratives regarding risk, returns, and 20th century inflection points in the macroeconomy. While idiosyncratic events impact German data – including the Thirty Years' War – the closely comparable asset composition across advanced economies and related evidence make the German findings plausibly representative for broader Western market trends.

The final section analyzes the relationship of capital returns and the economy over time. Exploiting recent advances in long-run consumption data for Germany, for the first time, we can reconstruct the correlation of capital returns (and its components) with consumption growth in annual frequencies over centuries. This allows not just the proposition of a new key stylized empirical fact: capital returns used to be closely correlated with the wider economy (consumption) in the early modern economy – but over time, capital returns and the real economy "divorce" from each other, a trend that accelerates in particular from the 19th century. This section also suggests key revisions to the long history of "safe" and "risky" asset features. While strongly positively correlated with the economy well into the 18th century, capital returns became less and less correlated with consumption growth measures ever since. In addition, for extended periods of time over the past half-millennium, real estate returns were evidently *less* correlated with consumption risk,

[1] According to the Monte Carlo simulations in Rogoff et al. (2024), a researcher would correctly reject nonstationarity only 49 percent of the times in a sample of 75 annual interest rate observations and only 30 percent of the times in a sample of 50 annual observations, while he/she would reject it with probability almost equal to 1 in their long-horizon sample of roughly 700 observations.

relative to government bond returns. Only in the first half of the 19th century did a 200-year period come to a close during which government bonds were "riskier" than real estate based on consumption risk: since then, there has been a secular reversal in which government bonds are showing ever weaker correlations with consumption growth. This makes the observation of secularly rising excess returns plausible: government are becoming ever "safer" relative to other parts of the economy.

Regarding the paper outline – after discussing literature, and briefly presenting new capital stock decompositions over the long run, Part 4 will be concerned with the first key capital stock component, presenting what arguably constitutes the most comprehensive long-horizon real house price index, based entirely on German primary sources and replicating modern construction methods (repeat-sales): nuanced variations of this index are also presented, including property size and urbanization adjusted series. Afterwards in Section 5, I will proceed to the second key component, sovereign bonds – followed by a brief analysis of cash as a capital component over time (Section 6). From primary source bond *yields*, I will propose a new method to derive bond *total returns* over the very long run, resulting in the very first existing bond total return series for such a horizon in existing literature. Afterwards, I will fuse components into "capital return" series and present formal econometric analyses, and conduct tests for structural breaks, stationarity, and adjustment speeds in capital returns over the long-run – suggesting that capital returns are strongly trend stationary and exhibit economically significant adjustment speeds. Section 9 turns to the correlation of capital returns with the macroeconomy: I link the new asset return data to new long-horizon German consumption growth data, and reconstruct the relationship between capital returns and the economy on the one hand, and the secular history of both relative "risk" in the two key asset classes on the other, suggesting new stylized facts. Part 10 concludes.

## 2 Literature

The analysis of long-run capital returns, and its components, can in recent times be traced back to the seminal work of Piketty (2014). In this and preceding contributions, Piketty proposes "stagnant" long-run returns on wealth prior to the 20th century – based on British and French data post-1750, and infrequent nominal interest rate data in Homer and Sylla (2005).

From a macro-finance perspective, Jorda et al. (2019) comprehensively reconstructed capital returns from a bottom-up composition of capital stock components, covering advanced economies since 1870 and including systematic housing returns. Here the authors break much empirical ground, but they echo other studies in finding "no visible" trend in risky returns or risk premia, coupled with a clear downward trend in government bond returns.

Importantly, there is no consensus on the long-run trajectory of excess returns and the (ex post) risk premium.[2] An older literature – predating debates on "secular stagnation" and "returns on

[2]Only a subset of the relevant literature discussed here distinguishes sharply between ex ante risk premia – that is, operates with *expected* rates of return – and ex post realized risk premia. Siegel (1992) without further discussion

capital" – investigated trends in excess returns (risk premia) mainly from a portfolio perspective: for instance, presenting U.S. data spanning 1806-1990, Siegel (1992) suggested that the phenomenon of a moderately rising premium was not driven by positive trends in risky assets (risky returns were "constant"), likely only applied to the U.S., and was not related to greater economic volatility or slower growth – but rather more likely due to rising risk aversion.[3] We recall that in Jorda et al. (2019), "no visible trend" was proposed for (ex post) global risk premia since 1870. Recently, Wachter et al. (2026) have echoed this view: by trying to jointly explain the decline in sovereign interest rates and "the stability in the equity premium" over the past century, the authors conclude that secular movements in interest rates are predominantly due to falling sovereign default risk – rather than falling "true real rates".

On the opposing end, Farhi and Guorio (2018) suggested a rise in the risk premium over past decades – "driven mostly by a confluence of two factors: rising market power and rising macroeconomic risk". Echoing this more narrowly for equities, Kuvshinov and Zimmermann (2020) find that the "equity risk premium has increased materially since 1990". Reis (2022) stresses that measures of private aggregate capital returns in the U.S. and broader advanced economies are either stagnant or moderately upwards trending since the 1980s, as opposed to the downward trend measured for government interest rates. His baseline estimates include real estate but exclude the government sector, though variations in these choices do not meaningfully change results, and financial frictions are emphasized as an explanatory factor. Marx et al. (2021) also observe elevated returns to aggregate capital in Western markets relative to risk-free interest rates, resulting in a rising risk premium: the authors contend that improved hedging properties of government bonds, as well as rising risk aversion can plausibly account for the phenomenon. This recent literature often derives private returns to capital from national product accounts, at times including dwellings and/or intangibles.

An extensive literature focuses separately on the individual performance of the two key components of capital returns, real estate and sovereign bonds. Real estate returns are the most important component in Jorda et al. (2019), and the underlying empirical basis for the time series is Knoll et al. (2017): the latter contribution reconstructed house price series for 14 advanced economies over the years 1870-2012, concentrating on urban housing total returns, which are adjusted with fixed estimates for expenditures to yield net returns. The authors find that real house prices for residential properties in advanced economies were stagnant in the pre-1945 era, and afterwards began a long-run acceleration from the second half of the 20th century. Specifically, the authors conclude that:

---

reports ex post (realized) excess returns as the "equity risk premium". Jorda et al. (2019) proceed along the same lines. Note that relevant contributions that do operate more formally with *expected* returns mainly derive expected returns from past realized ranges or lagged values. As emphasized throughout, I follow the notations in Siegel (1992) and Jorda et al. (2019) in operating strictly with *ex post* risk premia and excess returns.

[3] "The 1929-32 stock market crash and the Depression left a legacy of fear; most investors clung to government securities and insured deposits, driving their yields down. Redistribution policies undertaken by the government subsequent to the Depression may also have lowered real rates by shifting wealth to more risk-averse segments of the population", see Siegel (1992).

- "From the last quarter of the nineteenth to the mid-twentieth century, house prices in most industrial economies were largely constant in real (CPI-deflated) terms. By the 1960s they were, on average, not much higher than they were on the eve of World War I. They have been on a long and pronounced ascent since then, giving rise to a hockey-stick pattern of house prices in the long run."

To underpin this proposition, Knoll et al. (2017) undertake a widely used structural break test for the German and 13 other real house price indices and identify the years around 1964 as the key inflection point from which house prices began their secular "hockey stick" acceleration.[4] The findings on historical real estate returns were built on – and repeatedly challenged – in subsequent work: nuanced measurement improvements, including data on 20th century agricultural returns in Chambers et al. (2021), U.S. newspaper data in Lyons et al. (2024), and urban data for Amsterdam and Paris over 1700-1914 in Eichholtz et al. (2021) all qualified propositions on housing's performance over recent decades, which we can built on in this present study. However, this literature has not dealt with aggregate capital returns, nor spanned multi-century horizons.

For bonds, one classic long-horizon reference remains Homer and Sylla (2005) – on whom not least Piketty (2014) relies on for his proposition of "stagnant" capital returns. A widely used contribution, Homer and Sylla (2005) contains key limitations – including its conjectural empirics for pre-1800 data, its omission of any high-frequency (annual) data series, and the absence of primary source (archival) data basis. The new work of Schmelzing (2026) reconstructs annual-level real and nominal sovereign interest rates as far back as the 14th century using comprehensive archival material – on whose nominal yield data points I can thus draw here: the book also features reconstructions of capital stock decomposition for advanced economies (utilized here, too), but does not attempt to construct total return series for bonds, either.

Meanwhile, for the extension to bond total returns in Section 6, I propose a new methodological approach that *combines* primary historical data with machine learning techniques. Machine learning advances have overwhelmingly been used to project *forward* price and return variables in finance and economics literature, for instance Gu et al. (2020). One methodological innovation of this paper is to use machine learning to "backcast" financial variables, and to cross-validate any general results via new primary (archival) sources. Most recently in economics, a new interdisciplinary study exists in Koch et al. (2024), who attempt to reconstruct per capita GDP and improve on the classic Maddison income data set. A currently unpublished study discussing a similar approach for historical England/U.K. wage reconstructions exists in Paker et al. (2025).

With regards to the final section that investigates the long-horizon relationship between capital returns and the economy, attention has predominantly focused on the "R-G" spread – that is, the difference between capital returns and the (aggregate) GDP growth rate (Piketty, 2014; Jorda et al., 2019; Rogoff and Schmelzing, 2024). While new facts could be gathered here, this particular measure is not our main focus. Rather, canonical literature emphasizes the procyclicality of housing

[4] Eight out of the 14 advanced economy real house price indices record the key structural break between 1952-1964 on their basis; the authors use a standard Bai and Perron (2003) break test, allowing for up to three breaks.

with consumption as key factor for housing risk premia (Piazzesi and Schneider, 2016). But we know very little – if anything – about the long-run evolution of real estate procyclicality, including whether any clear trends exist. In this context, the proposition of improving hedging properties for bonds as a driver for risk premia via Marx et al. (2021) is also worth reiterating. We also note that Gomme et al. (2011) analyzed the relationship between business capital returns and equity returns, finding much greater volatility in the latter: but the focus is on U.S. cylical variation and on much shorter horizons (1954-2008).

When it comes to the long evolution of *relative* asset safety, no empirical reconstruction exists to the knowledge of the author. An extensive micro literature has dealt with asset risk, and compared risk attributes in particular historical settings (McFarlane, 1957; Flandreau et al., 2009); what this existing economic history literature does posit is that real estate (including land) constituted an early "safe asset" prior to the rise of safe governments, with government assets gradually assuming full "safety status" departing from the British environment in the early 1700s (Gorton, 2017). It is also relatively firmly established that portfolio optimization does take place in "typical" household settings – that is, risk attributes are opportunistically evaluated and investment decisions based on the weighting of asset level risk factors: we can generally assume that over the observation horizon in this study, households and investors did have a clear notion about asset risk, and sought to "hedge" risk factors (de Moor and Zuijderduijn, 2013; Korevaar, 2021).[5]

## 3 Capital stock decomposition: a portfolio approach

Earlier work to estimate wealth and wealth returns typically rely on (indirect) balance sheet approaches. A classic resource is Goldsmith (1985). The contribution covers 20 advanced economies (including Japan and Russia), and presents point estimates of broad capital stock decompositions from secondary sources. The work has underpinned stylized views on national wealth – but comes with important drawbacks: the earliest point observation is for the U.K. as of 1688 (followed by a subsequent estimate for 1760). All other countries are covered from the late 19th century, with Germany's first observation reported for 1850, followed by one in 1875. A balance sheet approach is also followed in the work of Piketty (2014) and underlying work. Also following a balance sheet approach, more granular estimates were recently constructed for the U.K. by Broadberry and de Pleijt (2021): their data spans investable and non-investable assets for Britain (England) over 1270-1870 and reports decompositions in 50-year frequencies. The data is benchmarking estimates to the well-known Feinstein estimates for 1760/1860.

In more recent work, it has been argued convincingly that portfolio approaches to wealth returns are much cleaner to the earlier indirect balance sheet approaches.[6] In Schmelzing (2026), historical

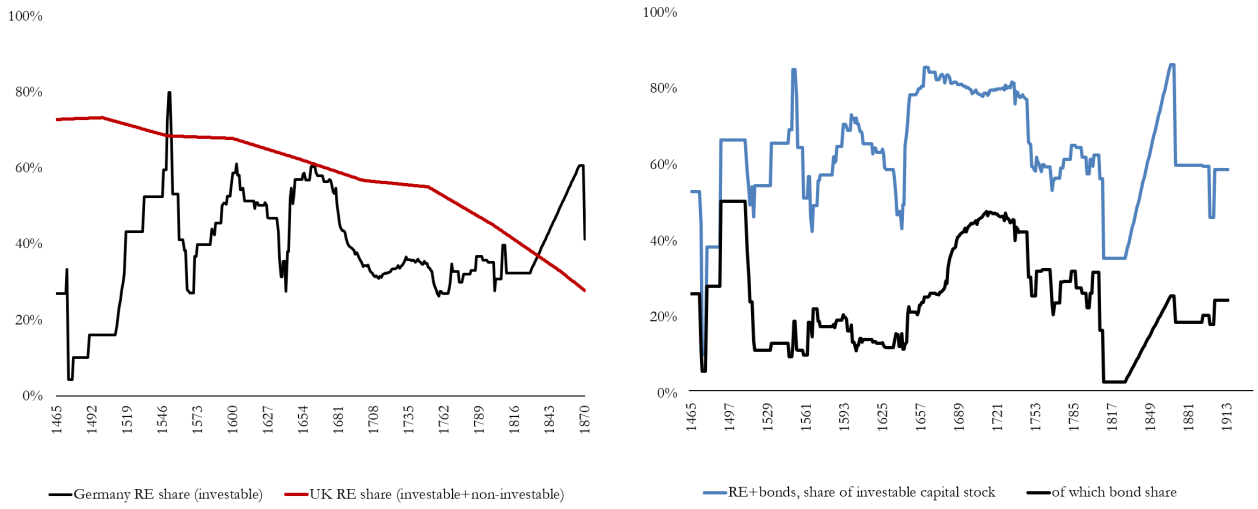
[5] In particular, based on investment behavior of Dutch households over the 15th century, de Moor and Zuijderduijn (2013) show that during economic or natural disasters, households systematically increased exposure to annuities and land – at the expense of "risky" assets such as livestock and corporate assets.

[6] E.g. Jorda et al. (2019) argue that a portfolio approach is superior to the previous indirect balance sheet approaches, by relying on actual market prices. Their portfolio approach spans real estate, equity, and sovereign bonds.

capital stock decompositions across six asset classes are reconstructed on the basis of actual primary sources (wills and household portfolios), for 288 leading investor households across European economies, in addition to aggregate wealth inventory data in secondary literature: the figures thus more closely represent actual investable capital stock shares, and also need not work backwards from modern benchmark estimates. The period covered is 1344-1914. Since they record cash as a separate asset category, the non-cash investable asset shares can equally be backed out.[7]

The left panel in Figure 1 displays the capital stock estimates in Schmelzing (2026) alongside the recent estimates for England/U.K. in Broadberry and de Pleijt (2021) – focusing here on the real estate and sovereign bond components. We see that if anything, the estimates for German real estate might be slightly conservative compared to the top-down estimates for the U.K. The right panel displays the actual share of the investable capital stock covered by the German asset return data over time: including cash, the share stands at an average of 77.6% of investable assets, and 72.3% of non-cash investable assets.

Figure 1: Capital stock decompositions, Germany, 1465-1870 (1914).




*Notes:* The two panels depict decompositions of capital stocks (national wealth). The left panel compares the primary source estimates of real estate in (investable) national wealth via Schmelzing (2026) with the recent U.K. (England) estimates for real estate (land+dwellings) in Broadberry and de Pleijt (2021). The right panel displays the relative importance for real estate and bonds using the capital stock decomposition of Schmelzing (2026), here as shares of the investable capital stock (including cash).

Excluded from the capital return aggregation are several components that are significant in today's capital stocks. In all cases, we have estimates of their relative size in the capital stock – but the limitation is sufficient availability of (total) return data. Among the most relevant are:

[7]See appendix for more detail on the archival data bases: the leading households covered include prominent merchants, political executives, and aggregated urban data points: archival data was only included if sufficient detail was available to distinguish the six asset classes. All values are current values. Among the published aggregate data points included are the wealth surveys in Leiden (Prak, 1985), Aalst (Ryckbosch, 2012), Überlingen (Schaefer, 1893), and Göppingen (Frey, 2000).

- Corporate returns. Prior to the establishment of institutionalized stock exchanges in Amsterdam, London, and elsewhere, data on corporate performance and returns come mainly from private corporations. Numerous profit data points are known for the early European (merchant) banks. The return evidence, however, remains too scattered to allow for a systematic incorporation – especially given the absence of evidence for the non-banking sectors. We should emphasize, however, that the important agricultural sector is at least indirectly incorporated through our coverage of (rural) land prices and total returns. On the basis of the historical portfolio wealth decompositions, the share of corporate assets stands at 10.9% of investable assets (incl. cash).

- Private fixed income. In Schmelzing (2026), mortgage rates for Germany are reconstructed on a high-frequency (annual) basis. Even though a reasonable argument can be made to derive absolute performance on the basis of yields, I abstained from calculating total returns here. On the basis of the historical portfolio wealth decompositions, the average share of private fixed income stands at 7.4% of investable assets (incl. cash).

# 4 Component 1: constructing a primary source repeat-sales house price index (RHPI), and total returns, 1465-2025

An important limitation of existing historical housing return data is that it is near-exclusively measuring urban return data: in an environment where urbanization rates are below 30% prior to 1900, that approach could meaningfully distort indices that purport to measure "national" housing total returns. In this paper, I combine urban and rural data, and adjust aggregates by urbanization trends. On the urban side, my basis for the construction of the new German housing price and total return indices are city level individual property records, known in German as *Häuserbücher* or *Hauschronik*. These are comprehensive compilations of cadaster and other official municipal records that allow a full chronological reconstruction of property level ownership, transactions, repeat sales prices, personal information on the occupants, and (occasionally) rental details. In effect, the data allows a construction of an index methodologically mirroring the widely used Case-Shiller repeat sales index for the U.S., which is still regarded as the most robust long-run approach controlling for heterogeneous housing characteristics.

Major progress has been made in recent years in the digitization of methodologically sufficient housing records over very long periods of time. German archives have spearheaded some of these initiatives, which – combined with new multi-century granular data for related asset classes – makes this country especially suited for new efforts at a holistic and robust general asset market reconstruction. Figure 2 displays an example, an entry from the new "Topo N" project for the city of Nuremberg (one of the key early modern financial hubs of Europe), which as of 2025 has systematically linked close to 4,000 property level records via unit-identifiers, and merged this cadaster information with ownership histories, sales prices, and other notable events related to the

property.[8] Such advancements mirror those in other cities and allow repeat-sales observations for thousands of properties over centuries for consistent individual housing units – for more than a dozen key German cities from the late medieval period.

These new sources have several advantages: perhaps most importantly, their level of detail on the property level allow a full repeat-sales index construction, in line with benchmark contemporary indices: such a methodology is considered the "gold standard" among various alternative approaches, since it allows a better control for quality changes in housing, a distinct concern for real estate as an asset class in existing literature. Second, as these records merge official with notary information, they contain *both* tax assessment values and market values of the property, with the latter being much preferred as a measure of actual price dynamics, though much literature (including the Case-Shiller index) has had to rely on assessment or self-reported values for earlier periods thus far, which contain problematic biases.[9] Third, the records are unusually consistent and complete: the average length for individual properties included in the index that is continuously documented stands at 201.5 years, with a maximum of 407 years, and more than one-third with over 300 years of continuous records. Typically, property related information is quite abundant in this source type, with a significant number of property level records reporting details such as building year of the property, underlying land values, incidences of significant material damages, and quality changes (significant renovations, additions, or mergers with other properties).

In the construction of the index, I focus solely on actual market transaction prices – ignoring tax assessment values of properties, which are often given, as these are noisy given variation in property tax rates, irregular assessment, and underreporting of actual property values by owners.

For the headline index, and all headline return figures, I value-weight observations across all properties for the headline indices, which rebalances index weights based on property values every single year (using interpolations for years where individual properties are not on the market). In the Appendix, I also show alternatively an arithmetically weighted version, which suggests higher (and upwards sloping) real house price growth prior to the Thirty Years War, but otherwise is closely comparable regarding secular trends, econometric properties, and inflection points.[10]

### 4.1 Summary Statistics

On the urban side, the German residential house sample covers twelve cities, spanning a heterogenous geographical and demographic range in the country over time. Four cities are in Northern Germany, eight in Southern Germany, and the size of cities ranges from 1,300 (Stengberg, as of 1500) to 538,000 (Munich, as of 1905). All the key financial hubs of Germany/the Holy Roman Empire over the early modern and modern period are part of the sample, including Nuremberg,

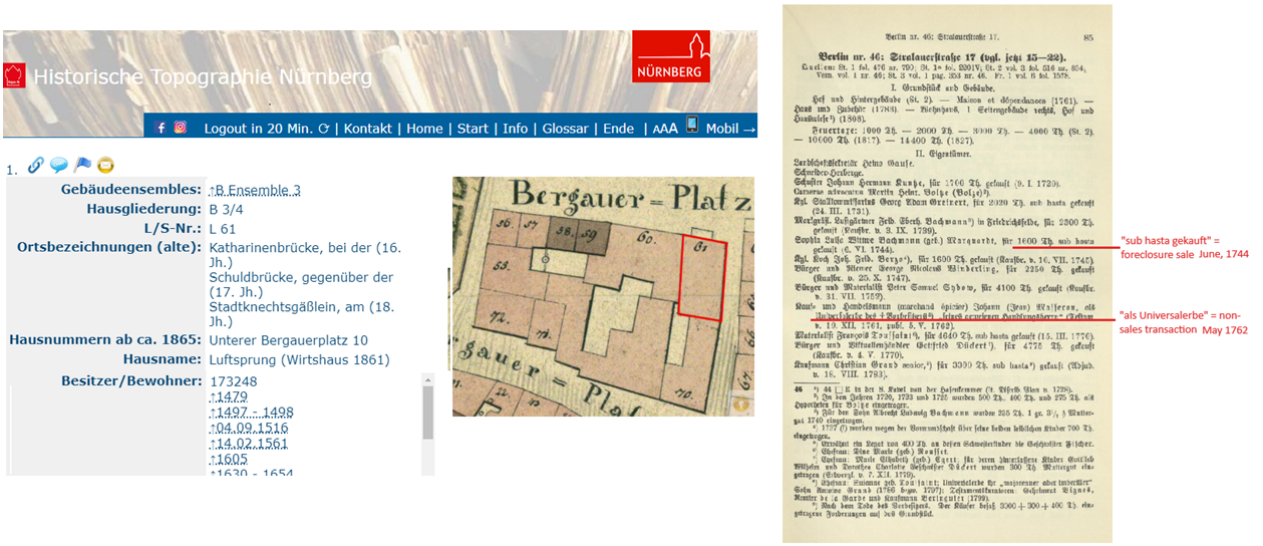
[8] See methodological details in Razum et al. (2023).

[9] Not least, non-sales transactions (inheritances, transfers, or property swaps). The Case-Shiller index for U.S. pre-depression home values uses self-reported survey data points via – these drawbacks have been critiqued in White (2009) and elsewhere.

[10] See further discussion and data in Appendix section 3 and Figure A.3.

Figure 2: Primary source exhibits I: Nuremberg TOPO N project, *Unterer Bergauerplatz 10*; and Berlin "Häuserbuch", *Stalauerstrasse 17*.

**[1]** **[2]**




*Notes:* The picture displays example entries from the primary sources I use to construct the new German house price index. On the left handside ([1]), a sample entry from the Nuremberg *Häuserbuch* ("Topo N" Project), for the address of *Unterer Bergauerplatz No. 10*. Consistent cadaster units are traced here in their ownership history, sales transactions, sales prices, and other notable events associated with the house. In this case, the first ownership entry (category *Besitzer / Bewohner*) is recorded in the year 1479. Entry via: https://online-service2.nuernberg.de/stadtarchiv/zeig.FAU?sid=574BA1843dm=6ind=2ipos=4220332f34. On the right handside ([2]), a sample page from the Berlin *Häuserbuch* that covers property level transaction histories spanning the 16th-20th centuries, published via Lüdicke (1933). Here the property recorded is *Stalauerstrasse 17*. To give a sense of the information covered (beyond traditional cash sales), the red callouts highlight a foreclosure sale in June 1744 ("sub hasta"), and below details on the property being passed on as inheritance in May 1762.

Frankfurt, Munich, and Berlin. While such ex post successful cities may feature some degree of "superstar city" bias, the inclusion of smaller and relatively stagnant towns – in addition to the full rural sample of estates and villages – balances the aggregate sample regarding these possible factors.[11]

In total, I rely on more than 7,800 early modern repeat sales at this point. The underlying urban properties are exclusively single or multi-family residential objects, and the rural estates and castles are predominantly held by old aristocratic families, who transact among each other. A significant share of the urban properties in the sample continues to exist as of 2026, based on a cursory survey – even though house numbers and other designations have repeatedly changed.[12] The houses in

[11] It has been argued that "Superstar cities" may exhibit faster house price growth, biasing indices upwards; ex post, my city sample includes both "successful" and "unsuccessful" cities with very stable or stagnating population growth (e.g. Stengberg).

[12] House numbers in any case are an innovation of the 18th century. In the pre-1700 era of our sample, the properties are typically designated by their owner, or by lengthy descriptions of their location relative to prominent buildings in the city (the church, city hall, or wells).

our sample have a median value of 594 Rheinflorin (Rfl., or *Rheingulden*) in 1465, a figure that has risen to 2,025 Rfl by the year 1650, 4,537 Rfl. by the year 1750, followed by 18,131 Rfl. by the year 1850, and finally 33,753 Rfl. at the eve of World War One (1910). The small town of Stengberg records the lowest median house prices over the sample period at 224 Rfl., while Berlin records the highest median house price, at 14,392 Rfl.

The 1,012 German properties between them account for more than 185,000 property-years, making ours arguably the most comprehensive sample in the real estate literature. The sample spans repeat sales observations from the late 15th century, all the way to year 1943, with an average continuous transaction history for each property of 201 years. Of course, even a sample of over one-thousand properties can still be considered "small" by the standards of the German housing stock, even in the early modern period. After all, more than 4,000 individual housing units existed in the early modern city of Nuremberg alone. Yet, our sample is designed to be representative of the German residential housing stock given its geographic composition – takes into account a separate sample of *non-urban* repeat sales (estates, castles, and village sales, via Appendix section 3.2, which incorporates the land component of housing – and also tries to adjust with new and nuanced data for changes in living space.

Tables 1.1-1.2 summarize the new German primary data, by period and across variables.

### 4.2 A new repeat-sales Real House Price Index (RHPI)

Existing repeat sale indices typically use CPI data to adjust nominal house price series to real bases. Shiller (2015, chapter 3), for one, uses year-to-year U.S. CPI index changes to deflate respective nominal home price changes and create a U.S. "real home price index" over 1890-2014. Similarly, Knoll et al. (2017) deflate nominal house prices with a 5-year lagged realized CPI (that is, the average over t-5 to t-1, to deflate nominal price at t=0), to obtain real house price indices. I broadly follow the latter approach, using lagged inflation, but also test in the Appendix (section 4) alternative inflation data and adjustments.

Such repeat-sales indices are regarded as "state of the art" indices, and they underlie benchmark U.S. and non-U.S. indices, though alternative "hedonic indices" can be preferable if a sufficiently large number of heterogeneous property characteristics are available (Lyons et al., 2024). Yet, while they can capture many of the quality-related changes in housing as a wider asset class – for instance, they are by design well equipped to deal with the land component of the home price, as they keep it constant at the property level – even repeat-sales indices are of course far from *perfectly* stripping out pure quality-related improvements. While acknowledging quality improvements from infrastructure, sanitation, access to hospitals and schools, as well as building quality improvements are relevant, but difficult to account for, two key quality adjustments over the long horizon can however be accounted for relatively cleanly, and are presented in variations of the German RHPI index: adjustments for urbanization (relocation from rural to urban), and property size trends. For instance, the long-run suggests that the median German homeowner only possesses a 30-50% larger home in terms of living space, while I also go beyond the near-exclusive

existing focus on *urban* housing returns and collect rural repeat-sales that can serve to incorporate urbanization trends.[13]

Figure 3 now displays three new benchmark long-run German house price index (RHPI), here in real terms (log), and value-weighted, over the full period 1465-2025 (1465=100). The index displays moderately falling real house prices until the early 1500s, followed by a sustained rise in prices, until a sharp visible inflection during the Thirty Years War, when structural inflation rates surge, a significant demographic shock takes place, and much existing housing stock is physically destroyed.[14] It takes until the early 1800s for this major shock to be fully reversed. Regardless of the specific inflation basis chosen, a key commonality of the various iterations of indices reconstructed is the sharp acceleration of the index values from the late 18th century. Afterwards, once more it appears that geopolitical shocks (World War One and World War Two) leave among the most notable imprints in long-horizon data, with sharp declines during 1914-8 and 1939-45 in the real price data (though showing smaller peak-to-trough declines than during 1618-1468).

Overall, this reconstruction suggests that real house prices – at least in one key geography where such returns can now be measured granularly over time – did *not* begin to accelerate only from the mid-20th century, nor do the 1980s appear a particularly obvious inflection point, as key literature proposes (Shiller, 2015; Knoll et al., 2017). While I test such propositions more formally in Section 7, a long-run "hockey stick" pattern is already visually difficult to maintain; at the same time, clearly housing markets of the past were subject to major shocks, with the 2007-8 housing crisis appearing almost moderate in the context of geopolitically driven crises of the past.[15] What we are observing here is that the restricted sample length at use in the literature so far has introduced biases towards rejecting "trends" – echoing the recent observation for interest rates which suffered from a similar limitation.[16]

While our benchmark approach in this section – to value-weight repeat-sales observations of the "Urban-Rural" sample and adjust with lagged realized inflation – is consistent with contemporary state of the art indices (notably the Case-Shiller index for the U.S.), key sub-periods and absolute index levels are certainly sensitive to particular sales weightings and also the specific inflation adjustment approach. Table 3.2 below details a range of plausible (real estate) adjustments and their effects for measures of aggregate returns, and the Appendix reports a number of additional

[13] See Appendix Sections 3.2 (Rural repeat-sales) and 3.3 (Living space data) for details.

[14] The financial analysis of the Thirty Years' War is still underdeveloped. A discussion of the initial 1618-23 economic shock that is showing up in our new index exists via Kindleberger (1991). A thorough attempt to measure property and demographic losses during the war exists via von Hippel (1978): his figures suggest demographic and physical capital losses exceeding either of the 20th century World Wars.

[15] The Rural-urban benchmark RHPI peaks in 1620 at 208.2, before entering its most dramatic decline ever, contingent upon the dramatic escalation of the Thirty Years' War: the index declines a full 86.5% to the bottom in 1639 (index value 29.1); the second most dramatic reversal ever in all variations of the RHPI relates to the 1914-20 episode: here the benchmark index loses 66.5% on the same basis.

[16] See Rogoff et al. (2024).

Table 1.1: Summary Statistics for German primary data, repeat sales, 1/2

| | # properties | # repeat-sales | Property-years | Med sales gain |
|---|---|---|---|---|
| 1465-1565 | 171 | 606 | 11,311 | 23.6 |
| 1565-1665 | 289 | 1,824 | 25,405 | 25.8 |
| 1665-1765 | 500 | 1,918 | 41,005 | 29.6 |
| 1765-1865 | 492 | 2,761 | 50,017 | 34.1 |
| 1865-1943 | 218 | 834 | 16,444 | 43.2 |
| Total | | | | |
| 1465-1943 | 1,012 | 7,875 | 185,174 | 33.3 |
| Of which rural estates, castles | | | | |
| 1465-1943 | 103 | 452 | 21,607 | 11.7 |

*Notes:* The table reports the sample composition of the repeat-sales observations for urban and rural residential real estate in Germany. Sub-periods do not necessarily sum to "total" summary statistics given overlapping property spans. Med sales gain measures the nominal change in price at the time of sale over the respective most recent previous sales price of the property, expressed in cumulative percentage terms (properties arithmetically weighted).

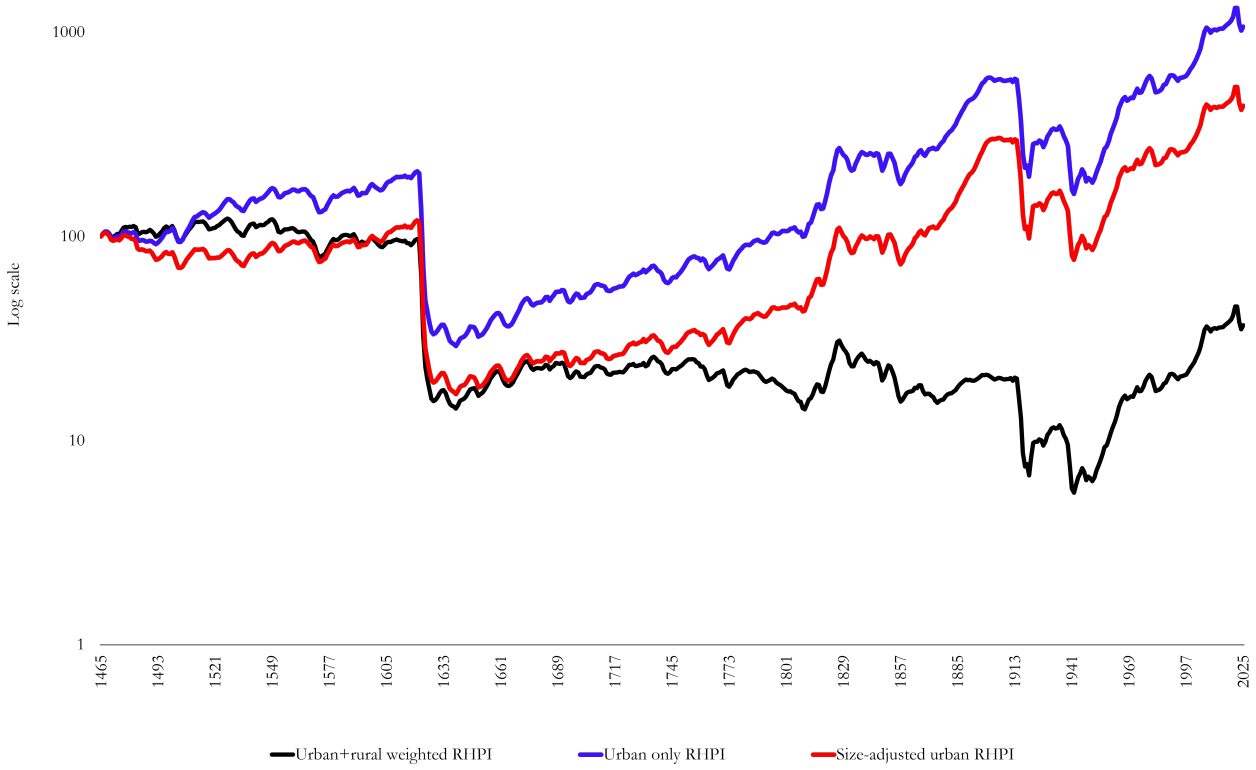
Table 1.2: Summary Statistics for German primary data, Urban-rural repeat sales, 2/2

| | gross nom TR | gross real TR | rental yd | real cap gain | hp |
|---|---|---|---|---|---|
| 1465-1565 | 7.75 | 6.60 | 6.49 | 0.11 | 23.8 |
| 1565-1665 | 4.57 | 2.33 | 3.98 | -1.65 | 33.2 |
| 1665-1765 | 4.35 | 3.92 | 3.62 | 0.29 | 42.6 |
| 1765-1865 | 8.63 | 7.91 | 7.00 | 0.90 | 29.5 |
| 1865-1943 | 7.84 | 6.10 | 6.66 | -0.55 | 22.2 |
| 1943-2025 | 10.23 | 7.49 | 5.08 | 2.42 | – |
| Total | | | | | |
| 1465-2025 | 7.14 | 5.63 | 5.44 | 0.20 | – |

*Notes:* The table reports performance measures for the repeat-sales observations for residential real estate in Germany. Measures are for the "urban-rural" variation of the data. All variables except "hp" are reported in percentage terms per annum. "hp" denotes the "holding period" (properties arithmetically weighted), measured as the distance between the observed repeat sales pairs, and is reported in years.

variations of the German benchmark indices presented in Figure 3.[17]

Figure 3: German real house price index (RHPI, log), urban / urban+rural / and size-adjusted urban basis, 1465-2025.




*Notes:* The Figure displays a German real house price index running over 1465-2025 (1465=100, log scale), value-weighting German data in three variations. Here we splice the new data introduced in this paper for 1465-1910 (using the "Schmelzing" inflation adjustment basis), with the existing German house price data in Knoll et al. (2017) over 1911-2020 (as updated in Jorda et al. (2019), and denoted "JST"), and GREX data over 2021-5. The gaps that JST record for selected German interwar years are filled by taking the arithmetic average of the real house price change for Denmark, France, Switzerland, and Belgium. The "urban+rural weighted" variation adds repeat sales from estates, castles and villages, weighted by urbanization share, as described further in Appendix Section 3.2. The "Size-adjusted RHPI" tracks real prices per square meter, using the data on representative evolving unit sizes in the sample (see Appendix Section 3.3).

## 4.3 Rental yields and owner-occupation

According to Wenderoth (2025), middle and even lower middle income households typically lived in their own single family home by the late 15th century – with around 30% of taxable citizens in German urban areas living in rented houses or apartments. This figure rises to around 60% by the

[17]For instance, multiple plausible iterations of particular weightings and inflation adjustments are to be found in the Appendix, particularly Section 3.1.

year 1800, and the development is relatively even across German cities.[18]

Overall, therefore, rental income constituted a clear investment motivation for buyers of real estate assets over the entire observation horizon, and should be an integral part of total return figures. This is equally true for rural real estate dynamics, where the extensive villages and estates usually feature households that rent their property from the owning landlord (who resides in the main estate).

As has been noted in related papers, historical rental data in primary sources is generally more elusive, hampered in part by the fact that rental contracts were privately negotiated and did not have to be registered with notaries or municipal authorities (unlike sales), in part by a convoluted nomenclature in early modern contracts.[19] All this complicates the construction of price and rental series featuring the *identical* housing properties: key recent literature has navigated this challenge in different ways, with some constructing separate housing and rental series (featuring different properties in each of the two samples) which are then merged to construct a rental yield series (e.g. Lyons et al. (2024)). However, our *Häuserbücher* contain a critical amount of detail, at least for multiple key cities, which allows a city-level linking of rental and price series – and several hundreds of additional rental-only data points have been identified in related primary and secondary sources – for instance new series in Wenderoth (2025) based on municipal rent contracts – which allow the replication of existing approaches that construct a separate rental index series that are subsequently merged with price series to derive rental yields (all rental observations are summarized in Appendix table A.2.2).

### 4.4 Costs, leverage, and net total returns

Thus far, we have not comprehensively measured the cost component of residential real estate for our long-run sample. The existing literature mainly works with general assumptions, with nuanced unit-level data being scarce. Generally, relevant costs include taxes on the municipal and state level, maintenance costs, depreciation, and costs resulting from vacancy. For Germany, the level of detail in the new sources allows dynamic, though not unit level, estimates which we will incorporate into a construction for net returns.

While there is regular evidence of leverage being involved in early modern real estate – mortgages being a financial instrument with deep and widespread historical use – I follow the vast majority of related studies and focus on non-leveraged real estate returns.[20]

[18]The Wenderoth data supersedes similar but patchier data in Dirlmeier (1978). I thank the author for sharing his data, which focuses on Southern German cities and which also suggests that generally the share of renters scales with city size in the early modern period.

[19]Housing contracts – besides mentioning the assessment or sales price – often refer to *Guelten*, or *Zins* ("interest") when denoting rental payments, with such terminology adding to confusion.

[20]In Schmelzing (2026), German mortgage rates are reconstructed, which in principle allow an incorporation of leverage costs. But apart from the data for Berlin in Lüdicke (1933) and Munich in Burgmaier and Schneider (1966), the associated data on loan-to-value ratios is not satisfactory enough. In particular the recent Jorda et al. (2019) report

### 4.4.1 Taxes

Beginning with the tax side, the information is relatively comprehensive. Some of the German *Häuserbücher* contain occasional unit-level tax information, most of which concur with our existing knowledge of aggregate city level rates.[21] For instance, among larger cities in our sample, we know that for Munich, an annual imposition of one Penny per Pound is levied from the mid-15th century, where it remains for centuries – a rate equivalent to 0.41% of the (self-assessed) value of the house. The city of Frankfurt charged 5/6 of one percent as real estate sales tax from the year 1613 onwards; the city of Augsburg charged 1/4 of one percent of the value of real estate and land in wealth taxes; the city of Leipzig charged 1 basis point of the (self-assessed) value of housing in irregular wealth tax impositions from 1481 onwards. It is to be stressed that these impositions occurred in irregular intervals – often connected to specific external threats for which cash was quickly needed and liquidity was otherwise constrained.[22]

Afterwards, property tax rates are rather stable from the 15th century, with evidence of a moderate upwards trend. For Berlin, Voigt (1901) reports tax rates of 0.3% of the property value in the 17th and 18th centuries, which rise to 1.25% by the first half of the 19th century. Afterwards, Müller (1881) provides annual rate of 1.9% of the market value for rented residential housing in Berlin during the 1870s and 1880s. With much corresponding data for other cities in our sample in broad agreement, Eberstadt (1903) provides rates for representative Hessian cities, over 1890-1901, arriving at a lower 0.6% of property value for rental properties. Against this backdrop, in the subsequent net return calculations, I operate with a constant 0.15 percentage point deduction from the gross rental yield (with more details also in Appendix section 2). Slight variations within the range of property rates provided in the literature, or the adoption of a more dynamic rate over time, do not alter any of our basic findings.[23]

### 4.4.2 Non-tax costs

With most existing literature operating with a uniform expenditure adjustment – without particular time trends – given the largely unobserved nature of precise maintenance, depreciation, and vacancy costs, what is the evidence for such non-tax costs for longer horizons?

It would be methodologically ideal of course to match property level price, vacancy, and rent data with matching property level expenditures – however, such information is sporadic. For

non-leveraged housing returns – with their argument of leverage adjustments in housing being "inconsequential" appearing questionable, however.

[21] For instance, in Nuremberg in 1500, we learn via TOPO N, B 6/5 that the house worth 190 Rfl was taxed at an annual 0.5 Rfl, that is, at a rate of 0.3%.

[22] For a systematic overview of tax rates, see Isenmann (2012). The Munich rates are via Solleder (1938, 208ff.). Further city level details in Mayr (1931) for Augsburg; Frankfurt data via Braeuer (1915).

[23] Wenderoth (2025, 91) also documents irregular direct impositions of taxes on rental income: for instance, Hamburg from the year 1627 imposed a 6.25% levy on rental income. Since these taxes are not documented more widely, I ignore them for the net yield calculations.

some cities in our sample, interesting internal cadaster calculations exist. For instance, the Berlin cadaster authorities during the 1870s apparently calculated with a 2% total gross expenditure figure relative to the current sales value of a rental residential property for total costs, including maintenance, labor, and fire insurance, but not adjusting for vacancy rates (Müller, 1881). A related official publication, also for Berlin, estimates an average vacancy rate for residential housing over the period 1880-1910 of 3% (Ascher, 1917). While providing some general reality checks, however, in line with existing literature, we will have to abstract from property and matched city level data and assume plausible aggregate figures. We will also have to abstract from some relevant city level provisions that are relevant on the margin, but cannot be generalized: for instance, Voigt (1901) reports that for Berlin in the early modern age, renovation costs are partially tax-exempt (up to 15% of total renovation costs) – lacking more specifics about the scope and duration of such provisions, we will disregard such details.

Given the less robust data situation across countries, recent long-run total return calculations thus have assumed time series based on general assumptions, rather than asset-level data. An example is Eichholtz et al. (2021), who estimate non-tax costs of 32% of rental yields as expenditure basis for Paris and Amsterdam, closely in line with the estimates in Chambers et al. (2021), who have actual property level data on expenditures for their U.K. sample. These estimates generally do not assume time trends and include maintenance, vacancy, insurance, and labor costs – as well as physical depreciation.

The Appendix (Table A.3) systematically compiles all expenditure share estimates in existing literature (and component breakdowns) that form the basis for these net return adjustments, for both urban and rural samples. Proceeding from here, I will report gross and net return estimates alongside each other, generally aligning expenditure estimates with existing assumptions in both secondary literature and specific historical sources – but applying such estimates uniformly across time and across urban and rural samples, in line with the bond return figures.[24]

### 4.5 Housing total ex post returns: nominal and real

The separate consideration of house price changes, rental yields, and expenditure components now allows the fusion into a total gross return series for German residential real estate since 1466.

We follow the total return definition in Jorda et al. (2019), specifically defining housing total return $R$ as:

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} + Y_{i,t} \tag{1}$$

with $P$ denoting the nominal price change year-on-year for housing index in country $i$, at time $t$, and $Y$ representing the rental yield. On this basis, gross housing total returns exhibit an upwards trend across all respective real estate components – ranging from 0.28 basis points per annum (urban+rural returns), to 0.52 basis points per annum (size-adjusted returns).

[24]See further Appendix section 2 and Table A.3.

Inflation has evolved dynamically in the early modern and modern periods, and therefore real total returns might represent a more accurate measure of actual asset performance. We once more follow Jorda et al. (2019), by defining housing real total returns $r$. Let $\pi$ denote the seven-year lagged CPI price change. Then $r$:

$$r_{i,t} = \frac{1 + R_{i,t}}{1 + \pi_{i,t}} - 1 \tag{2}$$

The literature at times operates with annual, at times with lagged multi-year inflation adjustments – given multi-year holding periods in residential real estate, the lagged indices appear more suitable than year-on-year inflation adjustments, and also make results more consistent with approaches for real interest rates.[25] For our German primary source series, we can use the Schmelzing (2026) inflation basis – and when that is done, this results in a real total gross return figure over the entire period of 5.9% p.a., of which 5.5% are rental yield returns, and just 0.4% p.a. result from capital gains (benchmark Urban-rural sample).[26]

## 5 Component 2: From bond yields to bond total returns

Neither for private nor sovereign debt, any long-horizon *total return* series are in existence.[27] We do have conjectural evidence that discounts or premia to the face value of the debt contract were applied in early sovereign debt markets: for instance, the series on Venetian government debt (*Monte Vecchio*) during the 14th century compiled by Gino Luzzatto.[28] But outside of these sporadic pieces of evidence, it is not evident that price variation relative to par was a general feature of sovereign debt in the early modern period, before the arrival of institutionalized bond markets. In general, the proposition of an environment of "hold-to-maturity" (HTM) bond investors (who do not re-invest proceeds into sovereign debt) is a representative one. In what follows, it should be kept in mind that "HTM" series are exclusively derived from primary sources – these do not rely on any regression or ML assumptions – and we report HTM results alongside the new total return series.

Two certain observations can be made: first, sovereign debt contracts were generally marketable and as such at least occasionally sold on from the original contracting creditor party (thus, the confirmation of secondary market transactions); second, the debt contracts served regularly as collateral in other financial transactions, where "haircuts" to the face value were typically applied.

[25] For instance, Schmelzing (2026) and Rogoff et al. (2024) use the same seven-year lagged adjustments.

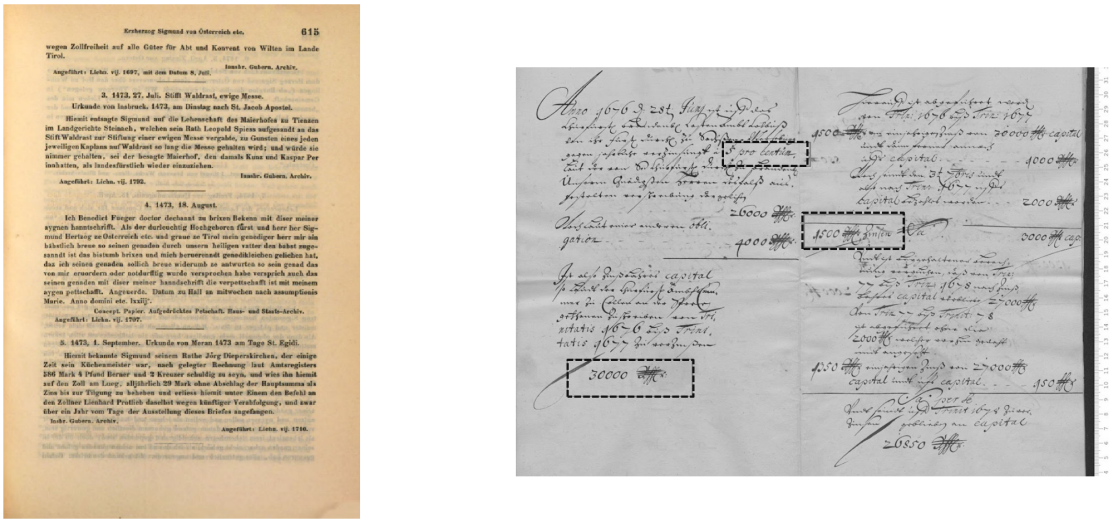
[26] Recall that an alternative German inflation series is tested via Appendix Figure A.2. Appendix Figure A.1 also displays separately the *net* total return series for Germany, spanning 1465-2025, incorporating our cost and tax estimates (as per section 4.4).

[27] Modern series include "JST" starting in 1870.

[28] According to the classic resource of Luzzatto (1929), who compiled monthly averages based on archival material, Venetian long-maturity bond prices fluctuated between 60 lire of par (in 1299) and 102 (September 1344).

In this sense, a de facto reassessment of par values took place regularly in the market context. And in this sense, an exposure to a total return index was in practice possible in the early modern period – beyond the exposure to coupon income – despite this not representing the default form of exposure. With these nuances in mind, we proceed, then, to the reconstruction of long-horizon German sovereign bond *total returns*.

Figure 4: Primary source exhibits II: German sovereign debt contracts, 1473 and 1676 transactions.



*Notes:* The exhibit displays two examples from the primary sources I use to construct the new German interest rate data series. Underlying source is Schmelzing (2026): the left hand side displays an excerpt of the contract collection of Emperor Sigismund, documenting a loan transaction with Joerg Dieperskirchen, over 586 Mark, dated September 1, 1473, via Chmel (1858); the right hand side dates from June 1676, a debt contract between Brandenburg and Domstift Meissen, via GStAPK I HA GR Rep 56 Nr. 40, Fasz 2 (5-year loan).

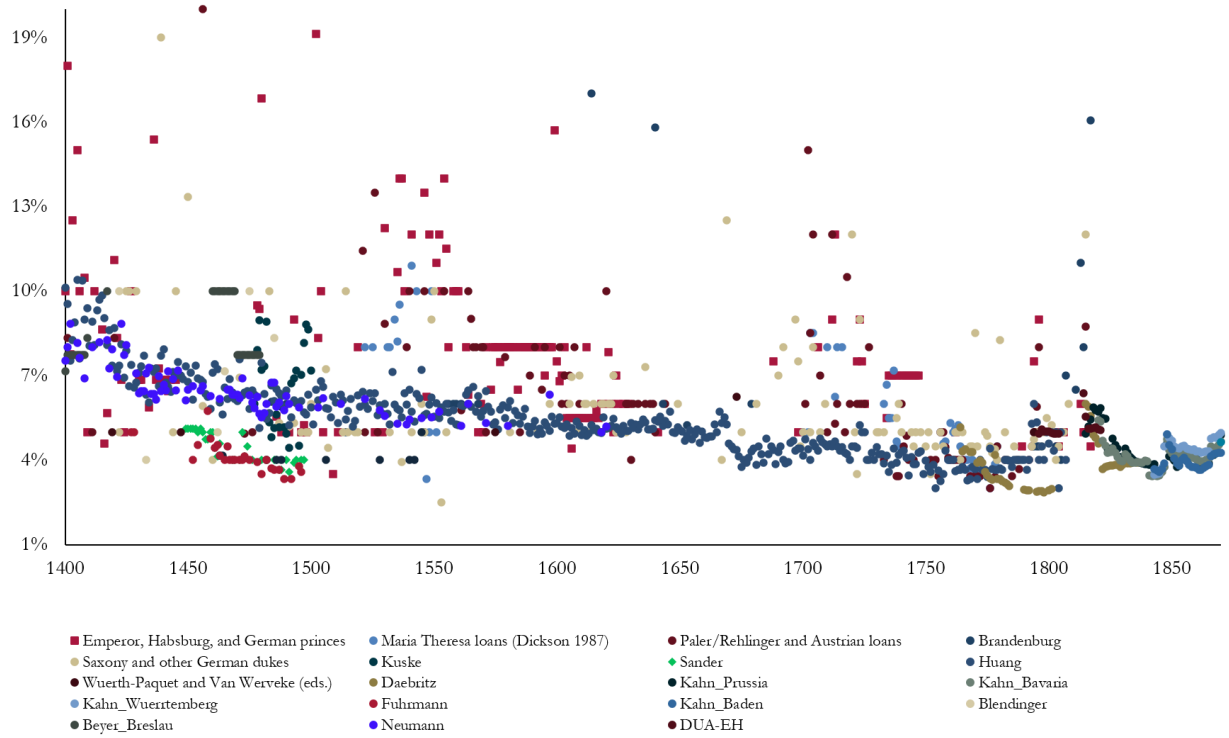
In this section, I reconstruct long-maturity government debt total returns for the first time. There are no long-run reconstructions of government debt total returns in existing literature. Nominal long-maturity sovereign yields are based on the new work of Schmelzing (2026). This work compiles multi-century interest rates for leading countries, including a total of 3,215 annual long-maturity data points underpin the German nominal sovereign yield series over 1400-1870. The debtors include the central executive (the Holy Roman Emperor), regional political executives (Dukes, Counts, constitutent states of the Holy Roman Empire), as well as independent imperial cities (*Reichsstaedte*). "Long maturity" is defined as an explicit or de facto maturity of two years or more, with perpetual contracts included in this category, and with the average maturity over time recorded at just over ten years.[29]

Figure 4 displays two sample files that go into the overall data set for Germany in the underlying Schmelzing (2026). On the left, an early data point dated September 1, 1473: we see the record of a

[29]Excluded are any transactions with non-voluntary elements (loans from court officials, from early central banks, or loans made under duress). All annual transactions are equal-weighted to form the aggregate German level.

loan transaction between the German Emperor Sigismund, and Joerg Dieperskirchen, a wealthy German citizen, over a principal of 586 Mark. An annual interest payment of 29 Mark is specified, with the tolls of Lueg specified as a security.[30] On the right hand side – based on a contract held in the Prussian Secret State Archives (GStAPK) – a 5-year contract dated June 1676, between the Elector of Brandenburg and the *Domstift* Meissen, over a principal of 30,000 Rheinflorin.[31] The interest rate is given as 1,500 Rheinflorin (=5%) per annum.

Figure 5: German nominal sovereign long-maturity yields, 1400-1870.




*Notes:* The Figure displays the total of 3,215 German nominal sovereign long-maturity yields. The source is Schmelzing (2026), who only includes transactions without apparent involuntary elements, aggregates all annual observations as equal-weights, and defines "long maturity" as equal to or above two years (contracted or de facto).

Figure 5 displays the underlying nominal yield basis, spanning the years 1400-1870 as sourced from Schmelzing (2026), by issuers. The full-sample average nominal long-maturity rate observed stands at 5.8%: as discussed in Schmelzing (2026), these long-maturity sovereign rates exhibit a downward trend over time (in both real and nominal terms). Yet, it is well conceivable that capital gains may at the same time compensate investors for lower and lower coupon rates, and therefore generate an upwards sloping total return profile. We simply do not know thus far – to date, there is not a single long-horizon reconstruction of sovereign bond *total returns* in existence.

[30]The town of Lueg was strategically located along the Brenner merchant route, and thus a valuable imperial toll checkpoint.

[31]*Domstift* is a Cathedral Chapter. "GStAPK" is the official abbreviation for the Prussian Secret State Archives, the *Geheimes Staatsarchiv Preussischer Kulturbesitz*.

### 5.0.1 Taxes

What we do know is that sovereign debt assets, like real estate discussed above, were subject to at least irregular taxation by public authorities: however, from what we can observe from historical records, when it was imposed, taxation generally took face values – not capital gains – as its taxation basis.

Similar to the situation in real estate, we have cursory – but arguably representative – information on the precise tax rates levied. For instance, via Gindeley (1868), capital taxation was fully suspended in Bohemia between 1574-92, but then activated between 1593-1617, when a rate of 0.6% on the face value (6fl for every 1,000fl) was raised.

Yet, unlike real estate, modern benchmark series for bonds are overwhelmingly pre-tax (Jorda et al., 2019): so for sovereign debt, there is no need to distinguish between "gross" and "net" debt total returns.

## 5.1 "Backcasting" financial time series

While the majority of financial applications are currently concerned with optimizing the prediction of future returns on the basis of a limited set of past realized data points, the financial historian faces a simpler problem: estimating *past realized returns (prices)* on the basis of a limited set of past realized known covariates (which are strictly already available for end of the $t$ period). In other words, we are operating in an easier environment because instead of a look-ahead forecast, we are interested in a backcast target variable reconstruction – an environment where important covariates of the target variable for the given period are already known. In this sense, the approach can also be thought of as a simple computational optimization of a nonparametric regression exercise – with the advantage that the supervised ML environment can optimize the nonparametric environment for a large set of covariates, and determine the ideal statistical structure. However, it is important to note that this means the model setup as used here does not inform about actual *forward-looking* (that is, beyond the cutoff of the training data set, this being the year 2020) predictive conditions.[32]

Importantly, while advanced machine learning models have already been prominently used for similar time series reconstructions in related fields – notably long-run climate trend and demographic reconstructions – a relevant application to finance has been fully absent, even though very recent published and unpublished studies demonstrate the potential for key economic time series applications – including advancements over the classic long-run Maddison per capita GDP series, as well as benchmark wage series.[33]

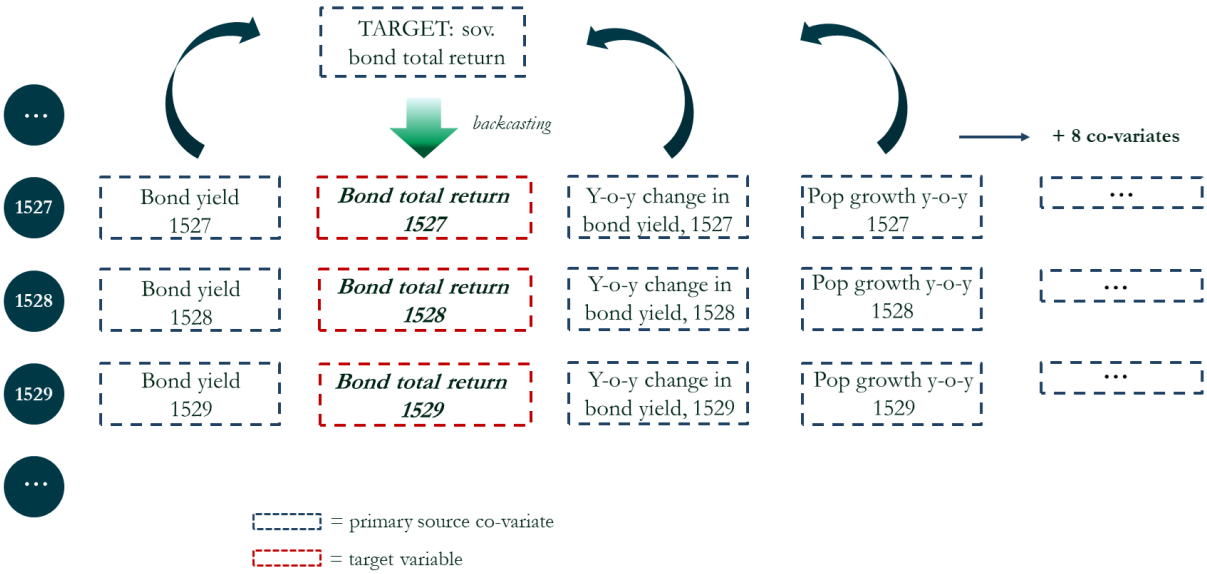
While each methodology faces its own set of challenges, machine learning approaches could offer a powerful cross-check to, and extensions of, our primary source housing reconstruction

[32]For this approach, all forward-looking covariate data would not yet be available at the time of prediction, time $t$.

[33]Here see Koch et al. (2024) for backcasting per capita GDP data, and Paker et al. (2025) for an unpublished demonstration for U.K. wage data; somewhat related are Bochow et al. (2025) and Borup et al. (2023).

approach that concentrated on Germany (as well as cross-checks to previous early modern price reconstructions in the literature, per the appendix). The situation is particularly promising given the new availability – or available reconstructions – of large sets of (thus far unexploited) known covariates for key advanced economies over centuries.

Figure 6: Backcasting approach, stylized.




*Notes:* A stylized illustration of the backcasting approach underpinning the reconstruction of German nominal bond total returns. Blue frames represent historical variable data points that are available from primary sources, and red frames represent data points that are "backcasted" on the basis of blue frames. This backcasting model is set for a context window of 0.

## 5.2 Training the model on 1870-2020 bond total return data

Specifically, I train a total of 13 nonparametric machine-learning models on both modern nominal and real housing capital gains as their target variables. A plethora of available models exists, of course all within a highly dynamic field – our focus will be on a subset of latest generation models that are either optimized for multi horizon tabular forecasting applications, or otherwise have appeared prominently in recent finance and economics literature.

Among the models I test are a Time-series Dense Encoder model (TiDE) (Das et al., 2024), a class of long-horizon forecasting models considered to be computationally more efficient than others, but more simplistic; a Seq2Seq+ model (Keneshloo et al., 2020), which utilizes deep reinforcement learning approaches; and a Temporal Fusion Transformer (TFT) model (Lim et al., 2021); both the Seq2Seq+ and TFT models incorporate Long Short-Term Memory (LSTM), a deep learning feature shown to be particularly potent at asset pricing applications such as predicting bond risk premia, for instance (Cong et al., 2021; Chen et al., 2024). Other models tested include a Random Forest (RF) model, tuned with conventional parameters – a type of ML approach that has been shown to perform impressively in forward-looking financial prediction environments (Gu et al., 2020); as

well as Gradient and XGBoost models that are emphasized for historical exercises in Paker et al. (2025).

The advantage of these models is not just their optimization for long-horizon fore- and backcasting exercises – but their strictly nonparametric foundation: all models assume no particular functional form of the data, and batch predictions resulting from such training therefore contain no particular biases on the econometric features of time series properties that we are interested in.[34]

The first step consists of the training stage, to train the machine-learning model, with annual residential house price change (in both nominal and real terms) as its target variable. I train the model on sovereign bond total returns on the basis of the most recent modern and early modern research datasets where a sufficient sample of additional covariates exists. Training and validation splits are optimized automatically, to generate the highest (out of sample) r-squared: most commonly, such optimizations use a 80:10:10 training:validation:test split.[35] The r-squared reported refers to the performance of the training data set on the validation and test data set – hence is an out-of-sample (OOS) measure. By design, the r-squared for backcasting environments will be higher than in classic forward-looking prediction environments, and ML approaches achieve elevated OOS more generally, as emphasized by Mullainathan and Spiess (2017) and others.

Figure 6 displays a stylized version of the deployment of the (fully trained) model on the basis of primary data points for the respective backcasting data point. The covariates importance across our specific ML models is then detailed via Figure 8. Specifically, the 13 covariates include long-maturity bond yields, the government deficit-ratio using Dincecco (2013), a financial crisis dummy using Metrick and Schmelzing (2026), population growth, 3-year lagged population growth, real long-maturity interest rates, inflation, 7-year lagged inflation, change in real per capita income, a FX peg dummy, the Geopolitical Risk Index (via Caldara and Iacovello (2022)), and respective construction cost indices (y-o-y change) that are sourced from national agencies.[36]

The source of the training dataset target values (annual nominal government bond total returns, 1871-2020) is the widely used "JST dataset" (Jorda et al., 2017) and sources therein.

Using this set of time series for training, I report the model training performance in Table 2, including model tuning parameters. Here we observe that out-of-sample model performance ranges from .14 to .389., with the Hist Gradient Boost marginally outperforming several peers. Mean Average Error (MAE) ranges from .042 to .059 and is broadly consistent with the relative ranking of models in r-squared metrics.

In the forecast horizon and context window columns, I record the model tuning delivering the

[34]Where such LSTM models struggle are environments with more meaningful data gaps in the training set – our environment uses the (comprehensive) JST environment where such data gaps are virtually absent (even though, as we argue for the case of Germany, the estimation procedures for given periods can be challenged).

[35]These splits are generally in line with existing forecasting literature, Bianchi et al. (2021) use a training:validation/test dataset split of 85:15. The backcasting p.c. GDP model in Koch et al. (2024) uses 80:20.

[36]For the early construction cost data, I use the new index spanning 1500-1855 by Pfister (2022).

Table 2: Machine-learning model performance. Target: German bond total return

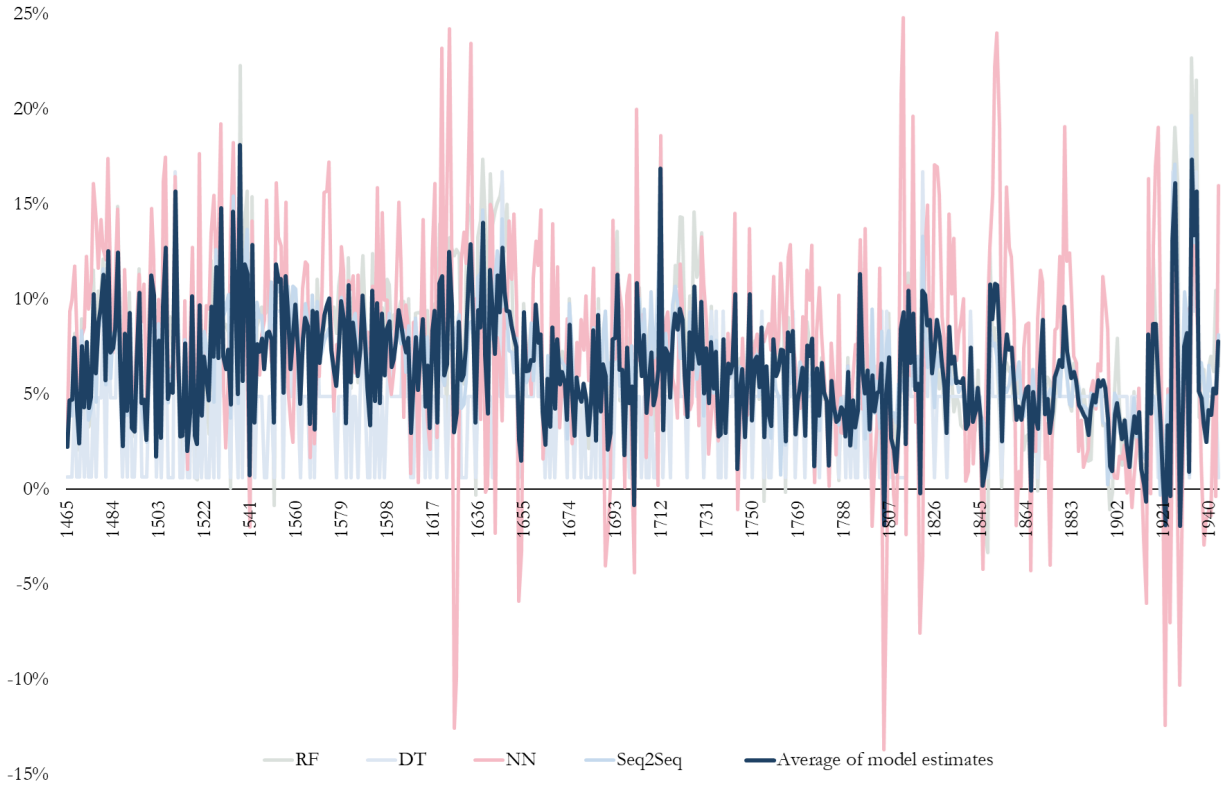
| | $r^2(OOS)$ | MAE | forecast horizon | context window |
|---|---|---|---|---|
| Target: nominal bond total return | | | | |
| Seq2Seq+ | .152 | .059 | 12 | 36 |
| Temporal Fusion Transformer (TFT) | .218 | .055 | 3 | 12 |
| Time Series Dense Encoder (TiDE) | .140 | .058 | 2 | 7 |
| Random Forest (RF) | .361 | .042 | 0 | 4 |
| Gradient Boost | .386 | .043 | 0 | 4 |
| ARIMA | -0.045 | – | – | – |
| Decision Tree (DT) | .311 | .046 | 0 | 4 |
| MLP Neural Network (NN) | .216 | .052 | 0 | 5 |
| XGBoost | .372 | .044 | 0 | 1 |
| Light GBM | .387 | .043 | 0 | 1 |
| CatBoost | .377 | .043 | 0 | 1 |
| Hist Gradient Boost | .389 | .043 | 0 | 3 |

*Notes:* The table reports training results from several machine-learning (ML) approaches, all with annual nominal (long-maturity) government bond total return as their target variable. The training data set consists of bond total returns in addition to ten further co-variant variables, annually over 1870-2020, primarily using "JST" (Jorda et al., 2017), Schmelzing (2026).

optimal r-squared performance: such inputs can be varied manually. While Seq2Seq models optimize on longer context windows and forecast horizons, we generally observe that most standard models prefer smaller parameters, optimizing on same-year forecasts (=0) and 1-4 years of context windows.

In a first, 7 now displays four variations of the resulting nominal German total sovereign bond returns, using leading machine learning models in the existing finance literature. This includes the "big four", the Seq2Seq, in addition to the prominent Random Forest (RF), Neural Network (NN), and Decision Tree (DT) models which have been widely used in forward-looking predictive exercises, including for U.S. stock market returns. I form an arithmetic average across these four to yield the dark blue line in Figure 7.

Figure 7: German sovereign bond total returns, four leading models, and average, 1465-1943.




*Notes:* The Figure displays four reconstructions of nominal German weighted total returns on capital, weighting real estate and bond total returns arithmetically. In addition, the dark blue line displays the average across the four models. RF=Random Forest, DT=Decision Tree, NN=Neural Network, Seq2Seq=Sequence to Sequence model.

## 5.3 Variable importance

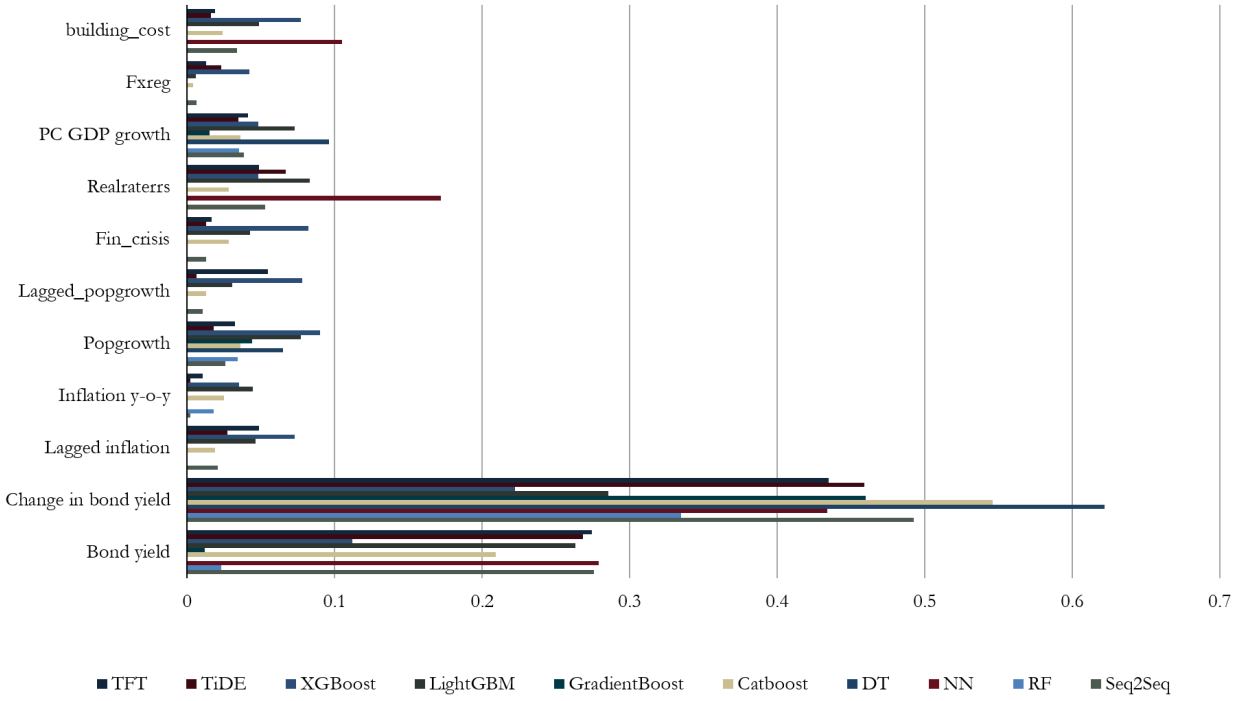
As reported, via Figure 8, we provide the variable importance, by model, for the most powerful machine-learning methodologies. This exercise is analogous to determining variable importance in a classic regression analysis. The importance refers to the training dataset that utilizes 1870-2020 data, across the ten leading machine-learning models. For the pre-1869 co-variates and data points I mainly rely on published GDP and population time series in the updated Maddison database (Bolt and van Zanden, 2020), fiscal time series in Dincecco (2013), interest rate and inflation data in Schmelzing (2026), and Metrick and Schmelzing (2026) for early financial crises observations. Additional variables with minor importance across models include inflation (year-on-year and lagged), the FX regime ("Fxreg"), year-on-year change in building cost index, and war events.[37]

We generally note that variable importance can vary substantially across models, but that a handful of variables retain high important across approaches – most notably the nominal (long-maturity) sovereign interest rate, year-on-year change in the sovereign interest rate, and (at some distance)

[37]The FX regime is coded on the basis of Karaman et al. (2020), assigning a binary value to "fixed" or "floating". The war variable is a binary one, too, using war codings in Clodfelter (2017). The building cost index is newly constructed on the basis of secondary source price data, see Appendix section 3.1 for details.

lagged real interest rates.

Figure 8: Variable importance, by ML model.




*Notes:* Variable importance for the top-11 most influential variables in each model. Variable importance within each model is normalized to sum to one, excluding "time".

In Figure 9, I compare the nominal yield series with the ML-derived model average bond total return series.

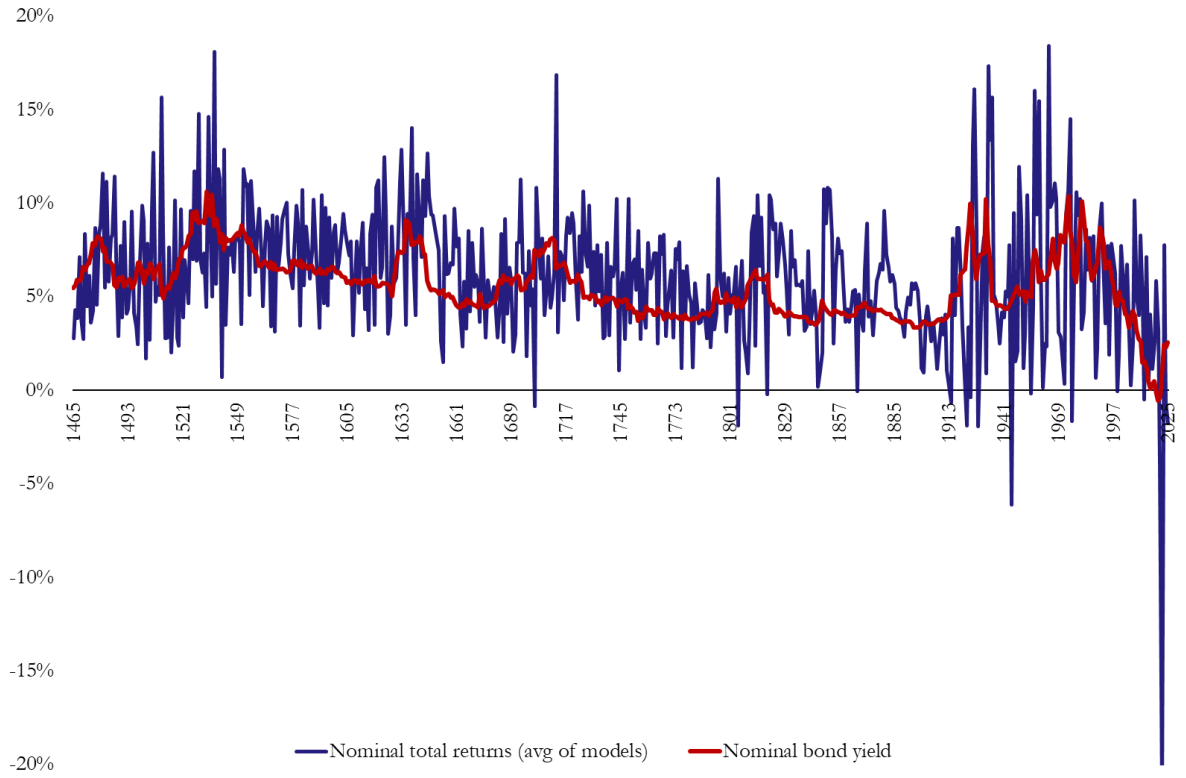
Generally, we should stress that the newly derived bond total return series here does not bias results in foregoing or subsequent exercises – but it enriches the panorama and refines many of the empirical propositions. It is critical to emphasize that a pure "hold-to-maturity" assumption for bond total returns – an assumption that entirely relies on the German yield data separately used in Schmelzing (2026) and Rogoff et al. (2024) – delivers identical results when it comes to econometric robustness, general slope propositions, and correlation trends (later detailed in section 9).[38]

## 6 Component 3: Cash

Existing benchmark series on capital returns generally focus on investable asset components, e.g. Jorda et al. (2019). Yet, given its dual role under the physical commodity standard – physical gold being the means of payment, as well as an investable asset – it could be worthwhile to entertain long-run reconstructions that incorporate cash shares: our underlying portfolio data does

[38]See in particular Table 5 – for "HTM" structural breaks; Appendix Table A.7 – for the confirmation of "HTM" trend stationarity / ADF-GLS results. And Appendix Table A.10 – for the confirmation of "HTM" correlation results.

Figure 9: German bond yields vs. Bond total returns, 1465-2025.




*Notes:* The Figure compares the primary-sourced German bond yield data with the ML-generated bond total return series (average of four leading models). Bond total returns after 1943 use "JST".

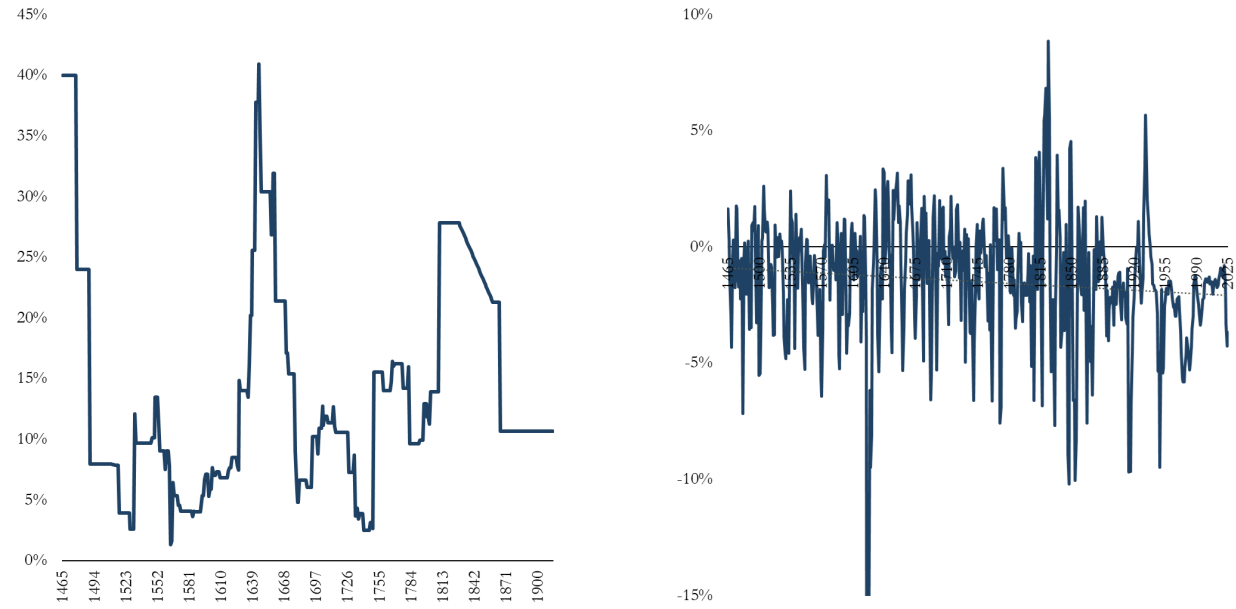
systematically include cash shares and thus allows for a version of capital returns with additional cash returns over time.

Under the pre-1914 gold standard – which, in practice, is better defined as a gold-silver standard –, "cash" consists of a mix of physical gold and silver coins. Most regularly, the "Goldflorin" (also known as the "Rheinflorin") is mentioned in the German household sources as a specific cash denomination: the Goldflorin served as the German benchmark gold coin modeled on the Florentine Florin, with a weight of 3.54 grams of pure gold as of the early 1500s. Over time, as was the case for other gold and silver currencies, Goldflorin purity was eroded through debasements. Exact gold weights are reconstructed in Karaman et al. (2020) (for Austrian coins) and Schmelzing (2026) (for German coins). After 1914 and the end of the gold standard, cash is an increasingly marginal component of the capital stock (and portfolios), and from this point onwards, I do not separately impute gold and silver stock returns (but only associate as "cash" paper money).

Cash shares are displayed via Figure 10: the all-time average (1465-1914) for cash stands at 13.9%, with sharp rises in this case share during the Thirty Years' War (to a peak of 38% over 1642-4), and the late stage of the Napoleonic Wars (1810-5 at 27.8%). In the post-1914 era, cash accounts for a mere 4.7% of the German capital stock.[39]

[39]In Goldsmith (1985), German cash shares bottom out at 3% in the 1970s. I keep a constant 3% share of cash over the

Figure 10: German cash share, and real cash return, 1465-1914 (2025).




*Notes:* The Figure displays the cash share of early modern portfolios (left hand side), as well as the ex post return on cash in Germany, the inverse of the year-on-year inflation rate (right hand side).

In my reconstruction, cash returns the inverse of the inflation rate: -1.3% p.a. on average over 1465-1914 (-1.49% over 1465-2025). I thus ignore storage and other costs – as well as the question whether true market prices at times deviated from the fixed nominal mint price (say, anticipating debasements or de facto inflows of foreign gold).[40]

## 7 The Fusion: Capital Returns and their Decomposition, 1465-2025

We can now proceed to the final step in the reconstruction process: the fusion of bond and real estate total returns, into a weighted capital return series. And on the basis of the preceding sub-section, the portfolio approach to capital stock over time now also allows us to present a variety that incorporates cash stocks.

With the Appendix providing much additional nuances about components and intermediary results, Figure 11 now displays our benchmark inflation-adjusted capital return versions, across two weighting methods, over the full 1465-2025 horizon. As discussed previously, I use a lagged inflation adjustment, as is standard in respective literature.[41] The investable asset share-weighted

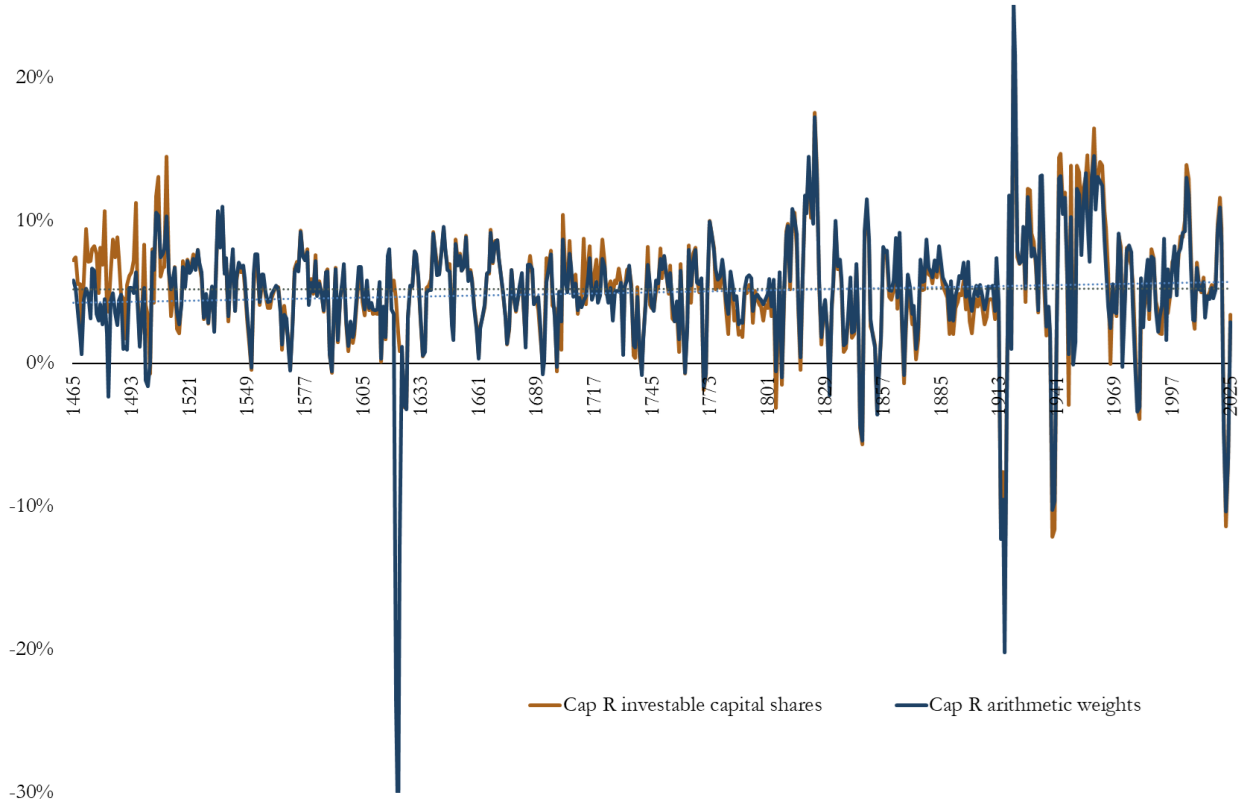
post-1971 era.)

[40] I also ignore potential service flows of gold, which may be relevant at least in the more recent era, according to Barro and Misra (2015).

[41] For further decompositions, including nominal aggregate returns, see Appendix section 4. On inflation specifically, a progressively lagged seven-year CPI inflation construction for Germany is used, sourced from Schmelzing (2026):

return series records an average of 5.25% over time, and increases by 0.24 basis points per annum. The arithmetically weighted counterpart stands at an average of 5.57%, and rises by 0.62 basis points per annum. I regard these two series in Figure 11 as the most appropriate series to capture "nationwide" returns to capital – however, per Table 3.2, we should also note that slopes and directions of the trend are sensitive to several key underlying features of the capital stock basis.

Figure 11: Real Capital Returns, investable asset / arithmetic weighting, 1465-2025.




*Notes:* The Figure displays the German Real Capital Return series, across two alternative weighting methodologies. The orange line uses the rolling investable capital shares of both assets (sovereign bonds, real estate) as described in chapter 3. The blue line uses equal weights (arithmetic shares) of the two assets. Real returns are gross (pre-housing expenditures, pre-tax), and both variations use the benchmark "Urban-rural" version of the real estate series. For a net return version see Appendix Figure A.1.

## 7.1 Decomposition: capital yields and capital gains

Recent debates have illuminated the precise decomposition of total returns over the past century: while Knoll et al. (2017) in their real estate work have suggested a decisive role for rental returns – one that plays the main role in accounting for superior returns of housing.[42] Similarly, Kuvshinov

the inflation data here includes rent components, but is generally based on urban prices. See discussion in Appendix section 3.1 and Figure A.2 for inflation variations.

[42]Even though Chambers et al. (2021) find rental returns that are just three-quarters the Knoll et al. (2017) estimate (3.0% p.a. vs 4.0% p.a.), but they also revise downwards real capital gains contributions for U.K. data over 1901-1983

Table 3.1: Summary Statistics for Capital Returns, S basis (ex cash), 1465-2025.

| | nom TR | real TR | capital gains | yield |
|---|---|---|---|---|
| 1465-1603 | 6.79 | 5.50 | 0.93 | 5.86 |
| 1603-1703 | 5.61 | 4.09 | 0.75 | 4.86 |
| 1703-1803 | 6.20 | 5.22 | 1.14 | 5.06 |
| 1803-1903 | 7.43 | 6.03 | 2.25 | 5.18 |
| 1903-2025 | 8.22 | 7.95 | 2.62 | 5.59 |
| Total | | | | |
| 1465-2025 | 6.90 | 5.40 | 1.54 | 5.36 |

*Notes:* The table reports summary statistics of the weighted capital returns, across relevant measures. Figures reported are using the urban+rural housing return basis, and assets are weighted by the "Schmelzing" ("S") (investable) capital stock shares over 1465-1914, ex cash. For a version with the cash component and alternative weighting, see Appendix.

and Zimmermann (2020) found a rising role for capital gains, at the expense of equity yields, in equity total returns over 1870-2020. Our new data may illuminate therefore the longer context of the precise decomposition return elements, and whether the predominant narrative of a "low yield environment" is affected.

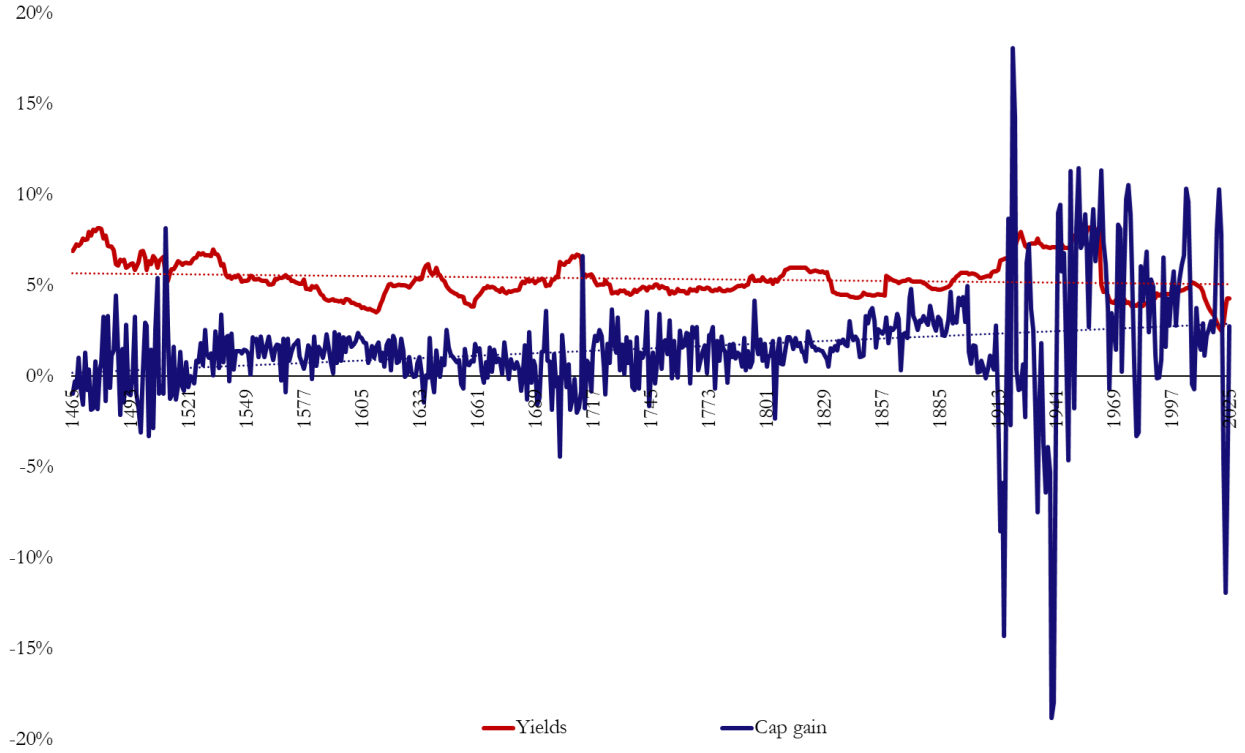
Indeed, via Table 3.1 that decomposes capital returns on the aggregate basis, a key observation stands out: historically, the majority of the return on capital over time was accounted for by the yield component. Indeed, over extended periods of time in the early modern era, capital gains are contributing less than one-fifth to the total capital return – with the share only rising gradually from the late 17th century. Yields, in comparison, appear to be far less volatile – over the course of centuries fluctuating in a relatively tight range of 5.1 – 6.9% – and show a moderately declining trend, on the order of -0.11 basis points per annum. In light of the recurrent debates over a "low rate environment" in the early 21st century, this perspective suggests that:

- Lower and lower asset yields appear to be a rather fundamental attribute of the asset universe over centuries – but even against the historic slopes, the decline in yields over the second half of the 20th century appears noteworthy.
- the proposition of a "low rate environment", meanwhile, appears nevertheless misleading – in the sense that it usually omits the secular total return context: we can now show that this "low-rate environment" actually does co-exist with a secular "high capital gains environment" which permeates different asset classes, and is clearly measurable on the level of aggregate capital returns, where it dominates the moderate secular decline in yields.

These basic insights are robust to many definitional variations, and underlying assumptions on costs, and weightings. But they are not universal. Table 3.2 records the linear trends associated with different variations of the benchmark return figures over time, as well as excess returns

(-0.7% p.a. vs 0.7% p.a.); rental growth rates are close to zero.

Figure 12: Decomposition of Capital Returns: yields and capital gains, 1465-2025.




*Notes:* The Figure displays the decomposition of German capital returns, into yield and capital gains components. Both components weighted according to investable capital stock shares.

(subsequent section).[43] We note that capital returns slope upwards in four out of the six variations in weightings and cost assumptions: specifically, when not adjusted for any of rural, cash, net, and size factors, the real return on capital in fact slopes downward (series 3); similarly, when adjusting for rural, net, and weighting real estate and bonds arithmetically, the slope is equally downward (series 5). This highlights the sensitivity of capital returns in particular to the real estate size- and capital stock-share adjustments – but affirms the base case of treating capital returns as a generally upwards sloping variable over the long run.

## 7.2 Excess returns / Risk Premia over time

The availability of the two main portfolio components over time allows a granular reconstruction of ex post excess returns. Our interest is particularly the application to the literature over recent years that emphasized moderately rising excess returns (risk premia) in Western economies "since the 1970s". But this literature is near-exclusively focused on equity risk premia, with more recent debates predominantly focusing on *levels* of excess returns, and similarly exclusively on the "short" horizon. An exception is Jorda et al. (2019), where housing is the key component of aggregate risk premia – which are found to be "trendless". Specific "housing excess returns" have been

[43] See Appendix Figure Figure A.10 for an illustration and some further details.

Table 3.2: Capital return iterations and trendlines, 1465-2025

Real Aggregate Capital Returns, 1465-2025

| TYPE | Rural | Cash | Net | Arith | size-adj | slope (BPS) | EX (BPS) |
|---|---|---|---|---|---|---|---|
| series 1 | X | | | | | 0.249 | 0.730 |
| series 2 | | | | | X | 0.146 | 0.627 |
| series 3 | | | | | | -0.191 | 0.291 |
| series 4 | | X | | | | 0.261 | 0.743 |
| series 5 | X | | X | X | | -0.138 | 0.344 |
| series 6 | X | | X | | | 0.217 | 0.699 |

*Notes:* The table reports variations of aggregate real capital returns, and the resulting linear trendline slopes over 1465-2025 associated with the variation. "X" denotes "included". "Rural", "Net" and "size-adj" are real estate - specific adjustment choices, as discussed per section 4. The "Cash" and "Arith" adjustment choices occur on the cross-asset level. "EX" = excess return slope. BPS = basis points per annum.

investigated prominently in the Case-Shiller canon (Case and Shiller, 1990), though in a narrow market setting.[44]

In what follows, two housing excess return definitions are analyzed, following standard literature that calculates housing returns over long-maturity bond yields (Campbell et al., 2009; Eichholtz et al., 2021) – specifically non-leveraged gross excess returns EX for year $t$ are:

$$EX_{i,t} = R_{i,t} - i_{i,t} \tag{3}$$

With $i$ representing the nominal long-maturity interest rate.[45]

Taking into account the notion that governments, especially in the initial period, may not necessarily have been identified as the "safe" asset provider, I calculate another variety of excess returns that takes into account cash holdings. In this case, the "safe" basis represents a weighted measure of bond yields + cash returns (c), the weights being the rolling relative share (s) in the capital stock.

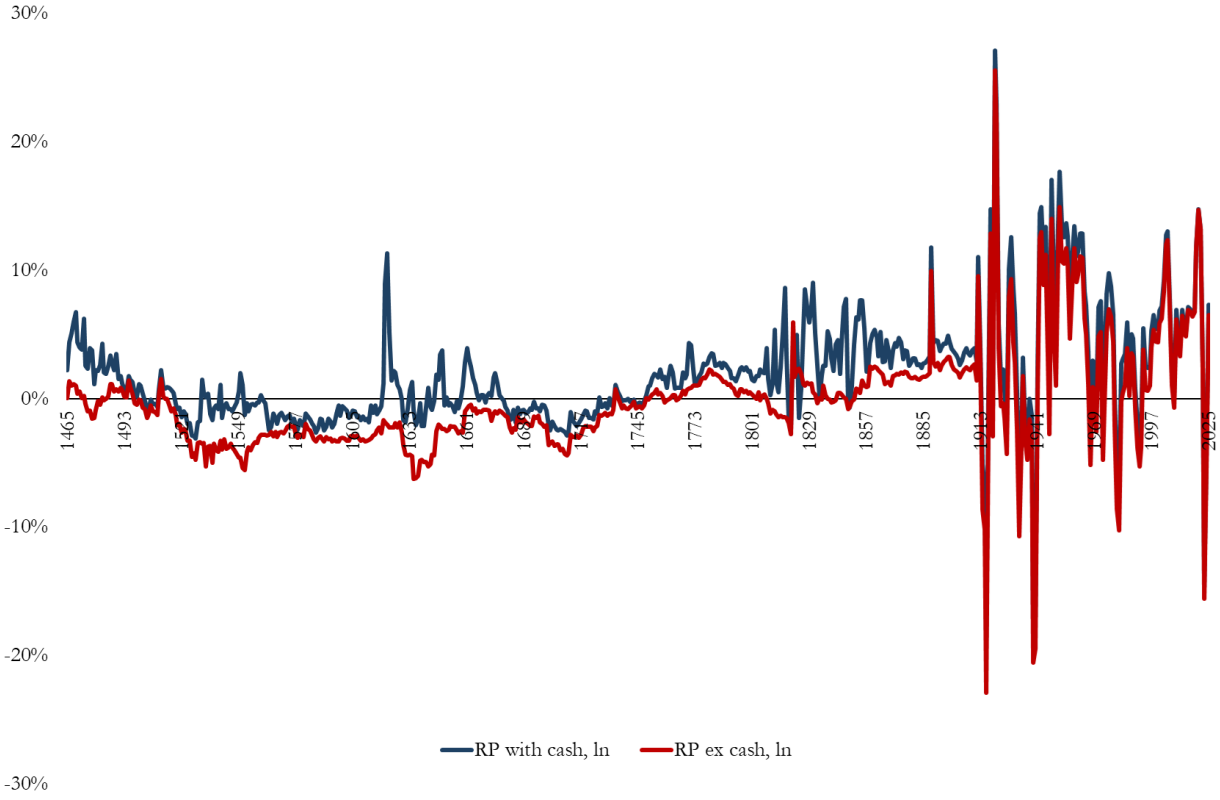
$$EX_{i,t} = R_{i,t} - [(s_{i,t} * i_{i,t}) + (s_{c,t} * c_i, t)] \tag{4}$$

We see these two variations displayed in Figure 13. Visually, both series appear to be upwards loping over the long-run, and display increasing volatility over the past twoc centuries. Between circa 1500-1750, both series appear to suggest sustained period of negative excess returns – that is, lower real estate total returns compared to the "safe" basis (whether or not cash is included). From circa 1800, excess returns are generally positive, though subject to occasional sharp downward

[44]In Case and Shiller (1989), the U.S. housing market between 1970-86 is analyzed: it is found that both U.S. real house price changes, as well as pre- and post-tax U.S. housing excess returns (housing total returns less the bill rate) do exhibit forecastability on the aggregate index levels. Individual house prices and excess returns, on the other hand, fail to exhibit forecastability.

[45]Note that in Eichholtz et al. (2021), the authors calculate excess returns on a net housing return basis, but subtract gross nominal long-maturity yields.

Figure 13: Excess returns, log, Germany, 1465-2025.




*Notes:* The Figure displays log German excess housing returns in two variations. The first variation is calculated as housing net total returns less a weighted cash + government bond yield basis (blue line, ln of equation 4); the second variation excludes cash and is defined as housing net total returns less government bond yields (red line, ln of equation 3). Housing returns in both cases refer to the urban+rural construction, post-tax.

shocks. These downward shocks appear to be consistently related to geopolitical and/or inflation shocks (including 1636, 1820, 1916, 1939, 1980, 2023).

Formal analysis confirms that the visual trends are indeed also econometric trends, with strong significance for trend stationarity (Table 4). Excess returns *do* exhibit long-horizon trends, as opposed to recent literature, which posited "trendless" excess risky returns based on shorter time horizons Jorda et al. (2019). Conversely, however, Table 4 confirms the findings for the U.S. housing market over the much shorter 1970-86 period in Case and Shiller (1989) for a much longer horizon: it is not exclusively within modern housing markets, or just the U.S. setting, where forecastability appears to exist. Once the sample lengths common in existing literature are significantly extended, the finding of forecastability (trend stationarity) appears to strongly replace notions of "trendless" risk premia – thus echoing recent findings for global interest rates that highlighted similar biases stemming from limited sample lengths (Rogoff et al., 2024).

# 8 Econometrics: formalizing trends and structural breaks

We turn to econometrics – and the question whether the key visual trends proposed in the foregoing can be formally confirmed. That includes an investigation of trend stationarity, as well as structural breaks ("turning points") in capital returns over time. If the answer is affirmative, consensus narratives on "trendless" returns and particular inflection points in key components of capital returns (prominently housing in the mid-20th century) can be more decisively questioned.

We recall that in their seminal paper, Case and Shiller (1989) analyzed U.S. city-level real house price index changes and excess returns, and rejected random walks in both variables. In representative fashion, the authors use regression analyses on post-1945 U.S. residential real estate data, a horizon virtually all subsequent econometric work focused on, and typically using regression analyses (for a summary see Ghysels et al. (2013)). None of the existing studies to my knowledge identified secular *trends* in returns, however, and I am unaware of any consistent econometric exercises being applied to sufficiently long-run data: existing research emphasizes that only at such horizons, the full statistical power of conventional econometric tests can be leveraged.[46] At the very least, therefore, one would expect that our data can confirm existing consensus on a much more general level.

Where we would expect potentially more substantial revisions are on the debates of "structural breaks" or "inflection points" in global housing: the long horizon could reveal the sensitivities embedded in existing short horizon literature, including the proposition that the mid-20th century marks a decisive break in the asset class.

## 8.1 Stationarity and forecastability

We begin with ADF-GLS tests (Elliott et al., 1996), a standard approach to confirm or reject forecastability (unit roots) and the existence of stationary time trends. Via Table 4 we can, first, clearly confirm forecastability for all German total return (gross and net) as well as real price series. The stationarity holds for variations of the test without a time trend, always at the 1% significance level, per the Appendix: there I also report a variety of ADF-GLS robustness checks, all country level results not shown in the main table, and also total return decompositions.[47]

Next, we observe that the confirmation of trend stationarity extends to our non-German, non-primary source price and return time series – including clear 1% significance for our benchmark "global" constructions across AW and GW methodologies: once more, these results hold even when not assuming a time trend, and they are also confirmed in a decomposition exercise that tests rent growth separately (reported via Appendix table A.15), a finding that suggests assuming a degree of income growth forecastability and that I will exploit further in section 7.

[46] See Rogoff et al. (2024) for interest rates: here, the true positive share for the ADF-GLS test falls off when samples

Table 4: ADF-GLS for German capital returns and components, 1465-2025

| | no of lags | t statistic | optimal lag |
|---|---|---|---|
| **Housing TR, 1465-2025** | | | |
| Real gross housing log return, 1466-2025 | 3 | -7.7020 | Seq |
| | 2 | -7.4185 | SIC, MAIC |
| | 1 | -9.9690 | |
| Real net housing log return, 1466-1910 | 3 | -7.9546 | Seq |
| | 2 | -7.5909 | SIC, MAIC |
| | 1 | -10.149 | |
| **Bond TR, 1465-2025** | | | |
| Nom TR, log, 1465-2025 | 3 | -4.1419 | SIC, Seq, MAIC |
| | 2 | -5.3389 | |
| | 1 | -7.1538 | |
| Real TR, log, 1465-2025 | 3 | -8.2856 | MAIC |
| | 2 | -9.7336 | |
| | 1 | -11.2450 | SIC, Seq |
| **Weighted Cap R, 1465-2025** | | | |
| Nom Cap R, S weights, ln | 3 | -6.4193 | |
| | 2 | -6.6541 | SIC, Seq, MAIC |
| | 1 | -8.9064 | |
| Real Cap R, S weights, ln | 3 | -10.032 | SIC, Seq |
| | 2 | -9.0458 | MAIC |
| | 1 | -11.345 | |
| Real Cap R with Cash ln | 3 | -9.7449 | Seq |
| | 2 | -9.8674 | MAIC |
| | 1 | -12.304 | SIC |
| **Capital R components, 1465-2025** | | | |
| Yields, S weights | 3 | -1.4677 | |
| | 2 | -1.5171 | |
| | 1 | -1.4617 | SIC, Seq, MAIC |
| Cap gains, S weights | 3 | -9.7715 | Seq |
| | 2 | -9.694 | MAIC |
| | 1 | -12.054 | SIC |
| **Excess Returns, 1465-2025** | | | |
| RP1 with cash | 3 | -8.1288 | SIC, Seq |
| | 2 | -7.5129 | MAIC |
| | 1 | -10.307 | |

*Notes:* The table reports the ADF-GLS test statistic for several choices of the number of lags $k$ (with a maximum of three lags). The regression includes a constant. The test assumes a time trend. The critical values at the 1, 5, and 10 percent significance levels are the following: $-3.48$ (1 percent); $-2.87$ (5 percent); $-2.59$ (10 percent). "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. "S weights" are the investable share weights derived from Schmelzing (2026): see the Appendix for results for alternative weightings.

### 8.2 Bai-Perron

Next, we test the long-run series via the standard structural break test of Bai and Perron (1998). The Bai-Perron methodology tests time series for up to five structural breaks, has been widely used in the financial and economic literature, and has in particular been employed to identify "turning points" in the canonical housing literature. Specifically, Knoll et al. (2017) use Bai and Perron (1998) as their basis to suggest the 1960s as the inflection point in advanced economy housing markets.

The test is however sensitive to the sample length used, and meaningfully improved sample length in particular have been shown to revise canonical structural break narratives in recent macro-finance literature (Rogoff et al., 2024).

I report the results from our long-run series in Table 5, which employs a standard set-up of the test.[48] We observe that generally 2-4 structural breaks are identified, always short of the maximum allowed. This applies to the series that are purely derived from primary sources (real estate total returns), as well as those that feature combinations of the primary source and ML-derived components – and it applies to iterations with and without cash components. Most series feature one (rarely two) early modern structural break, in addition to one 20th century structural break that takes place either around World War One or World War Two. No structural break after 1941 is ever identified, in any setup. The series for sovereign bond total returns feature no structural break at all – in what appears to be somewhat of an outlier.

## 9 Capital Returns and the Economy: a gradual divorce

What is the long-run relationship between capital returns and the economy? On this debate, much attention has been centered around "R-G" dynamics, which proxy well for inequality tendencies, according to influential studies. Rather than reconstructing such R-G series for Germany (which is feasible), I am here more interested in the long-run trends in the correlation between the real economy (as measured by consumption and the business cycle) and capital returns.[49] We have previously seen that changes in the hedging properties of bonds and alternative asset classes have been invoked to explain rising risk premia in advanced economies since circa 1980. This makes a reconstruction of the very long relationship intriguing: might the secular rise in excess returns have its parallel in a secularly increasing divorce of bond returns with the economy?

On the asset-level, reconstructing such correlations also provides us with a first empirical picture of the *relative safety* of key assets in the capital stock. Indeed, while very rich with working

of "only" 100 years of annual observations are used (to 68.8%), and drops even lower for 50 years (29.8%).

[47] See Tables A.7-A.8, along with discussion there. There, I test variations without a time trend, and raw year-on-year (as opposed to log) change variations.

[48] As is conventional in the literature, the Bai-Perron is run with a maximum of five structural breaks allowed.

[49] But using more narrow data on "R" (only encompassing real government interest rates), some long-run R-G series were presented by Rogoff and Schmelzing (2024).

| Table 5: Structural break tests, Bai-Perron basis | | |
|---|---|---|
| | no of breaks | break years |
| Real Estate, Germany, 1465-2025 | | |
| Real gross total return, urban-rural, log | 2 | 1548, 1743 |
| Real net total return, urban-rural, log | 2 | 1548, 1917 |
| Real gross total return, size-adjusted, log | 3 | 1620, 1743, 1902 |
| Real net total return, size-adjusted, log | 2 | 1548, 1917 |
| Sovereign Debt, Germany, 1465-2025 | | |
| AVG of models, nom TR | 0 | – |
| AVG of models, real TR | 0 | – |
| Yields only, nominal ("HTM") | 2 | 1725, 1913 |
| Aggregate Capital Returns, 1495-2025 | | |
| Capital R, invest, nom, ex-cash | 2 | 1626, 1941 |
| Capital R, invest, nom, with cash | 2 | 1589, 1941 |
| Capital R, invest, real, ex-cash | 2 | 1548, 1941 |
| Capital R, invest, real, with cash | 2 | 1825, 1917 |
| Net real TR | 2 | 1548, 1917 |
| Components, 1495-2025 | | |
| Yields | 4 | 1548, 1681, 1857, 1941 |
| Capital gains | 2 | 1625, 1941 |
| Measures of log Excess Returns, 1465-2025 | | |
| RP1 – Housing TR over bonds+cash | 2 | 1551, 1917 |
| RP2 – Housing TR over bonds | 2 | 1548, 1917 |

*Notes:* 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 linearly detrended series using a trimming parameter of 5 percent and a heteroskedasticity and autocorrelation consistent variance estimator. The significance level is 5 percent, and the maximum number of allowed break points is 5 for all series. When the test rejects the hypothesis of no breaks, the estimated number of breaks and the break dates are reported in the second and third columns, respectively.

assumptions on "safe" assets, the literature thus far has (to my knowledge) never reconstructed such a history on a plausible quantitative basis. Long-standing financial and historical literature relates either explicitly or implicitly to real estate as a (if not *the*) "safe asset" in Western financial systems prior to the rise and consolidation of government credibility. In representative fashion, Gorton (2017) states that "Privately produced safe debt must be backed by credible collateral. Since roughly 1750, privately produced safe assets have consisted of different types of debt, usually mortgages, as land is usually most likely to maintain its value".

Both historical and economics literature over many years abound with circumstantial evidence that elite households, investors, and governments very much operated with a notion of "safe assets". As a representative case, we can take McFarlane's analysis of John Fastolf's sprawling portfolio, the famed English mercenary oligarch of the 15th century. Pondering the further purchase of castles, and their relative safety at the time, we learn that:

- "If a man was anxious to make a wise disposition of his fortune, land had at least one clear advantage. From the negligence or dishonesty of executors there was no sure refuge, but land was less easily misappropriated than cash. It did not evaporate as movables could during the nonage of the heir; it was better protected in the courts; whether its financial yield was high or low, it was altogether safer."[50]

The broader question this Section seeks to address, therefore, does not only speak to a very basic (but untested) finance and history notion – namely, reconstructing along modern notions the relative "safe asset history" in a key Western economy. But as importantly, we are after the *time trends* in whatever initial risk profile prevailed in this asset class: is the presence or absence of any risk evolution in fact consistent with the observation of rising risk premia (excess returns) over centuries?[51] We recall that there is debate not just over the existence of any, if mild, trend in modern times. But particular factors driving patterns in risk premia in macro-finance literature are associated with rising macroeconomic risks (Farhi and Guorio, 2018), declining sovereign risk (Wachter et al., 2026), or changes in risk aversion (Siegel, 1992; Marx et al., 2021).

## 9.1 Capital Returns and the Economy – A Long-Run View

Modern Housing returns as well as residential investment are recognized as being strongly procyclical. This procyclicality is widely seen to account for Housing's "risky" asset market properties in the context of consumption based asset pricing frameworks (Piazzesi and Schneider, 2016). Conversely, it has been proposed that the increasing detachment of sovereign bond returns from the business cycle – increasing hedging properties – account for the multi-decade decline in sovereign bond yields from the late 20th century (Marx et al., 2021): sovereign bonds themselves, in this sense, have become safer over time (Wachter et al., 2026). If we follow the basic assumptions

[50] McFarlane (1957, 109).

[51] The relevant literature (e.g. Jorda et al. (2019)) uses "excess returns" and "risk premia" interchangeably, in both cases referring to ex post realized returns. I will follow this nomenclature.

of consumption based asset pricing, and project its premises onto longer horizons, we may be able, therefore, to study the evolution of underlying asset-level "riskiness": that is, whether the procyclicality of housing and bonds shows any time trend itself. Validating any existing time trends can help distinguish proposed contemporary driver, but also shed new light on historical relative risk patterns.

To undertake such an exercise over the long run, we are critically aided by the new data on German per capita consumption growth published by Pfister (2022). The Pfister (2022) data spans 1502-1855 and focuses on per capita food consumption growth, with food comprising more than 64% of median income household non-rent consumption prior to 1914, and still more than 50% for the top 20% households.[52]

To take account of the declining share of food consumption in the overall consumption basket over time, Figure 14 displays two long-horizon correlation variations between capital returns and consumption: first, one that consistently focuses on per capita *food* consumption growth over time, irrespective of its shrinking share in aggregate consumption (red line); second, a variation that incorporates the "dynamic" evolution of consumption baskets over time, and moves to a general consumption growth basket after 1877 (blue line). The key observation in Figure 14 is that both definitions exhibit clearly cyclicality themselves, but exhibit a downward trend over time: the correlation was on average positive in the early modern economy, and has sharply declined over the past century. It does not appear that the secular "divorce" of capital returns from the business cycle is sensitive to either of the plausible constructions of consumption trends over time, despite switches in the underlying data set bases.

The long-run divorce on the aggregate level holds on each of the decomposed asset levels, for real estate and for bonds.[53]

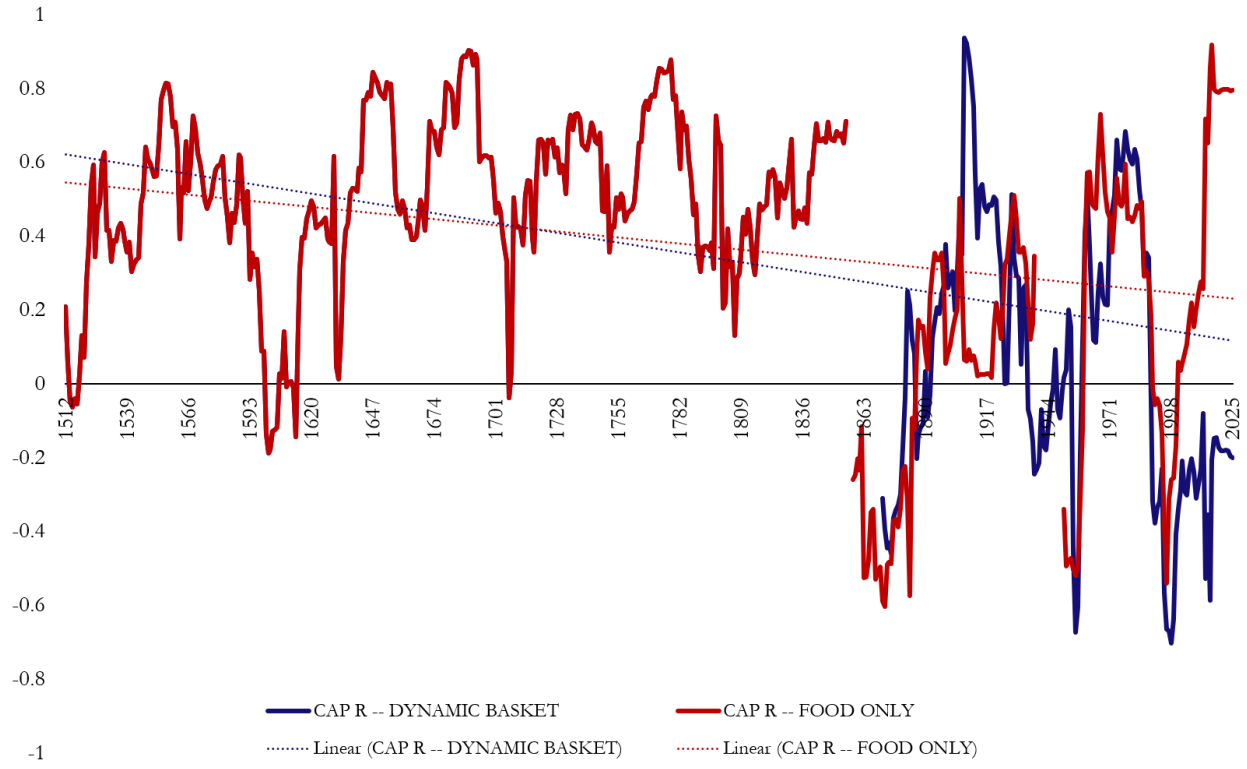
Based on the actual long-run procyclicality of the two key assets in the capital stock, then, we can derive new stylized facts about the "riskiness" – in both absolute and relative terms – of their returns over time. At least in measurable empirical terms, real estate (including land) did not *perform like* a safe asset in the early modern financial system: indeed, real estate total returns, across a variety of measurement variations, was highly positively correlated with per capita consumption growth in the pre-1618 era already. Real estate's procyclicality gradually fell over time, to reach moderately positive correlations with consumption growth as of the late 20th and early 21st century.

Interestingly – and here (to my knowledge) first documented – sovereign debt already exhibited a (moderately) lower average correlations with consumption growth compared to real estate in

[52]Historical consumption shares are reported in Hoffman et al. (2002, Table 1). The years 1856-1876 are excluded, and after 1877, the "food only" consumption basket (red line) correlation takes data from the Federal Statistical Agency (Destatis); the "dynamic" consumption basket (blue line) correlation takes real per capita consumption growth for Germany from the "JST" database.

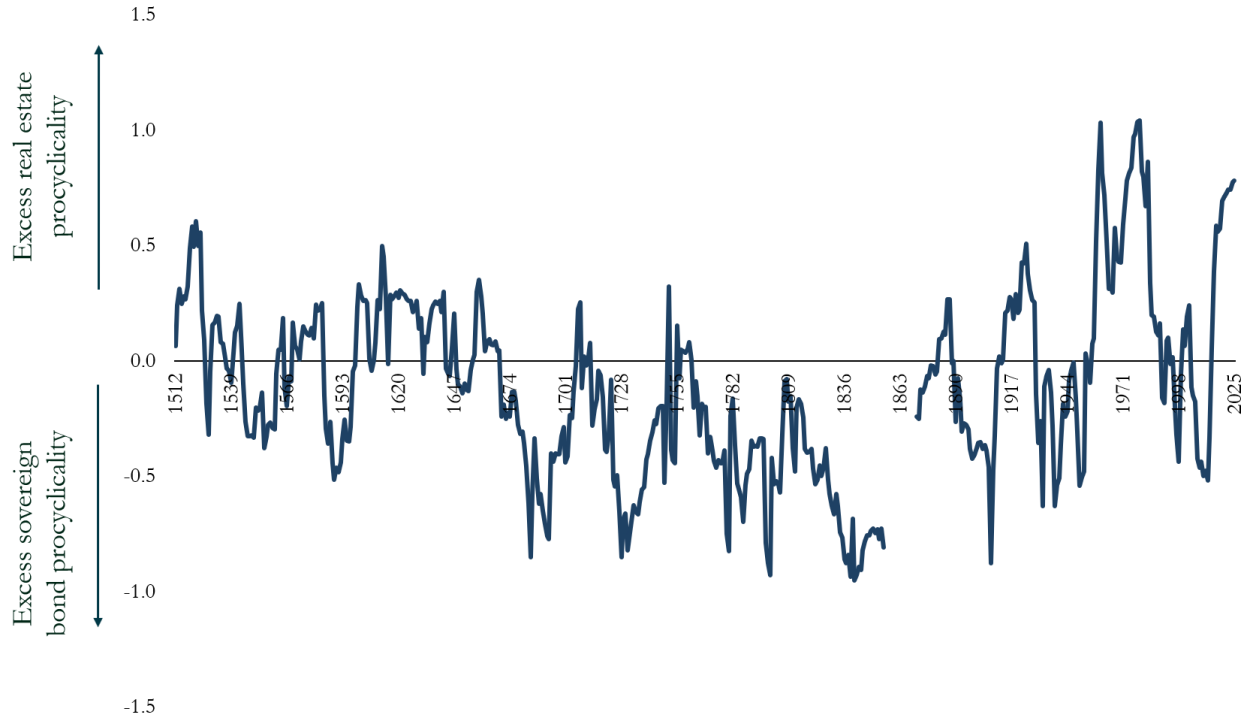
[53]See Appendix section 5.3 for the disaggregated correlation trends, confirming that the downward trend holds individually over time, for both real estate and bonds.

Figure 14: Aggregate Capital Return correlations to p.c. consumption growth, Germany 1512-2025.




*Notes:* Two variations of aggregate real Capital Returns and their correlation to real per capita consumption growth are shown. The red line depicts the rolling correlation between Capital Returns on the "urban-rural with cash" variation, and per capita food consumption growth, for Germany. The blue line depicts the rolling correlation between Capital Returns on the "urban size-adjusted" variation, and a "dynamic basket" of real per capita consumptions growth (which takes into account the declining share of food in consumption), for Germany. 15-year centered rolling correlation is used. Annual real per capita (food) consumption growth data is used over 1512-1855, via Pfister (2022). After 1871, food consumption data in Hoffmann (1965) and DeStatis, and aggregate post-1871 per capita consumption via "JST". The years 1856-70 are excluded given missing consumption data.

Figure 15: Excess Real Estate/Sovereign Bond pro-cyclicality, Germany, 1512-2025.




*Notes:* The line depicts the excess rolling correlation with consumption growth for real estate total returns over government bond total returns. More positive numbers in this sense indicate greater relative safety of bonds over real estate, and vice versa. Same sources and series variations are used as per Figure A.11.

the early modern era. Importantly, this result is again not sensitive to the machine-learning components that I used to derive bond capital gains in the early modern era: per the Appendix, even if we use "HTM" bond total returns (recall that they only depend on the archival yield data), the observation fully holds.[54] Like real estate, the secular procyclicality for bonds is declining, but unlike real estate, sovereign bonds exhibit a *negative* correlation to per capital consumption growth in the modern era (1950-2025). Sovereign bonds, in other words, have become a measurable consumption *hedge*.

So in absolute terms, falling procyclicality is a cross-asset phenomenon. But is there also a secular *relative trend*? Figure 15 provides a straightforward answer, displaying the spread between the earlier individual correlation trends – real estate correlation less the sovereign bond correlation over the same period (1512-2025).

What the Figure visualizes is the appearance of a "U-shape" of the relative safety evolution between real estate and sovereign bonds, over the past half millennium. Positive values indicate an excess procyclicality of real estate over sovereign bonds (keeping the same definitions as before): the intriguing observation is that sovereign debt seems to have lost a pre-existing relative safety status that appears to have been in place prior to 1618, when Germany plunged into the prolonged geopolitical distress of the Thirty Years' War. Real Estate "reclaimed" a (again, relative) safety status

[54]See Appendix section 5.3 and Table A.10.

over the period of circa 1618-1850. Only during the 19th century did real estates' primacy decisively end. Despite the extreme physical capital destruction associated with the Thirty Years' War, the default waves documented for sovereign issuers during and after the same period apparently left deeper financial scars for the economy at large: this particular "U-shape" might thus well be unique to the German setting – though the trigger of such an inversion, the sudden extreme geopolitical shock, may be indicative of a broader driver of such inversions.[55]

### 9.2 Rising "Risky Asset Risks"? Physical Asset Destruction over Time

Over time and space, a key idiosyncratic risk to housing is its physical destruction. Today, such destruction risk is prominently associated with climate change, including rising sea levels (Bernstein et al., 2019). Historically, fire outbreaks arguably constituted the greatest source of physical destruction – whether caused directly by accident, or indirectly through wars and natural disasters.[56]

An alternative driver in our observation of rising excess returns over centuries might therefore be rising housing-specific risks. That is, there are no fundamental shifts in preferences for "safe" assets – risk aversion might stay constant over time; but greater and greater probabilities of housing destruction would require secularly rising investor compensation. This would clearly change asset pricing interpretations and the policy implications of the long-run trends observed.

To test the idea of "rising risk asset risks", in this section, I utilize new data on historical urban fire outbreaks in German cities – as collected recently in a comprehensive effort by Zwierlein (2021): the data records property-level destruction sourced from urban records in two ways – the absolute number of houses destroyed annually by fires, and the share of houses in the respective city destroyed conditional on a fire outbreak. This long-horizon data is displayed in Figure 16, left panel. We observe sharp spikes in absolute and conditional housing destruction during geopolitical shocks, especially the Thirty Years War (aligning well with our independent RHPI slump) and the World Wars. Conditional destruction ratios show a clear long-run downward trend, from almost 100% in the late Renaissance, to less than 40% (running average) by the 20th century. Without attempting a deeper investigation of particular drivers and turning points – which is done with a historical angle in Zwierlein (2021) – the takeaway is that the data pointedly confirms a long downward trend in "housing destruction risk", and headline movements in this index align well (inversely) with all variations of our German RHPI.

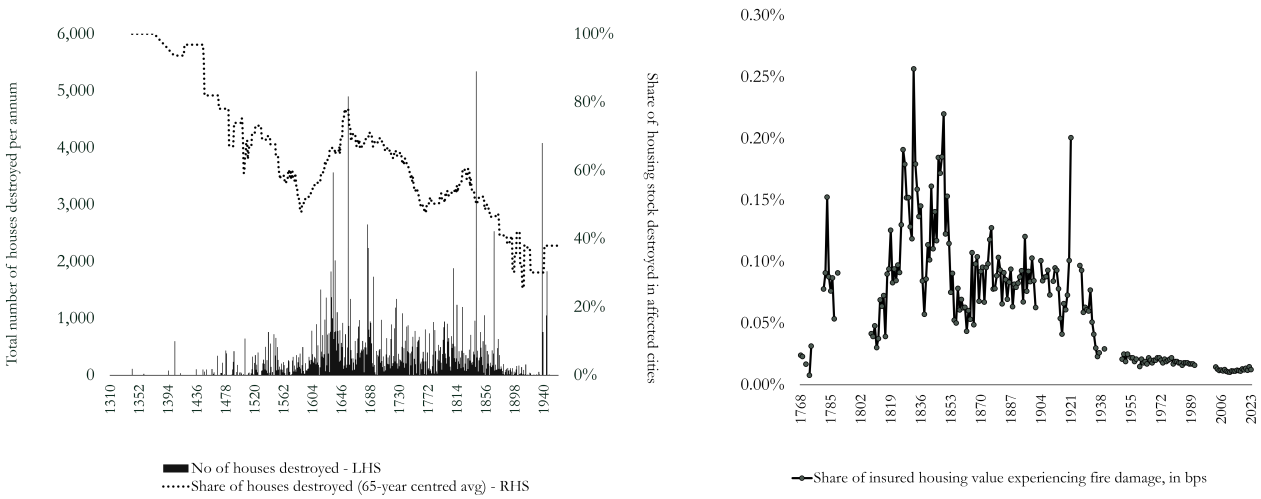
The right panel in Figure 16, meanwhile, displays independent data from German property insurance providers: the data is compiled from various state and regional property associations, and displays the share of property values that experienced fire damage claims on an annual basis, from 1768 onto present times.[57] We observe a similar long-horizon evolution in this variation

[55] Notice a similar, though more temporary, inversion associated with the interwar period, ca. 1918-1945.

[56] Of course, other low-frequency idiosyncratic risks exist – one might think of fundamental changes to rules and regulations (zoning, building standards), as well as technological innovations in building.

[57] Conditional on the property being insured – the share of properties being insured is sharply increasing over time.

Figure 16: Long-run housing destruction data: evidence from urban fire statistics and housing insurance claims, Germany 1310-2025.




*Notes:* The left panel displays German urban housing destruction by fire, in two variations, over 1310-1943: on the left y-axis (corresponding to black vertical bars), it measures the absolute amount of houses destroyed annually by fire in German cities; on the right y-axis (corresponding to black dashed line), it measures the share of houses destroyed conditional on a fire breaking out in a city. The data in this left panel is sourced from Zwierlein (2021). The right panel displays housing insurance data: the y-axis measures the annual claims related to housing fire damage, as a share of the total insured housing value in Germany, over 1768-2023 – here sources include Borscheid and Drees (1988).

of destruction risk: from annual ratios of 10-25 basis points of property value destructions in the 18th century, the ratio first stabilizes just under 10 bps throughout the 19th century, before experiencing another sharp decline, to an average of just 1-2 bps per year over 2003-2024. Without analyzing the granular turning points and drivers further at this point, the bottom line echoes the urban destruction data: housing has become a far safer investment in terms of its physical safety and longevity: at least directionally, therefore, housing thus echoes the rising absolute safety of government assets over time. The story, hence, it not once of *diverging* fundamental risk trajectories across the two key asset classes over time: rather, it appears that government assets managed to assume safety more intensely than any comparable asset in our universe.

## 10 Conclusion

Over 180,000 property-years merging both urban and rural real estate observations, combined with more than 3,200 long-maturity interest rate data points, allowed what arguably constitutes the most nuanced reconstruction of the "aggregate return on capital" over centuries for a single country. By focusing on a single country for which the most comprehensive archival sources were obtained, it was possible to test a wide range of underlying historical assumptions and components – and thus this paper enabled a new empirical panorama that contextualizes key macroeconomic

One important statistical insurance source used is Borscheid and Drees (1988).

debates over past years. It allowed for a dynamic weighted approach to the capital stock – creating meaningful improvements to balance sheet estimates in the process – by employing a portfolio approach over time. [58]

There are key revisions to a literature that has analyzed data spanning a century at best – and has thus been biased towards missing secular trends and robust inflection points. Among the propositions of this paper is the clear rejection of the idea that returns to capital, and measures of risk premia are fundamentally "trendless". Deeper fundamental trends appear to exist in Western economies, with the results for Germany holding clear implications for related countries (despite some obvious idiosyncrasies for Germany). These facts can not just be measured in-depth – they can be underpinned by now leveraging the greatly improved statistical power that come with long horizon data sets.

The "high R" environment of the 21st century, therefore, is no historical outlier – it stands in a multi-century line of macroeconomic developments. Not least, this long-run context qualifies the engrained narrative of a "low rate environment" – rates , indeed, appear to be (at least moderately) declining across asset classes over centuries – yet, the "low rate environment" co-exists with an increasingly "high total return and R" environment. The development appears to be driven by secularly rising capital gains across asset classes – coupled with at least moderately falling capital yields: this observation extends a finding emphasized in recent literature for a much longer horizon and on an aggregate level.

What is the relationship between capital returns and the macroeconomy? This question stands at the heart of canonical asset pricing models, as well as canonical macroeconomic debates (inequality, "secular stagnation", or the "K-shaped economy"). The reconstruction of "R" over time allows a new panoramic analysis of the interplay. Exploiting recent advances in the German data, I suggested that in the early modern economy, aggregate capital returns and its components were highly correlated to consumption growth – but that a secular weakening of this correlation in absolute terms is detectable, across variations of capital return measures, and for each component individually. Capital markets and the real economy secularly "divorced".

A separate prominent line of inquiry concerns the *relative* safety of the major asset classes – when did sovereign debt *measurably* attain hedging properties, and are such patterns plausibly related to excess return trends? I reconstructed the excess real estate correlation to consumption, over that of government debt returns. Contrary to stylized assumptions in traditional literature, the data suggests that real estate was already a measurably "risky" asset in both an absolute and a relative sense in the early modern era, at least as determined by its hedging properties. Yet – after a sharp inversion that appears to be related to the Thirty Years' War, and which triggered a two-century predominance of real estate – the ultimate fall in consumption correlation was clearly steeper for sovereign bonds, with the "Excess procyclicality" of real estate rising continuously since the mid-19th century. This observation is plausible in light of the finding of secularly rising "excess

[58] For instance, as Schmelzing (2026) and Rogoff et al. (2024) document, the downward trend in sovereign long-maturity rates since the Renaissance is a cross-country phenomenon.

returns" for real estate over bond yields. It does suggest that explanations invoking relative risk patterns for current 21st century market phenomena or rising risk aversion may both constitute promising lines of further inquiry: it also adds a notion, however, that an established "safe asset status" can reverse fundamentally in light of (geopolitical) macro shocks.

## References


Allen, Robert C. 1988. The Price of Freehold Land and the Interest Rate in the Seventeenth and Eighteenth Centuries. *Economic History Review* 41(1): 33–50.

Allen, Robert C. 2001. The great divergence in European wages and prices from the Middle Ages to the First World War. *Explorations in Economic History* 38: 411–447.

Ascher, Siegfried. 1917. *Die Wohnungsmieten in Berlin von 1880-1910. Eine Statistische Untersuchung*. C. Heymann.

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

Bai, Jushan, and Pierre Perron. 2003. Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics* 18(1): 1–22.

Ballerstedt, Maren. 2024. *Häuser und Menschen: Höfe- und Häuserchronik von Pretzien seit etwa 1600*. Ziethen Verlag.

Barro, Robert, and Sanjay Misra. 2015. Gold Returns. *Economic Journal* 126: 1293–1317.

Becker, Moritz A. 1888. *Hernstein in Niederoesterreich, sein Gutsgebiet und das Land im weiteran Umkreise*. Adolf Holzhausen.

Becker, Moritz A. 1889. *Geschichte von Hernstein in Niederösterreich und den damit vereinigten Gütern Starhemberg und Emmerberg*. Holzhausen.

Berlin, Statistics. 1897. *Statistisches Jahrbuch der Stadt Berlin, 24. Jahrgang, 1897*. Selbstverlag Stadt Berlin.

Bernstein, Asaf, Matthew Gustafson, and Ryan Lewis. 2019. Disaster on the horizon: The price effect of sea level rise. *Journal of Financial Economics* 134(2): 253–272.

Bianchi, Daniele, Matthias Büchner, and Andrea Tamoni. 2021. Bond Risk Premiums with Machine Learning. *Review of Financial Studies* 34: 1046–1089.

Bochow, Nils, Anna Poltronieri, Martin Rypdal, and Niklas Boers. 2025. Reconstructing historical climate fields with deep learning. *Science Advances* 11: 1–11.

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

Borscheid, Peter, and Anette Drees. 1988. *Versicherungsstatistik Deutschlands 1750 - 1985*. St. Katharinen: Scripta Mercaturae Verl.

Borup, Daniel, David Rapach, and Erik Schuette. 2023. Mixed-frequency machine learning: Nowcasting and backcasting weekly initial claims with daily internet search volume data. *International Journal of Forecasting* 39(3): 1122–1144.

Braeuer, Karl. 1915. *Studien zur geschichte der Lebenshaltung in Frankfurt a.M. während des 17. und 18. Jahrhunderts* . Joseph Baer.

Broadberry, Stephen, and Alexandra de Pleijt. 2021. Capital and Economic Growth in Britain, 1270-1870: Preliminary Findings. *CEPR Discussion Paper 15889* .

Brückner, Franz. 1989. *Häuserbuch der Stadt Dessau, 24 vols.* Rat der Stadt, Stadtarchiv.

Bundesamt, Statistisches. 2023. Wohnungsbestand im Zeitvergleich. *Zeitreihen, Stand 28. Juli 2023* .

Burgmaier, Andreas, and Gustav Schneider. 1966. *Häuserbuch der Stadt München, 5 Vols.* Oldenbourg.

Bürger, Sven-Uwe. 1992. Burg Amlishagen. III. Anmerkungen zur Besitzgeschichte. *Jahrbuch des Historischen Vereins für Württembergisch Franken* 76: 39–60.

Caldara, Dario, and Matteo Iaciovello. 2022. Measuring Geopolitical Risk. *American Economic Review* 112(4): 1194–1225.

Campbell, Sean, Morris Davis, Joshua Gallin, and Robert Martin. 2009. What moves housing markets: A variance decomposition of the rent–price ratio. *Journal of Urban Economics* 66: 90–102.

Case, Carl, and Robert Shiller. 1989. The Efficiency of the Market for Single-Family Homes. *American Economic Review* 79(1): 125–137.

Case, Karl, and Robert Shiller. 1990. Forecasting Prices and Excess Returns in the Housing Market. *Real Estate Economics* 18(3): 253–273.

Chambers, David, Christophe Spaenjers, and Eva Steiner. 2021. The Rate of Return on Real Estate: Long-Run Micro Evidence. *Review of Financial Studies* 34: 3572–3607.

Chen, Luyang, Markus Pelger, and Jason Zhu. 2024. Deep Learning in Asset Pricing. *Management Science* 70(2): 714–750.

Chmel, Joseph. 1858. *Actenstücke und Briefe zur Geschichte des Hauses Habsburg im Zeitalter Maximilian's I. Dritter Band*. Königlich-Kaiserliche Hof- und Staatsdruckerei, Wien.

Clark, Gregory. 2002. Shelter from the Storm: Housing and the Industrial Revolution. *Journal of Economic History* 62(2): 489–511.

Clodfelter, Micheal. 2017. *Warfare and Armed Conflicts: A Statistical Encyclopedia of Casualty and Other Figures, 1492-2015 (4th ed.)*. McFarland.

Cong, Lin, Ke Tang, Jingyuan Wang, and Yang Zhang. 2021. Deep Sequence Modeling: Development and Applications in Asset Pricing. *Journal of Financial Data Science* 3(1): 28–42.

Corelogic, SPGlobal. 2024. SP CoreLogic Case-Shiller Home Price Indices Methodology, April 2024. *available at https://www.spglobal.com/spdji/en/documents/methodologies/methodology-sp-corelogic-cs-home-price-indices.pdf* .

Das, Abhimanyu, Weihao Kong, Andrew Leach, Shaan Mathur, Rajat Sen, and Rose Yu. 2024. Long-term Forecasting with TiDE: Time-series Dense Encoder. *Transactions on Machine Learning Research* .

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.

Dincecco, Mark. 2013. *Political Transformations and Public Finances: Europe, 1650-1913*. Cambridge University Press.

Dirlmeier, Ulf. 1978. *Untersuchungen zu Einkommensverhaeltnissen und Lebenshaltungskosten in oberdeutschen Staedten des Spaetmittelalters*. Carl Winter.

Earle, Peter. 1989. *The Making of the British Middle Class*. U California Press.

Eberstadt, Rudolf. 1903. *Rheinische Wohnverhältnisse und ihre Bedeutung für das Wohnungswesen in Deutschland*. G. Fischer.

Eichholtz, Piet, Matthijs Korevaar, Thies Lindenthal, and Ronan Tallec. 2021. The Total Return and Risk to Residential Real Estate. *Review of Financial Studies* 34: 3608–3646.

Eichholtz, Piet, Thies Lindenthal, and Matthijs Korevaar. 2023. Limited Growth: 500 Years of Urban Rents. *Working Paper* .

Eisfeldt, Andrea, and Andrew Demers. 2022. Total Returns to Single Family Rentals. *Real Estate Economics* 50(1): 7–32.

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

Farhi, Emmanuel, and Francois Guorio. 2018. Accounting for Macro-Finance Trends: Market Power, Intangibles, and Risk Premia. *Brookings Papers on Economic Activity* 147–223.

Flamm, Hermann. 1905. *Der wirtschaftliche Niedergang Freiburgs i. Br. und die Lage des städtischen Grundeigentums im 14. und 15. Jahrhunderts*. G. Braunsche Hofbuchdruckerei.

Flandreau, Marc, Christophe Galimard, Clemens Jobst, and Pilar Nogues-Marco. 2009. The Bell Jar: Commercial Interest Rates between two revolutions, 1688-1789. In *The Origins and Development of Financial Markets and Institutions*, edited by Atack, Jeremy, and Larry Neal, 161–208. Cambridge University Press.

Frey, Dennis. 2000. Industrious Households: Survival Strategies of Artisans in a Southwest German Town during the Eighteenth and Early Nineteenth Centuries. *International Review of Social History* 45: 115–135.

Freyberg, Max. 1874. *Geschichte der ehemaligen Hofmark Jetzenberg bis 1848*. Wolf und Sohn.

Ghysels, Eric, Alberto Plazzi, Rossen Valkanov, and Walter Torous. 2013. Chapter 9 - Forecasting Real Estate Prices. In *Handbook of Economic Forecasting*, edited by Elliott, Graham, and Allan Timmermann, 509–580. Elsevier.

Giersch, Robert, Andreas Schlunk, and Bertold von Haller. 2006. *Burgen und Herrensitze in der Nürnberger Landschaft*. Altnürnberger Landschaft.

Giersch, Robert, Andreas Schlunk, and Berthold Haller von Hallerstein. 2007. *Burgen und Herrensitze in der Nürnberger Landschaft*. Altnürnberger Landschaft Verlag.

Gillissen, Manfred. 2016. *Geschichte der Koblenzer Profanbauten*. Stadtarchiv Koblenz.

Gindeley, Anton. 1868. *Geschichte der Boehmischen Finanzen von 1526 bis 1618*. Denkschriften der Kaiserlichen Akademie der Wissenschaften.

Goldsmith, Raymond. 1985. *Comparative National Balance Sheets: A Study of Twenty Countries, 1688-1979*. University of Chicago Press.

Gomme, Paul, B. Ravikumar, and Peter Rupert. 2011. The return to capital and the business cycle. *Review of Economic Dynamics* 14: 262–278.

Gorton, Gary. 2017. The History and Economics of Safe Assets. *Annual Review of Economics* 9: 547–586.

Grillmeyer, Siegfried. 2005. *Habsburgs Diener in Post und Politik*. Philipp von Zabern.

Grimm, Alois. 2001. *Aschaffenburger Haeuserbuch, Vols. I-VI*. Geschichts- und Kunstverein Aschaffenburg e.V.

Gu, Shihao, Bryan Kelly, and Dacheng Xiu. 2020. Empirical Asset Pricing via Machine Learning. *Review of Financial Studies* 33: 2223–2273.

Hamberger, Edwin. 1988. *Das Fuerstliche Landschloss zu Rimpar im 17. und 18. Jahrhundert*. Historischer Verein Schweinfurt.

Hartung, Bernhard. 1861. *Die Häuser-Chronik der Stadt Erfurt*. Hennings und Hopf.

Hildebrandt, Ludwig. 2007. Neues zur Geschichte der Burg Streichenberg bei Stebbach. *Jahrbuch des Heimatverein Kraichgau* 20: 53–72.

Hoffman, Philip T., David Jacks, P.A. Levin, and Peter H. Lindert. 2002. Real inequality in Europe since 1500. *Journal of Economic History* 62: 322–355.

Hoffmann, Walther G. 1965. *Das Wachstum der deutschen Wirtschaft seit der Mitte des 19. Jahrhunderts*. Springer Verlag.

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

Isenmann, Eberhard. 2012. *Die deutsche Stadt im Mittelalter 1150-1550, 4th edition*. Klett Cotta.

Jorda, Oscar, Katharina Knoll, Dimitri Kuvshinov, Moritz Schularick, and Alan Taylor. 2019. The Rate of Return on Everything. *Quarterly Journal of Economics* 134(3): 1225–1298.

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.

Kain, John F., and John Quigley. 1970. Measuring the Value of Housing Quality. *Journal of the American Statistical Association* 65(330): 532–548.

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.

Kaspar, Fred. 1985. *Bauen und Wohnen in einer alten Hansestadt. Zur Nutzung von Wohnbauten zwischen dem 16. und 19. Jahrhundert, dargestellt am Beispiel der Stadt Lemgo*. Aschendorff.

Keneshloo, Yaser, Tian Shi, Naren Ramakrishnan, and Chandan Reddy. 2020. Deep Reinforcement Learning for Sequence-to-Sequence Models. *IEEE Transactions on Neural Networks and Learning Systems* 37(7): 2469–2489.

Kindleberger, Charles. 1991. The Economic Crisis of 1619 to 1623. *The Journal of Economic History* 51(1): 149–175.

Knoll, Katharina, Moritz Schularick, and Thomas Steger. 2017. No price like home: Global house prices, 1870–2012. *American Economic Review* 107(2): 331–352.

Koch, Philip, Viktor Stojkoski, and César A. Hidalgo. 2024. Augmenting the availability of historical GDP per capita estimates through machine learning. *PNAS* 121(39).

Korevaar, Matthijs. 2021. Reaching for yield and the housing market: Evidence from 18th-century Amsterdam. *Journal of Financial Economics* 148(3): 273–296.

Kuvshinov, Dimitry, and Kaspar Zimmermann. 2020. The Big Bang: Stock market capitalization in the long run. *Journal of Financial Economics* 145(2): 527–552.

Landau, Georg. 1832. *Die hessischen Ritterburgen und ihre Besitzer. 4 Vols.* Luckhard'sche Hofbuchhandlung.

Landesamt, Berlin-Brandenburg Statistisches. 2022. Zensus 2022. Statistisches Landesamt Berlin-Brandenburg, Basistabellen Berlin, Stichtag: 15. Mai 2022 .

Lehmann, Johann Georg. 1857. *Urkundliche Geschichte der Burgen und Bergschlösser in den ehemaligen Gauen, Grafschaften und Herrschaften der bayerischen Pfalz. 5 Vols.* Meuth.

Lichtenberger, Elisabeth. 1977. *Die Wiener Altstadt. Von der mittelalterlichen Buergerstadt zur City*. Franz Deuticke.

Lim, Bryan, Sercan Arık, Nicolas Loeff, and Tomas Pfister. 2021. Temporal Fusion Transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting* 37(4): 1748–1764.

Luzzatto, Gino. 1929. *I Prestiti della Repubblica di Venezia, Sec. XIII – XV*. R. Accademia Dei Lincei, Libreria Editrice A. Draghi.

Lyons, Ryan, Rowena Gray, Allison Shertzer, and David Agoratsos. 2024. The Price of Housing in the United States, 1890-2006. *NBER Working Paper No. 32593* .

Lüdicke, Reinhard. 1933. *Geschichte der Berliner Stadtgrundstücke seit der Einführung der Grundbücher Ende des 17. Jahrhunderts*. Historische Kommission Berlin-Brandenburg.

Marx, Magali, Benoît Mojon, and François R. Velde. 2021. Why have interest rates fallen far below the return on capital? *Journal of Monetary Economics* 124: 57–76.

Mayr, Anton. 1931. *Die Grossen Augsburger Vermoegen in der Zeit von 1618 bis 1717*. Selbstverlag der Stadt Augsburg.

McFarlane, K.B. 1957. The Investment of Sir John Fastolf's Profits of War. *Transactions of the Royal Historical Society* 7: 91–116.

Metrick, Andrew, and Paul Schmelzing. 2026. Banking Crisis Interventions Across Time and Space. *Review of Financial Studies* 39(4): 1054–1076.

Meyer, Johannes August. 1903. *Die wirtschaftlichen verhaltnisse des grund und bodens der stadt Giessen in den letzten 25 jahren.* Bonifacius Druckerei.

Mullainathan, Sendhil, and Jann Spiess. 2017. Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives* 31(2): 87–106.

Müller, Gustav. 1881. *Werth-Karte des Grundbesitzes in Berlin mit Rentabilitäts-Berechnung und Taxation der Gebäude und Grundstücke*. Alexius Kiessling.

Nather, Wilhelm. 2005. *Die Olmuetzer Haeuser-Chronik. 2 Teile*. Univerzita Palackeho v Olomouci.

Neal, Larry. 2018. The Variety of Financial Innovations in European War Finance during the Thirty Years' War (1618–1648). In *Financial Innovation and Resilience*, edited by Neal, Larry, 127–145. Springer.

Oesterreicher, Paul. 1830. *Geschichte der Burg und des Rittergutes Rabenstein*. Drausnick.

Paker, Meredith, Judy Stephenson, and Patrick Wallis. 2025. Predictive Modeling the Past. *LSE Economic History Working Paper* .

Pfister, Ulrich. 2020. Urban population in Germany, 1500–1850. *University of Münster CQE WP* .

Pfister, Ulrich. 2022. Economic growth in Germany, 1500–1850. *Journal of Economic History* 82(4): 1–42.

Piazzesi, Monika, and Martin Schneider. 2016. Chapter 19 – Housing and Macroeconomics. In *Handbook of Macroeconomics, Vol. 2*, edited by Taylor, John B., and Harald Uhlig, 1547–1640. Elsevier.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Harvard University Press.

Prak, Maarten. 1985. *Gezeten burgers: De elite in een Hollandse stad, Leiden 1700–1780*. De Bataafsche Leeuw.

Razum, Matthias, Harald Sack, and Gerhard Weilandt. 2023. Transferring the Nuremberg topographical and temporal model into the public (TRANSRAZ). *Abschlussbericht der Leibnizgesellschaft / IRS* .

Reis, Ricardo. 2022. Which r-star, public bonds or private investment? Measurement and policy implications. *LSE Working Paper* .

Renaud, Joseph Ritter. 1904. *Beiträge zur entwicklung der grundrente und wohnungsfrage in München*. Hirschfeld.

Richter, Jürgen. 1986. Der Niedergang der Reichserbtruchsessen von Waldburg-Friedberg-Scheer im 17. und 18. Jahrhundert. *Zeitschrift für Hohenzollerische Geschichte* 165–232.

Rogoff, Kenneth S., Barbara Rossi, and Paul Schmelzing. 2024. Long-Run Trends in Long-Maturity Real Rates, 1311-2022. *American Economic Review* 114(8): 2271–2307.

Rogoff, Kenneth S., and Paul Schmelzing. 2024. r-g before and after the Great Wars .

Ryckbosch, Wouter. 2012. *A consumer revolution under strain : consumption, wealth and status in eighteenth-century Aalst (Southern Netherlands)*. Dissertation University Groningen.

SAM. 2022. Statistisches Amt München (SMA) – Indikatorenatlas München, accessed November 1, 2024 .

Schaab, Karl Anton. 1841. *Geschichte der Stadt Mainz, 3 vols.* F. Kupferberg.

Schaefer, Friedrich. 1893. Wirtschafts- und Finanzgeschichte der Reichsstadt Überlingen am Bodensee in den Jahren 1550–1628. *Untersuchungen zur Deutschen Staats- und Rechtsgeschichte* 44: 1–196.

Scheiber, Artur Maria. 1948. *Häuserchronik der Marktgemeinde Strengberg (Bezirk Amstetten, Niederösterreich)*.

Schmelzing, Paul. 2026. *The Long Run: A New History of the International Financial System*. Yale University Press (forthcoming).

Schuette, Grete. 1930. Was geschieht fuer Frau und Kind in der Neuen Siedlung. In *Das Wohnungswesen der Stadt Frankfurt*, edited by Schoenberger, Guido, 158–168. Englert.

Schulz, Fritz Traugott. 1934. *Nürnbergs Bürgerhäuser und ihre Ausstattung. Das Milchmarktviertel*. Gerlach and Wiedling.

Schwarz, Klaus. 1968. Der Bremer Wohnungsmarkt um die Mitte des 18. Jahrhunderts. *VSWG* 55(2): 193–213.

Schwemmer, Wilhelm. 1970. *Die Bürgerhäuser der Nürnberger Altstadt aus reichsstädtischer Zeit. Erhaltener Bestand der Lorenzer Seite*. Nürnberger Forschungen.

Schötz, Hartmut. 2015. Das Schloss Lehrberg, Landkreis Ansbach: Ehemaliges markgräfliches Lehen. *Frankenland* 242–247.

Shiller, Robert J. 2015. *Irrational Exuberance. 3rd edition*. Princeton University Press.

Siegel, Jeremy. 1992. The Equity Premium: Stock and Bond Returns since 1802. *Financial Analysts Journal* 48(1): 28–38.

Solleder, Fridolin. 1938. *München im Mittelalter*. Oldenbourg.

Stahleder, Helmuth. 2006. *Aelteres Haeuserbuch der Stadt Muenchen. 2 vols.* Ph. C. Schmidt.

Taschner, Michael. 2013. Die Sieben Zeilen in Nürnberg – in vielerlei Hinsicht für die Bauforschung von Bedeutung. In *Nürnberger Altstadtberichte Nr. 38/2013*, 49–68. Altstadtfreunde Nürnberg e.V. – Osterchrist Druck und Medien GmbH.

Voigt, Paul. 1901. *Grundrente und Wohnungsfrage in Berlin und seinen Vororten : eine Untersuchung ihrer Geschichte und ihres gegenwartigen Standes*. G. Fischer.

von Hippel, Wolfgang. 1978. Bevoelkerung und Wirtschaft im Zeitalter des Dreissigjaehrigen Krieges: das Beispiel Wuerttemberg. *Zeitschrift fuer Historische Forschung* 4: 413–448.

von Rambaldi, Karl. 1893. *Geschichte des Schlosses Eurasburg und seiner Besitzer*. Wolff und Sohn.

von Schack, Hans. 1896. *Beiträge zur Geschichte der Grafen und Herren von Schack*. Wilh. Baensch.

Wachter, Jessica, James Paron, and Max Miller. 2026. Sovereign default and the decline in interest rates. *Review of Financial Studies (forthcoming)* .

Weidenbacher, J. 1926. *Die Fuggerei in Augsburg. Die erste deutsche Kleinhaus-Stiftung*. Selbstverlag.

Wenderoth, Thomas. 2025. *Mietwohnen in der Stadt Verbreitung, Entwicklung und Typologie am Beispiel Frankens 1400–1800*. Volk Verlag.

White, Eugene. 2009. Lessons from the American Real Estate Boom and Bust of the 1920s. In *Housing and Mortgage Markets in Historical Perspective*, edited by White, Eugene, Kenneth Snowden, and Price Fishback, 115–158. University of Chicago Press.

Wilde, Manfred. 1999. *Häuserbuch der Stadt Wolfen, Kreis Bitterfeld, 1550-1990*. Degener and Co.

Wolter, Helmut. 2010. *Das Häuserbuch der Stadt Coburg, 1400-1945, 5 Vols.* Peter Morsbach Verlag.

Zwierlein, Cornel. 2021. *Prometheus Tamed: Fire, Security, and Modernities, 1400 to 1900*. Brill.

# APPENDIX

## 1. Details on data sources, and *Häuserbücher*

In Table A.1.1, I provide full details on the underlying sources used for the urban repeat sales observations across the thirteen municipalities in our German sample.

I also provide contextual details for each city – whether accompanying information such as house transfers and inheritances, details on tax assessment values (as opposed to sales prices), and quality improvements on the property level are available in each case.

Table A.1.2 presents the same evidence (selected) for the rural repeat sales sample.

Table A.1.1: Overview of sources for Housing Index, repeat sales – URBAN

| | # Source | period | Source type | information |
|---|---|---|---|---|
| Erfurt | Hartung (1861) | 1300-1700 | Secondary, cadaster | F |
| Stengberg | Scheiber (1948) | 1465-1860 | Secondary | F, N |
| Mainz | Schaab (1841) | 1604-1808 | Secondary | |
| Frankfurt | ISG | 1378-1845 | Primary, archival | F, N |
| Koblenz | Gillissen (2016) | 1408-1792 | Primary, archival | |
| Berlin | Lüdicke (1933) | 1681-1850 | Secondary, cadaster | |
| Coburg | Wolter (2010) | 1568-1865 | Primary, archival | N, A |
| Munich | Stahleder (2006) | 1368-1571 | Secondary, cadaster | F, N, Q |
| | Burgmaier and Schneider (1966) | 1572-1904 | Secondary, cadaster | F, N, A, Q |
| Freiburg | Flamm (1905) | 1572-1904 | Secondary | |
| Dessau | Brückner (1989) | 1516-1939 | Secondary, cadaster | Q |
| Nuremberg | Topo N | 1492-1910 | Primary, archival | F, N, A, Q |
| | Schulz (1934) | 1436-1916 | Secondary | |
| | Schwemmer (1970) | 1650-1904 | Secondary | F, N, A, Q |
| Aschaffenburg | Grimm (2001) | 1591-1928 | Secondary, cadaster | F, N, Q |
| Olomuc | Nather (2005) | 1515-1891 | Secondary, cadaster | A, F, N, Q |

*Notes:* The table reports underlying sources for repeat sales observations. For column 5, the following codes are used: F = contains explicit foreclosure events; N = contains additional information on non-sale transactions (such as inheritances or transfers within families); A = contains details on assessment values; Q = contains details on quality improvements. "ISG" = Institut für Stadtgeschichte, Frankfurt.

Tables A.2.1-2 provide the corresponding summary statistics for the house price and rent observations, by city and period. Specifically, Table A.2.2 reports details on the rental observations for the German sample. As emphasized, the rental observations are often – but by no means always – matching the property level price observations. This is in line with the majority of existing literature that has typically abstained from such property level matches (an exception being the recent Chambers et al. (2021)).

Table A.1.2: Overview of sources for Housing Index, repeat sales – RURAL

| | # Source | period | Source type | information |
|---|---|---|---|---|
| Hessian estates, 13x | Landau (1832) | 1322-1772 | Secondary, cadaster | C, E |
| Jeztendorf | Freyberg (1874) | 1622-1848 | Secondary | V |
| Amlishagen | Bürger (1992) | 1369-1821 | Secondary | C, V, E |
| Starhemberg | Becker (1888) | 1410-1621 | Primary, archival | C, E |
| Nuremberg region, 29x | Giersch et al. (2006) | 1311-1907 | Primary, archival | C, V, E |
| Streichenberg | Hildebrandt (2007) | 1391-1596 | Primary, archival | C |
| Prillwitz, Kloxin, 3x | von Schack (1896) | 1737-1843 | Primary, archival | V, E |
| Waldburg-Scheer | Richter (1986) | 1446-1785 | Primary, archival | C,V, E |
| Eurasburg | von Rambaldi (1893) | 1609-1770 | Primary, archival | C |
| Lehrberg | Schötz (2015) | 1828-1934 | Primary, archival | C, V |
| Palatian estates, 10x | Lehmann (1857) | 1347-1934 | Secondary | C, V, E |
| Pretzien farms, 26x | Ballerstedt (2024) | 1585-1930 | Secondary | V, E |

*Notes:* The table reports selected underlying sources for rural repeat sales observations. For column 5, the following codes are used: C = castle; V = village; E = significant land estates.

## 2. Details on cost estimates: maintenance, taxes, vacancies

Here I provide details about the real estate cost adjustments, which determine the adjustment of gross returns: with some exceptions, even literature covering recent data relies more on uniform estimates, in light of the unobserved nature of internal rates of return. That said, all existing expenditure estimates – including those covering earlier markets – arrive at broadly comparable benchmark shares.

Table A.3 begins by presenting estimates of total expenditures associated with residential housing, as a share of either revenues or gross rent (GR). We observe that literature dealing with 18th to 20th century data has operated with relatively similar assumptions on this expenditure component across advanced economy cities, with shares ranging from 21-35% of rental income. Generally, this literature is explicit about the absence of any apparent time trends in expenditure shares (Jorda et al., 2019): there is no evidence to the contrary in the sources I evaluated, so that we maintain this assumption for the long horizon, too (including for the tax assumptions, as discussed in section 4.4.1).

What all recent literature has in common is the exclusive focus on urban real estate, including the associated expenditure estimates. Table A.3 also includes data in earlier literature for rural German expenditure shares: again, there is no indication of clear time trends, and given the general proximity of rural estimates to the urban data, I apply a uniform expenditure adjustment to urban and rural gross returns.

Specifically, recall that a uniform rate of 33% on gross rental yields is applied to adjust for all expenditure components – in addition to a uniform 0.15 percentage point tax deduction from the gross rental yield. Figure A.1 displays the resulting net real total return estimates for Germany, over the period 1466-1866. Here I adjust the gross figures presented in the main part of the paper with housing expenditure estimates, as detailed in section 4.4 of the main paper. Recall that these series incorporate depreciation in the baseline adjustment, include tax estimates, and are inflation adjusted with the new CPI series that include rent components in Schmelzing (2026).

Table 2.1: Summary Statistics for Housing Index (RHPI), urban repeat sales

| | # properties | # transactions | Property-years | city size |
|---|---|---|---|---|
| Erfurt | | | | |
| 1300-1670 | 5 | 6 | 564 | med |
| Strengberg | | | | |
| 1465-1860 | 17 | 108 | 4,240 | sma |
| Mainz | | | | |
| 1604-1808 | 2 | 5 | 140 | med |
| Frankfurt | | | | |
| 1358-1864 | 27 | 146 | 6,280 | lar |
| Koblenz | | | | |
| 1408-1806 | 18 | 58 | 2,742 | med |
| Berlin | | | | |
| 1658-1858 | 170 | 1,351 | 18,185 | lar |
| Munich | | | | |
| 1391-1904 | 211 | 2,299 | 49,474 | lar |
| Freiburg | | | | |
| 1572-1904 | 3 | 3 | 55 | med |
| Dessau | | | | |
| 1516-1939 | 156 | 741 | 20,083 | med |
| Nuremberg | | | | |
| 1387-1943 | 121 | 962 | 30,851 | lar |
| Aschaffenburg | | | | |
| 1591-1928 | 22 | 65 | 2,133 | med |
| Coburg | | | | |
| 1568-1865 | 26 | 164 | 4,541 | med |
| Olomuc | | | | |
| 1515-1891 | 62 | 1,053 | 17,057 | med |
| Total | | | | |
| 1300-1943 | 909 | 7,361 | 163,567 | — |

*Notes:* The table reports the sample composition of the repeat-sales observations for urban residential real estate in Germany. Property-years reports the sum of all property level spans of years between first and final sales transaction. City sizes are grouped into three categories: '"sma" (small cities averaging under 50,000 population across sample period), "med" (medium size cities averaging between 50,000 and 80,000 population across sample period), "lar" (large size cities averaging over 80,000 population across sample period).

Table A.2.2: Summary Statistics for rental yield observations

| | # properties | # rental obs | Property-years | av yd |
|---|---|---|---|---|
| Basel | | | | |
| 1352-1566 | 5 | 6 | 215 | 7.1 |
| Freiburg | | | | |
| 1407-1564 | 17 | 94 | 4,240 | 9.1 |
| Konstanz | | | | |
| 1604-1808 | 2 | 5 | 140 | 6.5 |
| Munich | | | | |
| 1401-1529 | 4 | 4 | 65 | 3.0 |
| Warendorf | | | | |
| 1772-1863 | 7 | 100 | 186 | 7.9 |
| Nuremberg | | | | |
| 1529-1805 | 31 | 32 | 40 | 4.6 |
| Bremen | | | | |
| 1650-1765 | 2,435 | 2,438 | 38,960 | 2.9 |
| Frankfurt | | | | |
| 1378-1902 | 31 | 40 | 49 | 10.9 |
| Aschaffenburg | | | | |
| 1691-1847 | 7 | 7 | 69 | 3.4 |
| Berlin | | | | |
| 1754-2019 | >500 | 29 | n/a | 4.6 |
| Dessau | | | | |
| 1661-1768 | 3 | 4 | 8 | 7.8 |
| Total | | | | |
| 1352-1863 | 2,534 | 2,713 | 43,849 | 7.4 |

*Notes:* Rental yields are expressed as a percentage of sales or assessment price throughout, taking the price observation that is closest to the nominal rent observation in cases where no direct matching is available in the underlying source itself. Bremen data records city-wide averages across all rental properties assessed via Schwarz (1968). Munich also includes Strassburg and Augsburg data; Frankfurt includes Giessen data. Nuremberg data is partially from Wenderoth (2025). Column 5 records the average gross rental yield in the respective sample, in nominal terms, as a percentage of the house market value (the total number being the arithmetic average of city levels). The modern Berlin data is based on census statistics via Bundesamt (2023), which does not provide the underlying sample size for the aggregate figures – hence we designated a "n/a" value to column 4 for Berlin.

Table A.3: Overview of housing expenditure share estimates

| | Source | period | EXP share | coverage |
|---|---|---|---|---|
| Urban shares | | | | |
| England | Allen (1988) | 1600-1646 | 7.4% | M+T |
| England | Earle (1989) | early 1700s | up to 40% | M+T |
| Paris | Eichholtz et al. (2021) | 1809-1943 | 25% | M+C+D |
| Berlin | Müller (1881) | 1870s | 20.7% | M+C+I+T |
| Amsterdam | Eichholtz et al. (2021) | 1900-1979 | 31.8% | M+C+D |
| U.K. | Chambers et al. (2021) | 1901-1983 | 32% | M+C+T |
| U.S. | Eisfeldt and Demers (2022) | 1986-2014 | 35% | M+T+I+R+V |
| DMs | Jorda et al. (2019) | 1870-2018 | "About one-third" | M+I+C+D |
| Rural shares | | | | |
| Rimpar estates | Hamberger (1988) | 1698-1803 | 1.0 p.p. of GR | R |
| Taxis estates | Grillmeyer (2005) | 1830s | 3.0 p.p. of GR | M+C+I |
| Elberfeld | Eberstadt (1903) | 1891-1901 | 32.6% | M+C+I |

*Notes:* The table reports estimates for the tax and non-tax housing expenditure share for various geographies and historical periods in secondary literature, together with details. The EXP share in column 4 is expressed as a percentage of the gross rent. Column 5 denotes the components included in the expenditure share by the respective authors: "M" = Maintenance; "I" = Insurance; "T" = Taxes; "C" = Vacancy. "GR" = Gross Returns. "R" = Repairs.

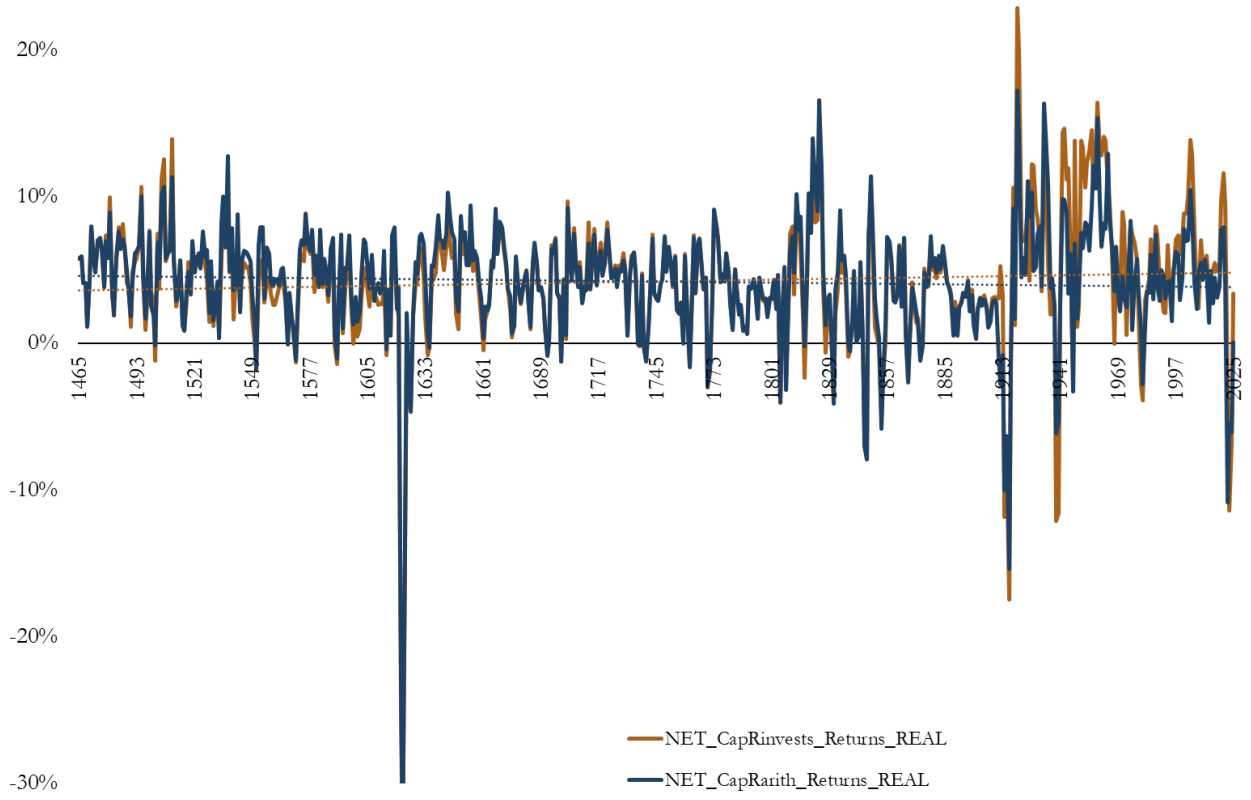
## 3. Comparison to existing indices, variations of weightings and inflation bases

How does my new German price index compare to existing work? Via Table A.4, I report correlation statistics for the new German House Price Index (nominal year-on-year change) to other historical house price indices in recent literature. It is of course not to be expected that very high degrees of correlation should exist – but such measures at least provide a first indication to what extent our results could be representative for broader samples of advanced economies over longer periods of time. Per the table, we generally observe low but positive correlation values when comparing the GHPI nominal changes to historical indices for Paris, Amsterdam, and Germany.

Prominent repeat sales house price indices are constructed on a price-weighted basis, including the prominent Case-Shiller index for U.S. cities (Corelogic, 2024). For the Case-Shiller index, in the words of the creators follows the following approach (Corelogic, 2024): "The composite home price indices are analogous to a cap-weighted equity index, where the aggregate value of housing stock represents the total capitalization of all of the metro areas included in the composite. The numerator of the previous formula is an estimate of the aggregate value of housing stock for all metro areas in a composite index".

While such a methodology has a different set of potential drawbacks, it could be relevant to test whether high level differences exist to the arithmetically weighted index presented in the main part of the paper. In this subsection, therefore, I present the main German index on such a comparable price-weighted basis – with an index rebasing every year based on shifting value weights. Each individual home enters the aggregate index sample with its own weight – that is, there is no arithmetic weighting on the city level – thus also analogous to the Case-Shiller methodology. Interpolated home values are used for all annual

Figure A.1: Total Net Real Returns on Capital, alternative weightings, 1465-2025.




*Notes:* The picture displays German net real returns on capital, across two weighting methodologies. The dark blue line displays the net return series based on the arithmetic weighting of the two key asset classes (real estate and sovereign bonds); the orange line displays the net return series based on the rolling investable capital stock shares (as detailed in section 3 of the main paper).

observations in between actual transactions.

## .1 Inflation variations for RHPI

What is the effect of alternative inflation adjustments on both our housing-specific results, and on the aggregate capital return features? The benchmark German Real House Price Index (RHPI) adjusted nominal price changes with a lagged inflation series in Schmelzing (2026) that includes rent price changes. But there are at least two alternative inflation series in the literature that could unearth potential biases in the benchmark series. First, Allen (2001) constructed reference European inflation series that have been widely used in the economic history literature: his constructions focus on urban areas, are based on existing secondary literature, use dynamic consumption baskets, and are transparent about the fact that they *exclude* rent components.

Second, more recently Pfister (2022) has constructed a German inflation series to deflate German GDP data, spanning 1500-1855. His data seeks to capture non-urban areas, but equally excludes rent prices.

Figure A.2 displays the headline RHPI index across the three alternative inflation adjustments. What we observe is that the Allen (2001) and Schmelzing (2026) inflation adjustments are rather closely aligned over the long run, with the terminal value in the former (1910=59.2) just slightly above that of the latter

| Table A.4: Correlations new German house price index (GHPI) and returns, other indices | | | |
|---|---|---|---|
| | period | correlation | source |
| Total | | | |
| GHPI – Amsterdam, annual index change | 1689-1748 | .181 | Korevaar (2021) |
| GHPI – Paris, annual index change | 1810-1910 | .093 | Eichholtz et al. (2021) |
| GHPI – Germany, annual index change | 1871-1910 | .014 | Knoll et al. (2017) |
| GHPI – U.S., annual index change | 1891-1910 | .099 | Lyons et al. (2024) |
| GHPI – Paris, log total returns | 1810-1910 | .109 | Korevaar (2021) |

*Notes:* The table reports correlation statistics between the new German House Price Index (GHPI), using the gross urban-only basis, with several existing historical indices. All index correlations in nominal terms, except U.S. which reports U.S. real correlation with German real basis.

(1910=41.2). Visibly, the outlier is the Pfister (2022) inflation adjustment, recording a terminal value significantly above the two alternative series (1910=831.0).

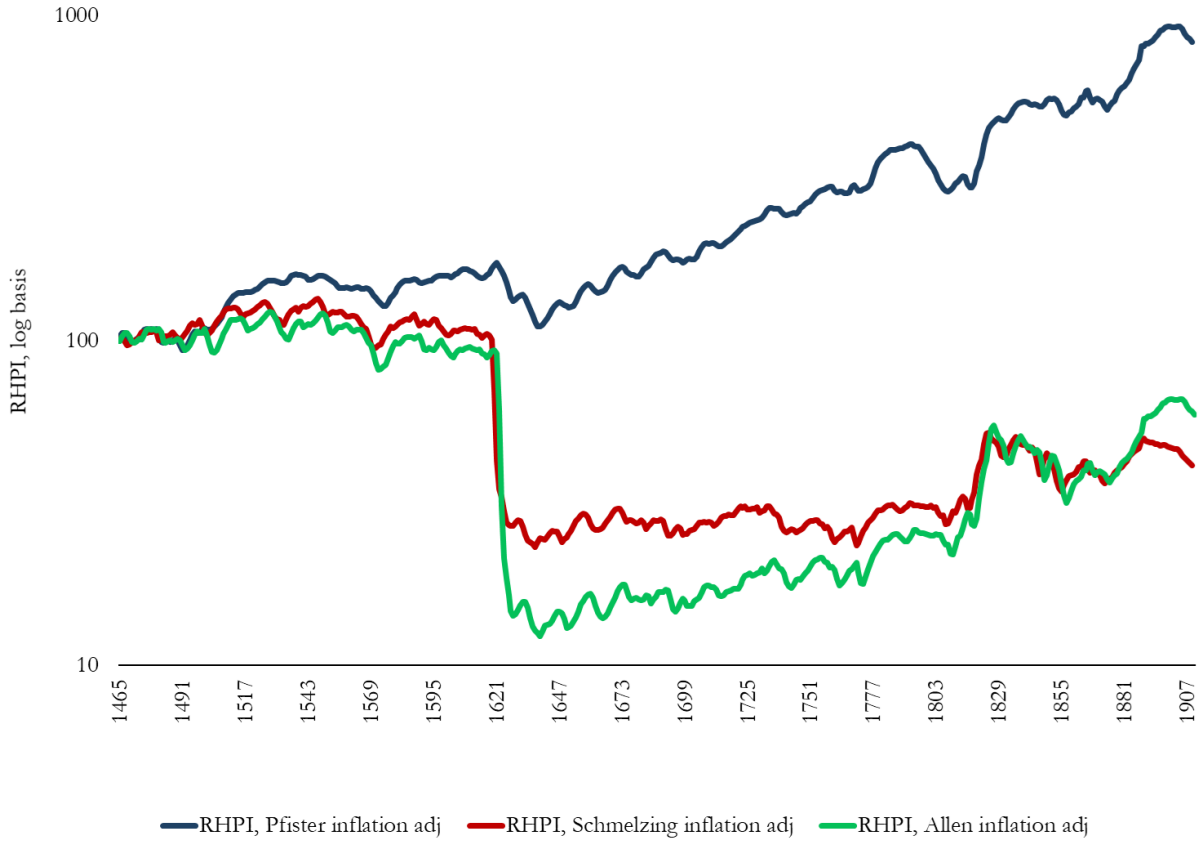
Where the series sharply diverge is above all the years around 1618, the outbreak of the Thirty Years' War. Conventionally, the years of 1618-23 are associated with the severe inflationary shock known as the "Kipper- and Wipper Crisis", initiated by the aggressive currency debasements upon war outbreak. The average inflation rate in Pfister (2022) over 1618-23 stands at -0.9%, while the average for Schmelzing (2026) stands at 15.4%, and for Allen (2001) at 13.7%. It is mainly this extreme index correction over 1618-23 that then carries over into the eventual sharp discrepancies for the terminal values observed for the Pfister (2022) adjusted RHPI.

Choosing an inflation series that *does* reflect the sharp inflation spike over 1618-22 appears to be the much better choice, against widespread evidence of actual price surges on the household and business level at the time, and inflation data being echoed in related time series, such as currency statistics. Detailed discussions in the economics literature are provided by Kindleberger (1991) and Neal (2018), and new currency debasement data also in Karaman et al. (2020), all of whom emphasize the sharp inflation spikes and provide associated data. I therefore acknowledge here that a uniform choice in favor of the Pfister (2022) data would change some of the results regarding total return slopes as well as those for aggregate capital returns – mainly in a reinforcing direction.[59] There is, however, little convincing reason to opt for such a choice. Meanwhile, opting for the Allen (2001) inflation series as our benchmark would confirm all benchmark results (stationarity, structural breaks, slopes) with only minimal differences.

Figure A.3 meanwhile compares the value-weighted to the arithmetically weighted German RHPI, log scale. Here, value-weighted ("VW") refers to weighting the price change of each house in the sample by the absolute value of each house in the total house sample, while "AW" takes the same weight for each house in each year. We observe that the main differences in the two bases are apparent during the pre-1618 era – when the AW basis shows a moderate *upwards* trends, as opposed to the moderate downward trend of the baseline VW basis; and secondly, we note that the terminal index value for AW is moderately higher, standing at 2020=699, versus 2020=257 for VW. The post-1618 shape and average real price gains y-o-y,

[59]When using the Pfister (2022) inflation basis, the gross real estate total return series (Urban+rural basis) displays a 0.17 basis point slope over 1465-2025. The 2-asset (real estate and bonds) weighted capital real return series over 1465-2025 records a moderately negative -1.12 basis point per annum slope.

Figure A.2: Inflation variations for value-weighted German RHPI, 1465-1910.




*Notes:* The Figure displays three variations of the value-weighted real housing index, taking 1465=100 (log basis). The series differ only in regards to the inflation adjustment chosen: first, we observe the RHPI adjusted with the Pfister (2022) inflation series (in blue); second the RHPI with the Schmelzing (2026) inflation series (in red); third the RHPI with the Allen (2001) inflation adjustment (in green).

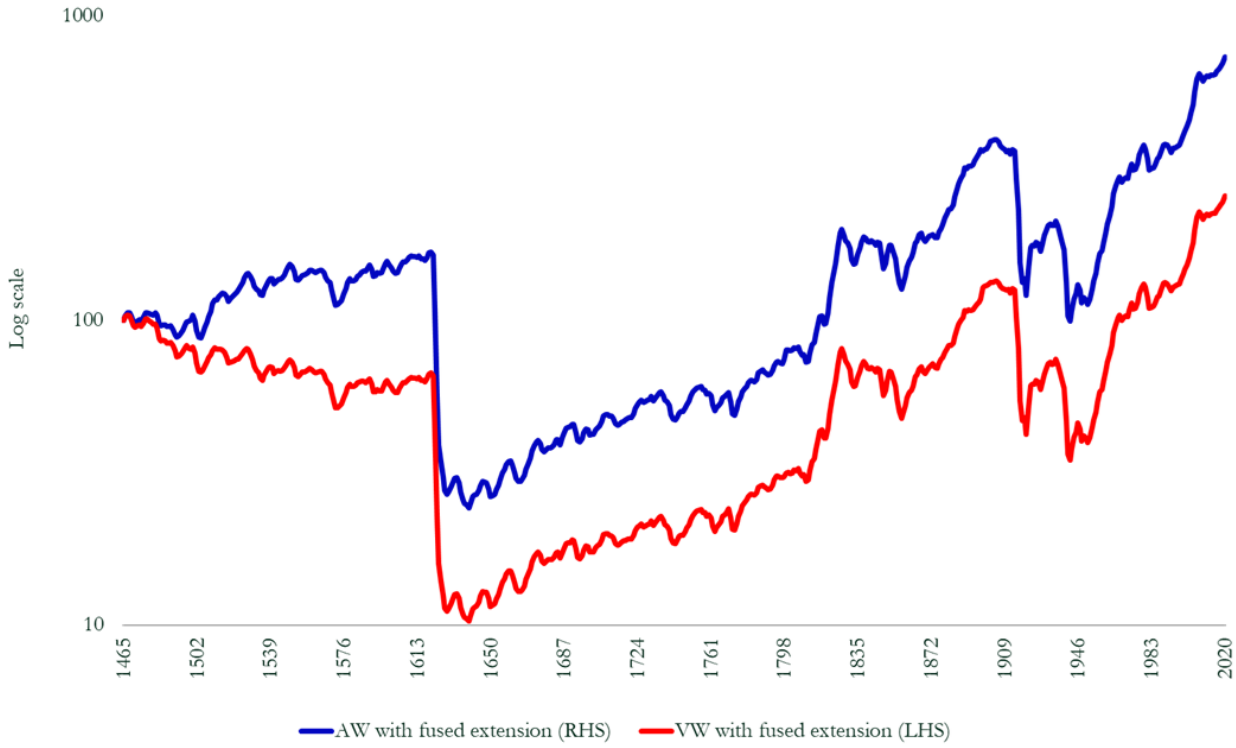
however, are closely aligned on both indices.

Overall, these exercises illustrate that the ML process generates time series within plausible ranges, as benchmarked against existing primary source approaches in the literature. Of course, there are multiple conceptual reasons why we should not expect any significant overlap in these indices: for instance, the Herengracht index measures a street-specific repeat sales sample – while the ML index is based off the country-wide "JST" reconstruction for Holland. In any case, the moderately positive correlation to existing city- and street-specific data is a plausible observation, especially in light of agreement on key (visual) inflection points.

## .2 Urban vs Rural trends – Repeat Sales of Castles, Villages, and Estates

Thus far, existing literature on both historical and contemporary housing focuses near-exclusively on *urban* repeat sales, in line with benchmark indices for the U.S. and elsewhere, the Case-Shiller index for instance taking the 20 leading U.S. city averages. Eichholtz et al. (2021), Lyons et al. (2024), and Knoll et al. (2017) are all examples for this exclusive urban focus. Chambers et al. (2021) are a notable exception for including

Figure A.3: AW versus VW German RHPI, Urban, 1465-2020.




*Notes:* The Figure displays value-weighted (VW) versus the arithmetically weighted (AW) German Urban RHPI, on log scale, and both indexed to 1465=100, and both adjusted with the same seven-year progressively lagged realized inflation figure for Germany.

agricultural returns in their U.K. data, though they only do so for the period of 1901-1983.

Housing being a bundle of structures and land, non-urban data can not least capture more properly the meaningful share of agricultural land in early capital stocks. Additional biases are known to exist even when purely considering within-urban data ("Superstar cities"). Such biases will also be corrected somewhat when incorporating rural data, especially when very long horizons with greater rural shares come into play. I addressed some of the "survivorship bias" concerns by including small cities that can be regarded as "losing out" from the secular urbanization process affecting Western economies since the 15th and 16th centuries.

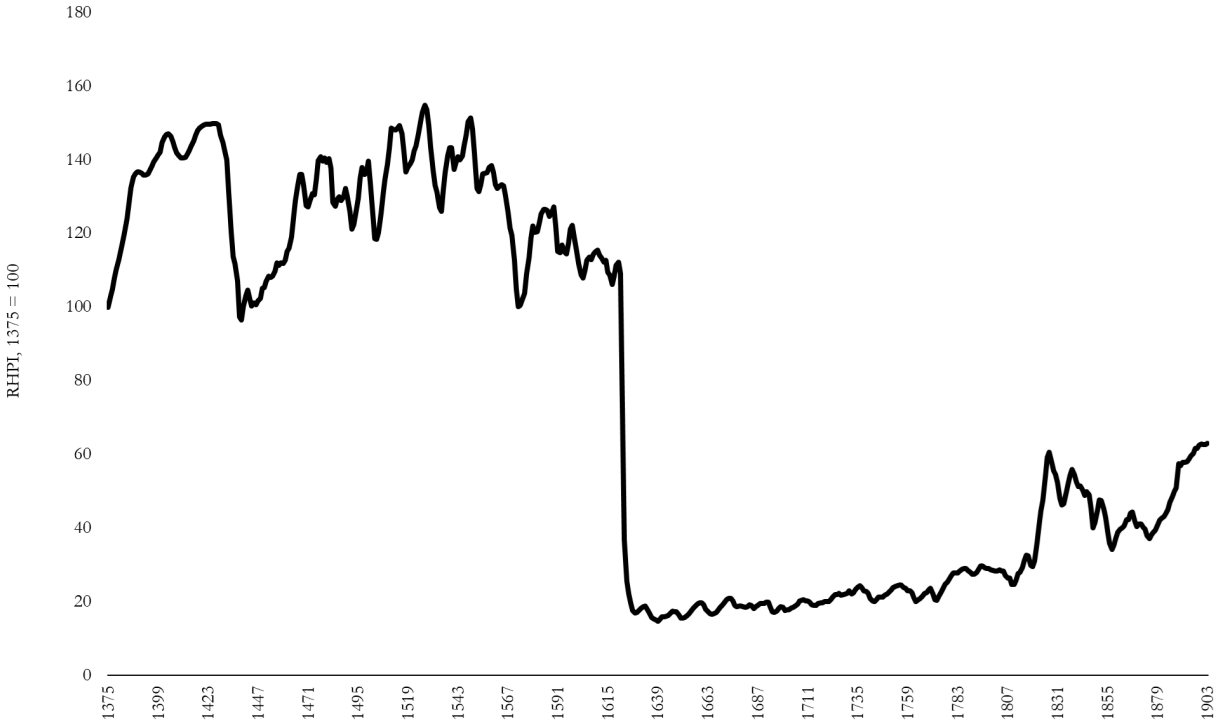
A more dynamic approach to adjust for the strong urbanization trend itself is now to formally reconstruct repeat sales for rural areas, estates, and villages. In a second step, trends in urbanization rates can then be utilized to weight the urban and new rural sub-indices over time. Urbanization rates for Germany released by Pfister (2020) can be utilized. This is done in this sub-section, which assembles and aggregates such rural transactions in a methodologically identical manner. According to Pfister (2020), urbanization rates fluctuated around a very low ratio just around 10% over the period 1465-1800, before accelerating more meaningfully: over the course of the 19th century, the ratio in fact triples to reach a terminal 28% by 1904. Figure A.5, left panel, displays the full timeline that is used (after 1910, when the cross over to "JST" series is undertaken, we retain an urbanization value of 100%).

Figure A.4 now displays the "Rural RHPI" individually, composed on the basis of repeat sales for 103 estates / farms / castles over 1375-1934. The sources include – among numerous estate-level publications – collective volumes such as von Schack (1896), Oesterreicher (1830), Lehmann (1857), Giersch et al. (2007) for

a comprehensive project on Franconian estates, as well as Becker (1889) for various Austrian estates (also see Table A.1.2). Combined, the 103 estates are associated with 442 repeat sales and 21,607 property-years. A total of 133 separate annual rural yield observations have been identified to construct rural total returns.

In Figure A.4, we observe that the rural RHPI, too, experienced a dramatic shock at the outbreak of the Thirty Years' War. Unlike the urban RHPI, however, we see that it never subsequently recovered the pre-1618 high watermarks, and in fact remained severely depressed until into the 20th century. The narrative sources frequently mention the dramatic physical destruction of these rural properties – a significant share of which were abandoned and turned into permanent ruins after 1648. This is despite the fact that the median rural estate even after 1618 remains more than twice as expensive in absolute terms than the median urban property in the sample (13,200 Gulden versus 6,085 Gulden post-1618). At the same time, this ratio stands at more than 4.5x in the pre-1618 era (rural estates median: 3,280 Gulden, urban properties median: 780 Gulden).

Figure A.4: Rural RHPI – Castles, Estates, Farms, Villages – 1375=100.




*Notes:* The Figure shows a Rural RHPI for Germany composed of repeat sales for 103 Castles, Estates, Farms, and Villages, indexed to 1375=100, and covering a total of 21,607 property-years. Lagged inflation adjustment using Schmelzing (2026), as per other benchmark indices. For details see text.

## .3 Size-adjustments for housing

All existing relevant housing indices continue to acknowledge the challenges with regards to quality changes in housing – and that remains true for both "hedonic" and "repeat sales" constructions of such indices. That is, it might well be that much if not most of the residual positive index changes over time could still capture quality improvements. Representative for most of the existing literature, Knoll et al. (2017) state that:

- "researchers using our dataset in the future should take into account that accurate measurement of

quality-adjustments remains a challenge."

These authors – in line with preceding work – undertake no specific quality adjustments for price indices in the early part of their sample, and instead splice early observations with quality adjusted series in the second half of the sample. My basis is therefore consistent in that no additional quality adjustments (beyond choosing the repeat sales methodology in the first place) are made: conceptually, that is of no concern for any of the preceding or subsequent discussions, though of course it would be *desirable* to know precisely the underlying decomposition in quality-related price changes and the residual components.

Yet, we can at least sketch the general empirical contours and focus on one recognized major quality component for which long-run reconstructions are feasible: specifically, floor space per capita.

Key contributions decomposing the "quality" input into residential housing have posited that the floor area per capita and the number of rooms constitute major ingredients for housing quality, for both renters and owner-occupiers – and indeed a statistically more significant one compared to factors such as neighborhood, schooling, or crime indicators (Kain and Quigley, 1970). This is in particular true for existing long-run studies which try to devise constant-quality indices, and which emphasized floor space per capita as a dominant input (Clark, 2002; Eichholtz et al., 2023).[60] Given that the *Häuserbücher* frequently inform us about house level parcel and floor sizing, we can derive some long-run statistics on representative quality changes on this key indicator: we know, for instance, that the Berlin houses in our sample run at an average 417 square meter parcel size, and know that in the two districts from which our houses are drawn (Dorotheenstadt and Friedrichstadt), persons per house are calculated at 14.9 for the year 1711, yielding a square meter living space figure of 28 per capita.[61] Overall, however, the data suggests that it is not implausible to assume an increase of representative floor space per person in German urban areas by a factor of 2-2.5x between the early 1500s and the early 21st century – from levels between 15-20 square meters in the era prior to the Thirty Years War, to contemporary averages of 38-40 square meters based on most recent census data for key German cities. That is an increase that falls far short of the increases in real property level house prices over the same period when using arithmetically weighted indices – however, interestingly, value-weighted indices suggest that living space growth could account for a not insignificant share of overall real price growth. Specifically, our value-weighted RHPI that fuses JST data post-1910 records a full-sample increase in real house prices of 3.9x over 1465-2020. To the extent that the sample over time captures "representative" housing samples, this suggests that living-space adjusted real house price growth on such a basis would only account for the residual 1-1.5x in real price growth.

Repeat-sales housing indices are considered as the best construction method to incorporate quality changes in housing assets – but they still remain subject to potential quality-related distortions. In general, the literature has settled on preferring repeat-sales indices to control for many quality-related changes in housing, but acknowledges that no perfect methodology exists to fully separate "true" housing index changes from housing quality changes. Indeed, some key literature dealing with long-run construction has

[60] As Eichholtz et al. (2023) state, "housing quality per capita [is] most relevant when discussing the evolution of the standard of living. Such concerns are relevant: Adam Smith explains the relative affordability of London housing as the result of mass-scale sub-letting of parts of dwelling homes (Smith, 1776). Smith's observation suggests that the number of people per house may have varied substantially across cities (and potentially over time). To address these concerns, we need to look at housing space per capita." And similarly, via Clark (2002), "to estimate housing rental values while controlling for quality, I have used the subset...of the observations where there are multiple observations on the same property at different times, and no indications of change in size or quality of the structure, or of change in lot".

[61] Earlier house and parcel sizes in Germany are typically measured in "Ruten" (1 Quadratrute = 14.1846 square meters) or "Fuss" (1 Fuss = 0.313 meters).

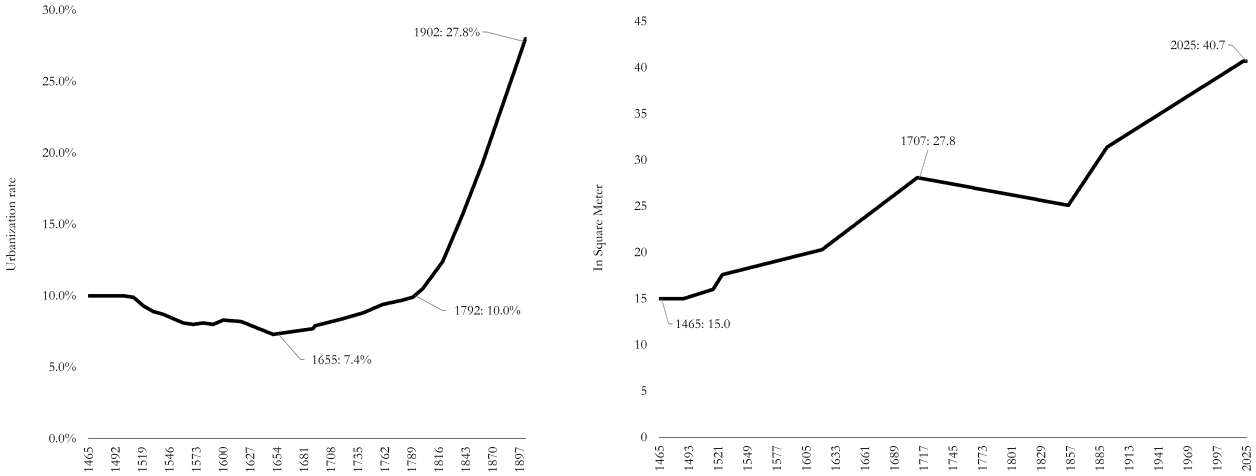
resorted to simply acknowledging the potentially distorting role of quality changes, but gone ahead with splicing quality-controlled with non-quality controlled sub-indices (Knoll et al., 2017).[62]

So it appears relevant to get a sense of the potential size of the quality effect, to the best extent possible. For this purpose, the following Table A.5 begins by systematically compiling available data for living space in early modern and modern German cities, converted to a per capita basis. I only include data points where authors refer to "typical" or "representative" housing units in the respective city, and have included precise space data.

Here we observe that there appears to have been a clear upwards trend in living space per capita in German urban areas over the very long run – perhaps an intuitive general observation. Our earliest data points indicate typical living space of around 15 square meter – a figure which then rises over the course of the early modern period, and by the eve of the Thirty Years War rose above 20 square meters for the first time. From there, another doubling of average living space takes place by the early 21st century, where current official city surveys for Berlin and Munich indicate living spaces per capita of 38-40 square meters.

This means that, indeed, some of our housing price appreciation is likely driven by house quality improvements related to size: given homes can be enlarged, either by adding "attached houses" on the same address, or through adding additional stories, say. That both effects took place is confirmed by anecdotal evidence in our cities. At the same time, at least on the basis of the limited number of concrete data points, it is not evident that such space dynamics themselves had any decisive role per se: assuming a doubling of living space for a representative housing inhabitant in our sample over half a millennium would translate into an annual increase of living space of just 0.2%.

Figure A.5: German urbanization rates and living space per capita square meters, 1465-2025.




*Notes:* The Chart displays on the left panel the evolution of German urbanization rates per Pfister (2020), which underlies the "urban+rural" weightings of the index, as well as on the right panel German living space per capita as derived from Table A.6 and the associated discussion. This factor underlies "size-adjusted" variations of the benchmark RHPI.

To provide some additional nuance to quality improvements, Figures A.6 and A.8 actually display the evolution of representative residential housing units that are featured in our sample. These are typical

[62] Knoll et al. (2017) acknowledge the fact that their indices may insufficiently control for quality dynamics, especially in the pre-1945 era, and write that "researchers using our dataset in the future should take into account that accurate measurement of quality-adjustments remains a challenge." For Germany, the paper claims that the pre-1945 data is non-quality controlled, while from 1962, their nationwide index controls for quality changes.

Table A.6: Average or representative living space estimates, German cities

| | Source | period | sqm per capita | type detail |
|---|---|---|---|---|
| | | Early Modern Era | | |
| Nuremberg | Dirlmeier (1978) | 1488 | 15 | representative M and LM |
| Augsburg | Weidenbacher (1926) | 1516 | 16 | representative LM |
| Lemgo | Kaspar (1985) | 1400-1650 | 17.6 | |
| Cologne | Eberstadt (1903) | ca. 1620 | 20.3 | "typical three-window house" |
| Berlin | Voigt (1901) | 1711 | 28.1 | representative UM and M |
| Frankfurt | Schuette (1930) | 1855 | 25.1 | representative M |
| Vienna | Lichtenberger (1977) | 1856 | 16.8 | weighted avg, inner city |
| | | Modern Era | | |
| Berlin | Berlin (1897) | 1890-5 | 31.4 | all-city average |
| Wolfen | Wilde (1999) | 1872-1908 | 27.5 | representative M and LM |
| Munich | Renaud (1904) | 1903 | 9.3 | all-city average |
| Giessen | Meyer (1903) | 1900 | 20.4 | representative M and LM |
| Frankfurt | Schuette (1930) | 1930 | 37.0 | Housing office Frankfurt |
| Berlin | Grundstuecks-MB | 1946 | 11.8 | all-city, census data |
| Berlin | Grundstuecks-MB | 1996 | 34.8 | all-city, census data |
| Berlin | Grundstuecks-MB | 2000 | 37.9 | all-city, census data |
| Berlin | Landesamt (2022) | 2022 | 40.7 | all-city, census data |
| Munich | SAM (2022) | 2022 | 38.4 | all-city, census data |

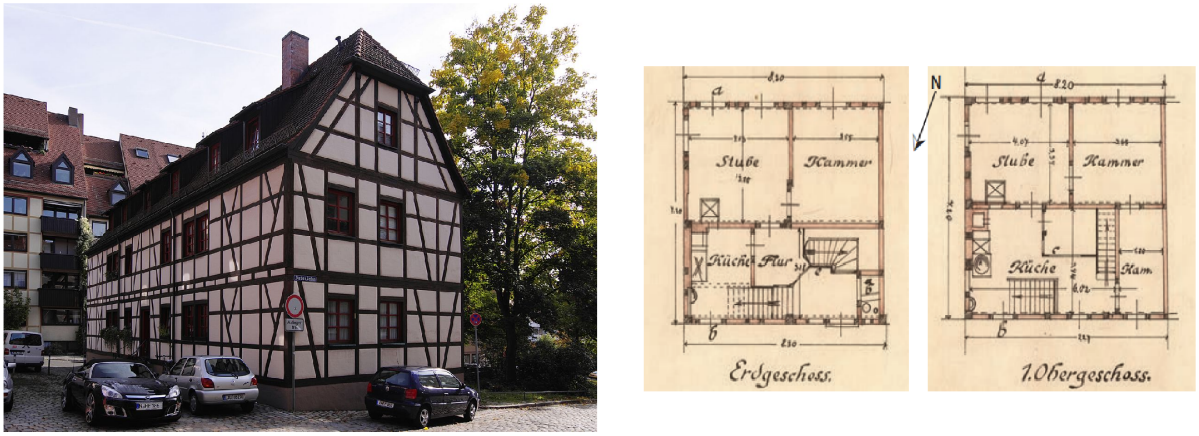
*Notes:* The table reports data for the residential living space per capita in contemporary and historical sources. Data prior to 1700 operates with average household size of 3.3 persons. um = upper middle class; M = middle class ; LM = lower middle class. The Vienna figure takes midpoints of the range in Lichtenberger (1977), and inner city population figure of 1834.

units behind our numbers over time and space. In the first, we exhibit a artisan building in Nuremberg, originally built in the year 1489 for newly arriving weavers from the Nuremberg region. The buildings have undergone an eventful history, but are still located at the same place in Nuremberg today, keeping the same address at the *Sieben Zeilen*. Thanks to in-depth studies by the Nuremberg Historical Association (Taschner, 2013) we have extensive detail about the precise quality adjustments and architectural details of the buildings.

The building contains two stories, with an additional roof area used for storage purposes, as well as a cellar containing work areas. Two households lived in such a building from the inception date. Per the architectural drawing displayed on the right handside of Figure A.6 – itself produced for a renovation project in the year 1912 – we know that the living space per story stands at 59 square meters (8.2 meters width, 7.2 meters depth, including staircases). Such dimensions are comparable with corresponding information from other German urban areas in our sample and beyond. Adjusting the living space for household size developments for urban middle class households yields several relevant data points for living space per capita.

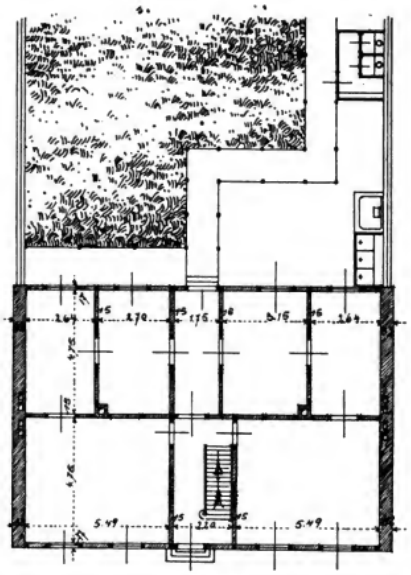
Next, Figure A.8 displays a representative residential street in Munich, shown in its condition in the year 1939. We observe that by the 20th century, residential houses in larger German cities feature four to

Figure A.6: Representative middle class residential house, built 1489, Nuremberg.




*Notes:* The Figure's left handside displays the exterior of a representative Nuremberg middle class home, located at the *Sieben Zeilen*; the left handside displays the corresponding architectural drawing, sourced via Taschner (2013).

Figure A.7: Representative lower middle class residential house, Rhineland, ca. 1848.

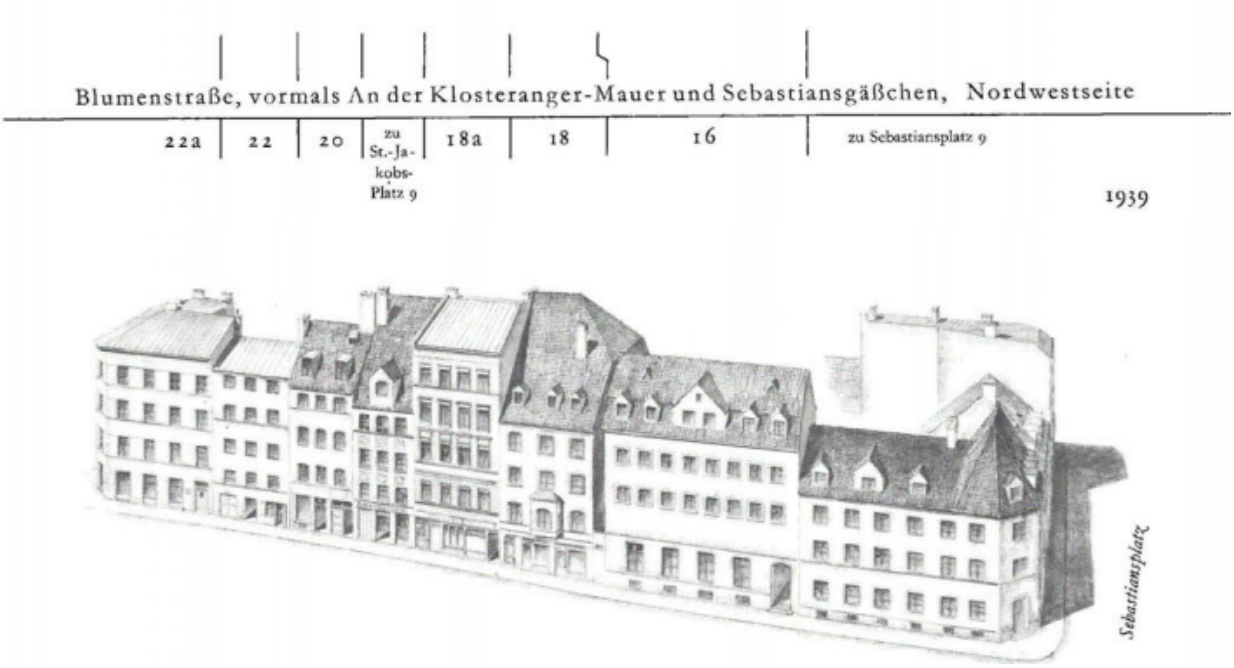



*Notes:* The architectural drawing from Eberstadt (1903) displays two representative middle class house types in the Rhineland, 1848. The house contains three storys, with two units per story. Displayed is the ground story, with the added backyard. The measurements per unit are 5.48 meters by 9.52 meters = 52.3 square meters of living space per unit.

five stories, including attics. At this time, one can assume three to four households per residential unit, that is 12-20 persons in a representative residential building, per Burgmaier and Schneider (1966). These residential units again contain unit-level heating, hot water, and kitchens. In the ground floors, we often find commercial real estate, and shared cellars are typically included below ground floors. The street length is about 80 metres, and the depth is 11 metres. Today, the buildings in the *Blumenstrasse* remain in place,

and can be located at
geoportal.muenchen.de/portal/plan – coordinates 11° 34 16 | 48° 07 55.

Figure A.8: Representative middle class residential street, 1939, Munich.




*Notes:* The architectural drawing from displays a typical street of residential houses in Munich, shown in their condition as of 1939 (*Blumenstrasse 16-22a*). Architectural drawing by Gustav Schneider in Burgmaier and Schneider (1966).

### .4 Source details – building costs

Next, I list the sources and methodology for my post-1870 building cost indices, by country. All post-1870s building cost figures used for the training data are sourced via respective national statistical agencies: note that I do not use the limited (post-1950) data available via Knoll et al. (2017). We recall generally, per Figure 8 of the main paper reporting variable importance, that building cost indices – several of which we newly construct in consistent ways from raw commodity data – exhibit some relevance for the machine-learning approaches presented. The early modern, pre-1870 German data is obtained as follows:

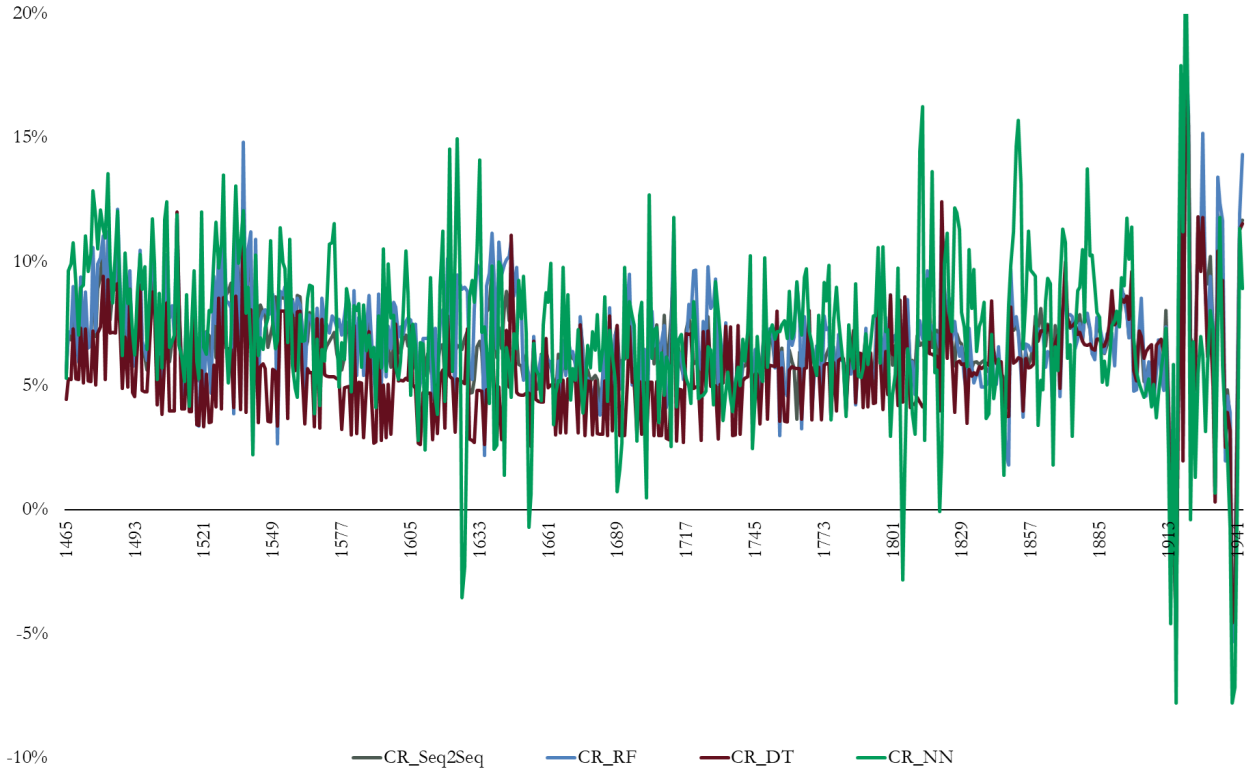
- Germany: I use the building cost index constructed by Pfister (2022), who reports index values on a quinquennial basis over 1500-1860 – we linearly interpolate annual frequencies.

## 4. Further details on aggregations

I provide further details on aggregate capital return components and variations in this section. The first additional group of components are the nominal total returns on capital (with the main section mainly focusing on the real bases): displayed in Figure A.9, we see nominal returns on capital, across the four leading bond total return models as reconstructed in section 7 of the main part. From a visual inspection, despite important nuances, all models appear to agree on key general outlines of capital returns – specifically,

a moderate long-horizon upwards trend, interspersed by low return periods such as the 17th century, and higher volatility towards the end of the observation period.

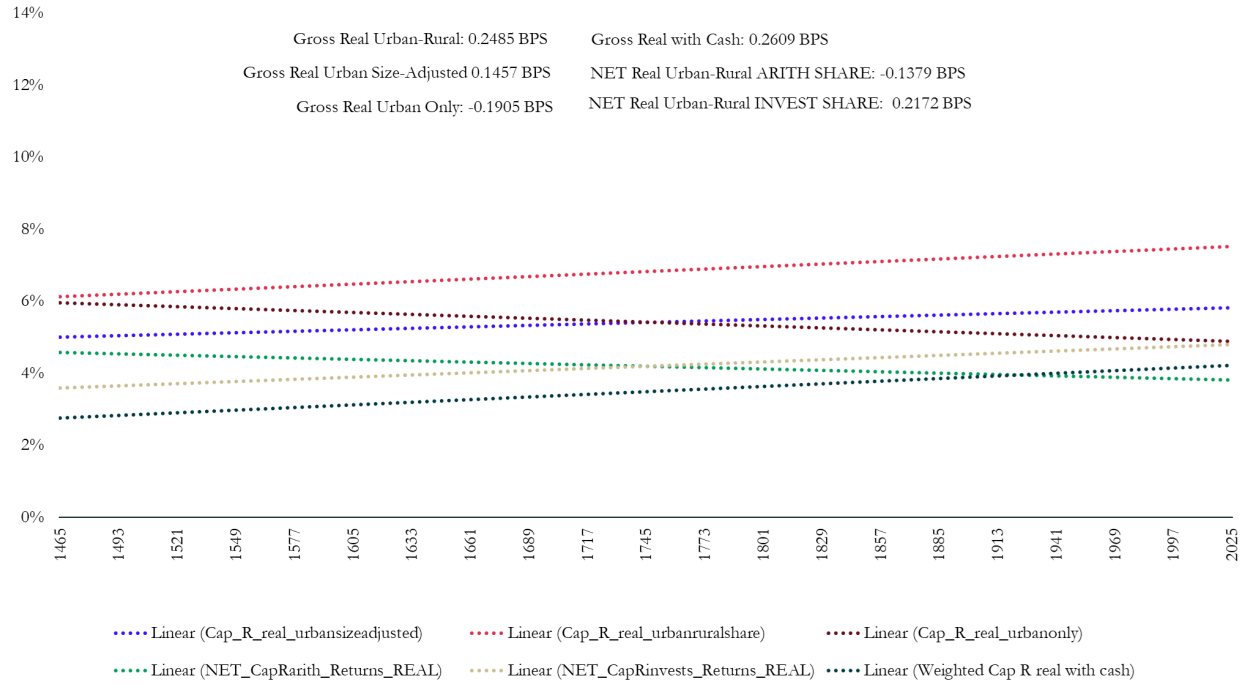
Figure A.9: Capital total returns, nominal, leading models, 1465-1943.




*Notes:* The Figure displays four reconstructions of nominal German weighted total returns on capital, weighting real estate and bond total returns arithmetically. The model variations refer to the bond component only, the real estate component remaining identical and primary source derived (rural+urban basis).

Next, Figure A.10 visualizes the linear trendlines for six variations of German real aggregate returns on capital over the years 1465-2025 – based on the variations we encountered in Table 3.2 of the main paper. The purpose is to illustrate the factors that influence the broader, stylized trends in capital returns over time. Across the six trends, four are trending upwards over time, between 0.15 and 0.26 basis points per annum: the steepest slope is measured for Gross Real Returns, investable shares, including cash. Two, meanwhile, trend downwards over time, between 0.14 and 0.19 basis points per annum. In this sense, we can conclude that *the majority* of capital return measures do suggest a long-horizon upwards trend, a trend that is robust to various modifications: but it is clearly not a universal observation, and slight – but plausible – modifications would have led a real investor to be exposed to widely different long-run cumulative returns.

Figure A.10: Capital Return Trends, Six German variations, 1465-2025.




*Notes:* The Figure displays the trendlines for six Real Return variations of German aggregate capital returns – per the details in Table 3.2 of the main paper. These iterations are non-log returns. The slopes are recorded in terms of annual change (in basis points, BPS) of the linear trendline.

# 5. Econometrics, further variations

In this section, I report additional econometric results and variations, to analyze the main paper results with regards to robustness, especially the main time series with regards to their stationarity (ADF-GLS) and structural break properties (Bai-Perron). Additionally, I undertake further variations of the correlation exercises discussed in section 9.3 of the main paper.

## .1 ADF-GLS variations

First, I assess further time series properties of key variables that were introduced and discussed in the main part of the paper (section 6), via Table A.7. Next, Table A.8 reports additional results for a standard ADF-GLS trend stationarity test, via Elliott et al. (1996), this time *not* assuming a time trend (unlike the basline table).

Across all these variations, we observe that housing total returns, as well as measures of aggregate capital returns and excess returns, all maintain clear trend stationarity, all at the 1% significance level.

| Table A.7: ADF-GLS, additional variations, with time trend | | | |
|---|---|---|---|
| | no of lags | t statistic | optimal lag |
| Housing TR, 1465-2025 | | | |
| Nom gross TR, log, 1465-2025 | 3 | -9.1310 | SIC, Seq |
| | 2 | -7.8541 | MAIC |
| | 1 | -10.764 | |
| Nom net TR, log, 1465-2025 | 3 | -9.2264 | SIC, Seq |
| | 2 | -7.9688 | MAIC |
| | 1 | -10.938 | |
| Bond TR, 1465-2025 | | | |
| Bond HTM, log, 1465-2025 | 3 | -3.1077 | MAIC, Seq |
| | 2 | -3.4964 | |
| | 1 | -3.6347 | SIC |
| Weighted Cap R, 1495-2025 | | | |
| Nom Cap R, arith weights | 3 | -8.4679 | |
| | 2 | -9.0084 | Seq, MAIC |
| | 1 | -10.820 | SIC |
| Real net Cap R, arith weights | 3 | -8.9463 | |
| | 2 | -9.6718 | MAIC |
| | 1 | -10.997 | SIC |
| Capital R components, 1465-2025 | | | |
| Yields, arith weights | 3 | -2.4531 | Seq, MAIC |
| | 2 | -2.7622 | |
| | 1 | -2.8388 | SIC |
| Cap gains, arith weights | 3 | -9.9261 | |
| | 2 | -10.475 | MAIC |
| | 1 | -12.366 | SIC, Seq |
| Excess Returns, 1465-2025 | | | |
| RP2 (ex cash) | 3 | -8.6907 | SIC, Seq |
| | 2 | -7.6620 | MAIC |
| | 1 | -10.429 | |

*Notes:* The table reports the ADF-GLS test statistic for several choices of the number of lags $k$ (with a maximum of three lags). The regression includes a constant. The test assumes a time trend (see appendix for ADF-GLS results without time trend). The critical values at the 1, 5, and 10 percent significance levels are the following: $-3.48$ (1 percent); $-2.87$ (5 percent); $-2.59$ (10 percent). "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.

Table A.8: ADF-GLS, benchmark series, without time trend

| | no of lags | t statistic | optimal lag |
|---|---|---|---|
| Real Estate, 1465-2025 | | | |
| Real gross housing log return, 1466-2025 | 3 | -5.9558 | |
| | 2 | -5.8473 | SIC, Seq, MAIC |
| | 1 | -8.0501 | |
| Real net housing log return, 1466-2025 | 3 | -7.9546 | Seq |
| | 2 | -7.5909 | SIC, MAIC |
| | 1 | -10.149 | |
| Bonds, 1465-2025 | | | |
| Real TR, log, 1465-2025 | 3 | -8.4537 | MAIC |
| | 2 | -9.9105 | |
| | 1 | -11.429 | SIC |
| Cap R, 1465-2025 | | | |
| Real Cap R, S weights | 3 | -9.9532 | SIC, Seq |
| | 2 | -8.9952 | MAIC |
| | 1 | -11.287 | |
| Components, 1465-2025 | | | |
| Yields, arith weights | 3 | -1.1076 | Seq, MAIC |
| | 2 | -1.4010 | |
| | 1 | -1.4735 | SIC |
| Cap gains, arith weights | 3 | -6.5112 | MAIC |
| | 2 | -7.2361 | SIC, Seq |
| | 1 | -8.9101 | |
| Excess Returns, 1465-2025 | | | |
| RP2 (ex cash) | 3 | -8.2565 | SIC, Seq |
| | 2 | -7.3283 | |
| | 1 | -9.9852 | MAIC |

*Notes:* The table reports the ADF-GLS test statistic for several choices of the number of lags $k$ (with a maximum of three lags). The regression includes a constant. The test assumes no time trend. The critical values at the 1, 5, and 10 percent significance levels are the following: $-2.58$ (1 percent); $-1.95$ (5 percent); $-1.62$ (10 percent). "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.

| Table A.9: Additional structural break tests, Bai-Perron basis | | |
|---|---|---|
| | no of breaks | break years |
| Real Estate, Germany, 1465-2025 | | |
| Gross nom TR, log | 2 | 1548, 1941 |
| Net nom TR, log | 2 | 1548, 1941 |
| Sovereign Bonds, variations, 1465-2025 | | |
| NN, nom TR (1465-2002) | 0 | – |
| RF, nom TR (1465-2002) | 3 | 1551, 1645, 1921 |
| DT, nom TR (1465-2002) | 0 | – |
| Real yields | 0 | – |
| Capital Return components, 1465-2025 | | |
| Yields, arith weight | 2 | 1558, 1898 |
| Capital gains, arith weight | 0 | – |

*Notes:* 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 linearly detrended series using a trimming parameter of 5 percent and a heteroskedasticity and autocorrelation consistent variance estimator. The significance level is 5 percent, and the maximum number of allowed break points is 5 for all series. When the test rejects the hypothesis of no breaks, the estimated number of breaks and the break dates are reported in the second and third columns, respectively.

## .2 Structural break variations

Next, I report further results on the benchmark structural break test, continuing to rely on the standard methodology of Bai and Perron (1998). Table A.9 reports further time series variations, while still allowing for up to five structural breaks and maintaining other parameters (such as trimming at 5%).

Generally, we observe that the key findings in the main iterations are echoed in these additional series. In particular, the break clusters in the mid-16th century and the early 20th century – the latter ones clustering around World War One and Two inflections – are once again showing up across the headline and sub-component series. And once more, we see that – at a maximum of three structural breaks (RF nom TR series for sovereign bonds) – the maximum allowed breaks is never reached. In fact, four additional series – including the comprehensive "Capital gains, arith weight" iteration show no breaks over the entire observation horizon. This generally reinforces the notion in the main part of the paper that we are dealing with relatively few clear "inflection points" in capital returns – and that the long-horizon fundamental trends emphasized throughout appear quite insensitive to particular weighting and component choices.

## .3 Correlation details – further data

Here I report the expanded data from the correlation exercise discussed in section 9.1 of the main paper. Table A.10 reports additional correlation variations, by asset and period, keeping the basic 15-year rolling correlation window and regime delineations.

We observe that the basic result reported on the basis of the benchmark series in the main paper do hold

| Table A.10: Correlation matrix for German total returns, p.c. consumption growth, 1512-2025 | | | | |
|---|---|---|---|---|
| Real Estate, real TR, 1512-2025 | | | | |
| Period | 1512-1618 | 1512-1750 | 1750-2025 | 1950-2025 |
| TR, RU (baseline) | 0.216 | 0.176 | 0.282 | 0.158 |
| Real TR, SA | 0.464 | 0.570 | 0.277 | 0.018 |
| Real TR, U | 0.373 | 0.414 | 0.177 | 0.157 |
| Net TR, RU | 0.448 | 0.557 | 0.263 | 0.017 |
| Sovereign Bonds, real TR, 1512-2025 | | | | |
| AVG of bond models (baseline) | 0.182 | 0.259 | 0.165 | -0.119 |
| Random Forest | 0.159 | 0.231 | 0.137 | -0.126 |
| Decision Tree | 0.191 | 0.271 | 0.198 | -0.079 |
| Neutral Network | 0.135 | 0.129 | 0.088 | -0.087 |
| HTM | 0.209 | 0.177 | -0.019 | 0.031 |

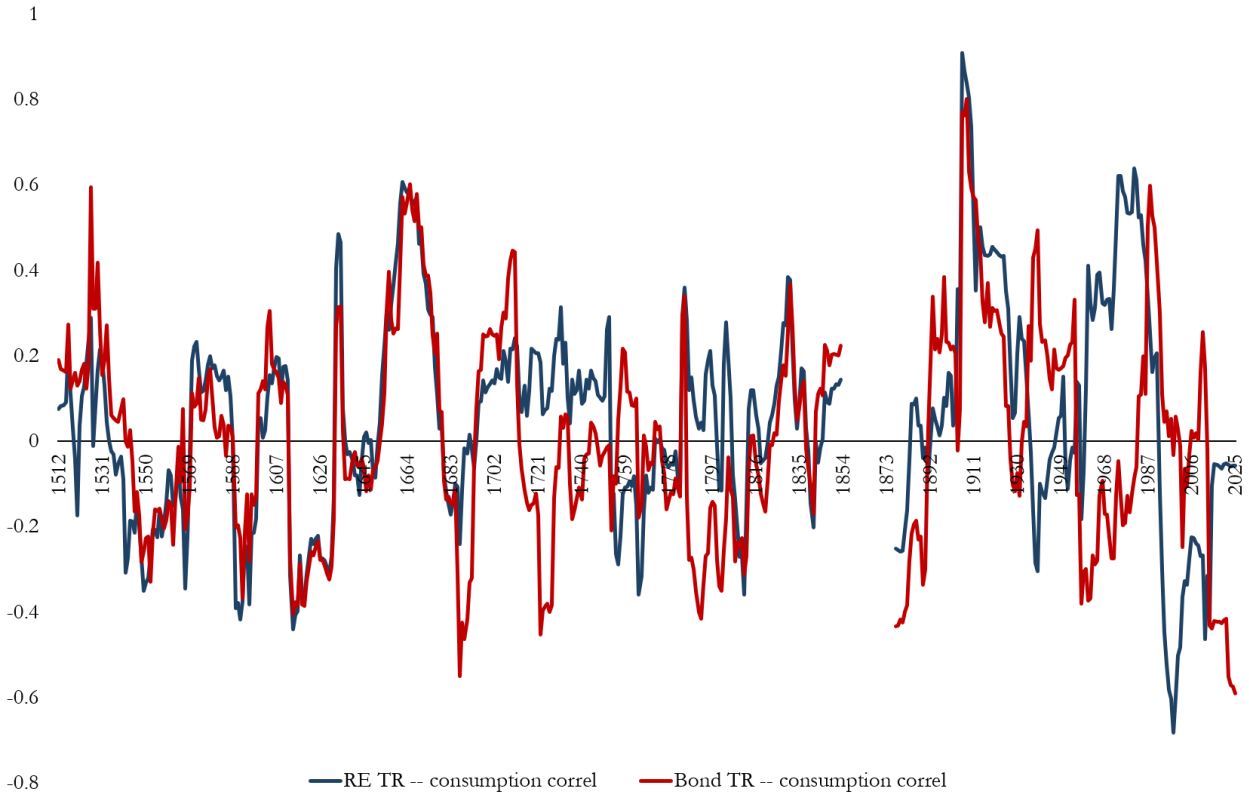
*Notes:* The table reports correlations between bond and real estate total returns on the one hand, and per capita consumption growth on the other (with all measures in real terms). Correlation reported as 15-year rolling windows. Pre-1856 real per capita consumption growth is sourced via Pfister (2020) in all variations, and period of 1856-70 is excluded. "RU" = Rural-urban; "SA" = size-adjusted; "U" = urban only. "HTM" = hold-to-maturity total returns.

across capital return components and aggregate return variations. In particular, the observation that housing procyclicality is rising over time – as expressed as a rising positive correlation with per capita consumption growth – is maintained across iterations of net/gross return bases, as well as variations of the housing sample (SA = size-adjusted; U = urban sample only; RU = rural-urban sample).

Equally, we observe that the counterpart – a falling procyclicality of sovereign bonds, as expressed as a falling correlation to per capita consumption growth over time – holds across additional variations of the sovereign bond return component. All iterations record a 0.12-0.27 correlation in the pre-1750, which drops to a generally negative correlation in the modern era. In particular, we note that this observation holds when we drop any use of machine-learning input – and instead assume a "hold-to-maturity" return profile for sovereign bonds (not an implausible assumption in the early period, as argued in section 5): even under this "HTM" variation, the positive sovereign bond procyclicality measured in the pre-1750 era has practically vanished by 1950-2025. In this sense, once more, results are not sensitive to the construction of our new total return series for bonds.

Figure A.11 displays the moving correlation between housing and bond total returns, and per capita real food consumption growth (using a 15-year rolling window). The all-time (1512-2025) correlation for housing stands at 0.025; for bonds, it stands at -0.017 on this basis. As interesting is the time trend in absolute and relative terms: real estate total returns begin the sample with a negative correlation to the business cycle (pre-1618: -0.05), at a time when bonds record a slight positive correlation (pre-1618: 0.005). By the end of the sample (post-1950), the situation has almost inversed: real estate is now positively correlated with per capita consumption growth (0.02), while sovereign bonds are negatively correlated (-0.07). The sample choices are based on key historical regimes (Thirty Years' War; World War Two), but the results are broadly echoed by alternative sampling and variations of return weightings.

Figure A.11: Correlation Real Estate / Bond total returns, per capita consumption growth, Germany 1512-2025.




*Notes:* The red line depicts the rolling correlation between sovereign bond total returns, and per capita consumption growth, for Germany. The blue line depicts the rolling correlation between real estate total returns (urban+rural basis), and per capita consumption growth, for Germany. 15-year centered rolling correlation is used. Annual real per capita food consumption growth data is used over 1512-1855, via Pfister (2022). The years 1856-70 are excluded given missing consumption data.