

**Historical monetary and  
financial statistics for policymakers:  
towards a unified framework**

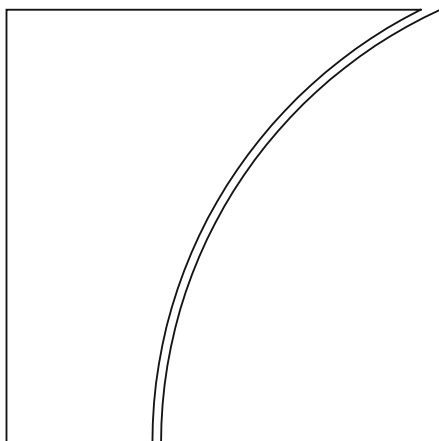
September 2022

The views expressed are those of the authors and not necessarily the views of the BIS.

This publication is available on the BIS website ([www.bis.org](http://www.bis.org)).

© *Bank for International Settlements 2022. All rights reserved. Brief excerpts may be reproduced or translated provided the source is stated.*

ISSN 1682-7651 (online)  
ISBN 978-92-9259-603-3 (online)



## BIS Papers

No 127

### Historical monetary and financial statistics for policymakers: towards a unified framework

by Vincent Bignon, Claudio Borio, Øyvind Eitrheim, Marc Flandreau, Clemens Jobst, Jan F Qvigstad and Ryland Thomas (editors)

Monetary and Economic Department

September 2022

JEL classification: C82, E42, E43, E51, N01, N10, N20, R31.

Keywords: Historical statistics, credit, house prices, interest rates, central banks, financial history.



## Foreword by Agustín Carstens

Following the outbreak of the 2007–09 Great Financial Crisis, central bankers and policymakers have become increasingly aware of the importance of macro-financial history. Nowadays, informed policymaking calls for constant mining of the historical record. This has led to extensive efforts to combine national historical data sets so as to produce international panels.

Few people realise how problematic this assemblage really is. Where they exist, individual countries' monetary and financial statistics have been constructed through nationally focused efforts. While these were often the product of heroic breakthroughs, they have never followed a common plan. The result is that the individual national series that comprise modern panels are not only imperfect, but unlike Tolstoy's unhappy families, which are all unhappy in their own way, each imperfect series is imperfect in its own style.

At the initiative of Claudio Borio, Head of the BIS Monetary and Economic Department, and Jan F Qvigstad, former Deputy Governor of the Central Bank of Norway, a group of central banks met in Basel in October 2016 and agreed to establish a small network on Historical Monetary and Financial Statistics (HMFS).<sup>1</sup> Building on long-standing intellectual synergies between central bank research departments and academic researchers on monetary and financial history, and soliciting input from leading international experts, the network has started to unpack the problem, making a start by addressing the issues involved in the construction of series for interest rates, credit aggregates and housing prices.

The present BIS Paper is born out of these travails. The result is both sobering and optimistic. It is sobering because it highlights the limitations in existing measures of our macro-financial history. It is optimistic because, while the gaps and imperfections are extensive, they can be overcome. In the end, the report offers a way forward by putting together a unified framework: using it as a how-to guide, future producers of macro-financial statistics – or “statistorians”, as we dub them here – will find step-by-step instructions. Even more importantly, the report will help economists and policymakers to find a safe passage around the pitfalls of statistical history.

Agustín Carstens  
BIS General Manager

<sup>1</sup> It has included the following countries: Austria, Canada, Denmark, France, Italy, Japan, Norway, Sweden, the United Kingdom and the United States



# Historical monetary and financial statistics for policymakers: towards a unified framework<sup>2</sup>

Vincent Bignon, Claudio Borio, Øyvind Eitrheim, Marc Flandreau, Clemens Jobst, Jan F Qvigstad and Ryland Thomas

## Abstract

In recent years, many central banks have engaged in data projects aimed at the collection and documentation of historical monetary and financial statistics (HMFS) for their respective countries. For these countries, long runs of data for key macroeconomic time series are now available and are increasingly being used in policy-oriented research of interest to various institutions. Information from these historical databases is used to draw parallels between current developments and historical events to shed light on today's policy issues in the areas of monetary and financial stability.

The aim of the Bank for International Settlements (BIS) HMFS project is to establish a network between interested central banks that have already invested in local national HMFS databases, using the BIS as a hub. The HMFS project has a decentralised structure, with each central bank providing the resources necessary to contribute information about its country. Moreover, the network rests on a partnership between central banks and academic experts, including an annual meeting in which we exchange views about the latest statistical developments in our area of interest at the member central banks. The focus of this undertaking has been fairly narrow, yet in essence it is ambitious. We have concentrated on conceptualising issues as well as reviewing the state of the art in international historical time series for interest rates, credit and housing prices. This BIS paper is an attempt to take stock of what counts as "good practice" and how it should be implemented in different contexts.

It is our understanding that there exists a large amount of tacit knowledge among economic historians of how best to construct historical time series. At the same time, we observe that different data series coexist and sometimes strongly differ in their methods and logic. Researchers and policymakers using historical statistics do not always have access to the proper expertise or are confronted with a multiplicity of choices. In this BIS paper, data sources and methods are documented and discussed for interest rates, credit and housing prices, respectively. We identify the main challenges facing the construction of such data sets. Where possible, we provide an overview of transversal issues and research-based evidence on what may constitute "good practice". A separate BIS Paper provides an overview of available historical data for the participating countries.

<sup>2</sup> Any views expressed are solely those of the authors and cannot be taken to represent those of the Bank of France, Université de Aix-Marseille, the Bank for International Settlements, the Bank of England, the Central Bank of the Republic of Austria, the Central Bank of Norway, the University of Pennsylvania or the University of Vienna.





## Contributors

### Authors team for the report

Vincent Bignon (Bank of France and Université de Aix-Marseille)

Claudio Borio (BIS)

Øyvind Eitrheim (Central Bank of Norway)

Marc Flandreau (UPenn and Wharton School)  
(Lead author)

Clemens Jobst (University of Vienna and Central Bank of the Republic of Austria)  
(Coordinator)

Ryland Thomas (Bank of England)

Jan F Qvigstad (Central Bank of Norway)

### HMFS steering committee

Claudio Borio (Bank for International Settlements)

Øyvind Eitrheim (Central Bank of Norway)

Marc Flandreau (UPenn and Wharton School)

Ryland Thomas (Bank of England) (Chair)

### HMFS country teams

Austria: Clemens Jobst

Canada: Patrick Alexander, Arthur Berger, Robin Brace, Kim Huynh, Laura Murphy,  
Trevor Sabeau, Larry Schembri

Denmark: Kim Abildgren

France: Vincent Bignon, Pierre Sicsic, David Demilly

Italy: Federico Barbiellini Amidei, Paolo Croce, Elena Genito, Paolo Piselli

Japan: Mariko Hatase, Haruko Kato, Mari Ohnuki, Shotaro Yamane

Norway: Øyvind Eitrheim, Mats B Fevolden, Jan F Qvigstad

Sweden: Tor Jacobson

United Kingdom: Ryland Thomas

United States: Mark Carlson



## In this volume

Searching for good practice in historical monetary and financial statistics –  
background and overview  
Vincent Bignon, Claudio Borio, Øyvind Eitrheim, Marc Flandreau, Clemens Jobst,  
Jan F Qvigstad and Ryland Thomas

Historical data on interest rates  
Marc Flandreau, University of Pennsylvania  
Clemens Jobst, University of Vienna and Central Bank of the Republic of Austria

Historical data on credit  
Vincent Bignon, Bank of France and Université de Aix-Marseille  
Marc Flandreau, University of Pennsylvania and Wharton School of Business

Historical data on real estate prices  
Øyvind Eitrheim, Central Bank of Norway  
Clemens Jobst, University of Vienna and Central Bank of the Republic of Austria



## Table of contents

Foreword by Agustín Carstens .....	i
Historical monetary and financial statistics for policymakers: towards a unified framework.....	iii
Contributors.....	v
In this volume.....	vii
The history of the HMFS .....	1
1. Searching for good practice in historical monetary and financial statistics.....	3
a. Introduction.....	3
b. Historical monetary and financial statistics: the state of play .....	4
c. Synthetic countries, synthetic objects, and the methodology of historical monetary and financial statistics.....	6
Box I.1 The need for “statisticians” .....	8
d. Setting the agenda.....	13
e. The HMFS going forward.....	14
References.....	15
2. Historical data on interest rates.....	19
2.1 Some general considerations.....	20
2.2 Long-term interest rates .....	25
Box II.1 Pricing terminable annuities: what we did not know they knew .....	28
Box II.2 When no clean government bond series are available.....	31
Box II.3 Dealing with conversion risk: three cases .....	34
2.3 Short-term rates .....	35
2.4 Conclusion .....	41
References.....	42
3. Historical data on credit.....	47
3.1 The lamppost syndrome: pitfalls in the construction of historical credit aggregates .....	49
3.2 Intensive margin and representative samples: do we measure bank credit adequately? .....	52
Box III.1 Dark matter in banking: a case study.....	58
3.3 Institutions that are not “banks” but act like banks .....	59
3.4 Conclusion .....	62

References.....	63
4. Historical data on real estate prices.....	68
4.1 Introduction and overview .....	68
4.2 Methods.....	71
Box IV.1 Quality-adjusted historical HPIs for Norway 1819–2020 .....	78
Box IV.2 The house sale timeline – which price to select? .....	81
4.3 Historical context and primary sources.....	81
4.4 Conclusions and lessons going forward .....	86
References.....	88
Previous volumes in this series .....	94

## The history of the HMFS

The historical monetary and financial statistics (HMFS) project is a network of central banks and academic experts hosted at the Bank for International Settlements (BIS) focused on policy-relevant long-run historical monetary and financial data. It operates under the aegis of the Irving Fisher Committee (IFC). The network's departure point is a recognition that historical data is less about *collection* than *production*. Production of data requires a thorough understanding of historical markets, instruments and institutions, and a sound methodological framework to ensure that data are intelligible and comparable across time and space. The ambition of HMFS is to provide a forum in which policymakers, statisticians, economists and historians can engage in debate about the proper approach(es) to historical data and encourage mutual learning.

A project like this is bound to have many origins. In the past, the construction of historical monetary and financial statistics has been a shared endeavour among academics and central banks. In recent years, many central banks have engaged in important data projects aimed at the production (collection and documentation) of historical monetary and financial statistics for their respective countries. Up until now however, there had been no attempt at discussing the results of these national efforts in a systematic fashion, beyond existing informal connections and positive research spillovers.

Hence the key rationale for the HMFS as an international, central bank supported project. Two milestones explain how these national efforts converged. First, in 2006, a conference was organised in Vienna at the Central Bank of the Republic of Austria by Marc Flandreau (then a professor at Sciences Po, Paris) and Edi Hochreiter (then of the Austrian central bank), to which representatives of a significant number of central banks were invited and participated. A special session considered the possibility of launching an initiative that would have been very closely related to HMFS.

One important issue raised by Flandreau was that the production of high-quality historical data was a global public good and that, absent proper design incentives, it tended to be in short supply. While national solutions had been found, there was a need for international arrangements. As policymakers, central banks had shown a keen interest in developing the tools that would enable them to remedy that specific "market failure", since they would be the primary users of the resulting data. Because of their public mandate, as the many examples indicate, central banks could take some responsibility as producers and repositories for historical monetary and financial data. What was needed therefore was an institutional mechanism to translate this logic at the international level. A suggestion was made to place the project under the aegis of an international organisation such as the BIS.

In parallel, the Central Bank of Norway (which Øyvind Eitrheim had represented at the Vienna conference), launched an ambitious project in anticipation of the celebration of its bicentenary 10 years later (in 2016). This project was led by Jan Qvigstad, then deputy governor of the Central Bank of Norway, and was managed by Øyvind Eitrheim. One concern, which the leaders of this project expressed from the beginning, was the need to ensure that the project and the national data it produced would also be of interest from an international point of view and for that reason, compatible with other data. To that end, the project coordinators appointed Michael Bordo and Marc Flandreau as advisors. The project, finalised in 2016, led to the

construction of new historical series and to several significant publications focused on Norway.

The completion of the project ushered in the launch of HMFS. At the initiative of Claudio Borio at the BIS and Jan Qvigstad, a group of central banks that had already invested in local HMFS databases (several of which had participated in the 2006 meeting) met in Basel in October 2016 and agreed to establish a network. The central banks of Austria, Canada, Denmark, France, Italy, Japan, Norway, Sweden, the United Kingdom and the United States participated in this initial phase. The aim was to share information and to identify and implement best practice approaches. Three subsequent meetings took place in Basel in October 2017, 2018 and 2019. The present volume is born out of these working sessions. In 2022, the group started operating under the aegis of the BIS Irving Fisher Committee, which brings together the heads of statistics at over ninety central banks.

Marc Flandreau (UPenn and Wharton School) supervised the finalisation of the volume with the help of Clemens Jobst (University of Vienna and the Central Bank of the Republic of Austria). The authors of this volume include: the steering group of the HMFS project, comprising Claudio Borio (BIS), Øyvind Eitrheim (Central Bank of Norway), Marc Flandreau and Ryland Thomas (Bank of England), as well as Vincent Bignon (Bank of France and Université Aix-Marseille), Clemens Jobst and Jan F Qvigstad (Central Bank of Norway). Their contributions, and that of other people who helped with the project at different points, are identified in the text.



# 1. Searching for good practice in historical monetary and financial statistics

Vincent Bignon, Claudio Borio, Øyvind Eitrheim, Marc Flandreau, Clemens Jobst, Jan F Qvigstad and Ryland Thomas

## a. Introduction

The subprime crisis led the international policy community to give closer attention to history. As was subsequently realised, the manner in which the securitisation of mortgages erred by considering too narrow a historical time frame for the computation of mortgage default risks led to a serious underestimate of potential losses, and, as a result, of the true exposure of financial investors and markets. This has been widely interpreted as a reminder of the need to consider history more seriously and has led to a healthy demand for “stronger history” and more rigorous production of historical data. However, as historians have long known, neither the study of history nor the collection of retrospective “data” are in themselves a safeguard against misguided inference.

The upshot is that, on the one hand, historical investigation will never produce the sort of definite guarantees that are generally desired for fact-based policymaking. On the other hand, historical investigation is, more often than not, the only way in which certain questions can be addressed. The study of history, rather than being a way to *establish* tangible facts, is a manner of *constructing* them. Construction takes place in a constant back and forth between source material and the investigator’s conceptual and methodological choices. Consciously integrating this process of construction into the process of research is the best way economic historians can deflect the risk of relying on the wrong historical precedent (as was done at the onset of the subprime crisis). This is where the main contribution of economic history can be found: by assisting the knowledge process, it can increase the odds that “history” will be used rather than misused (Eichengreen (2012, 2015)).

If a venerable tradition is to be summoned, the current relevance of history to macro-financial policymaking may be seen as deriving from the data-focused historical methodology articulated in the 18th century by Joseph Priestley. Priestley’s “New chart of history”, published as part of his *Lectures on history and general policy*, reflected his radical empirical agenda and had an enduring legacy in 19th century finance and financial policy (Priestley (1788)). Through the intermediary of Francis Baily, a founder of modern actuarial science and a disciple of Priestley, the “New chart of history” was subsequently revised and updated, and now provided a global chronology of nations and international conflicts (Baily (1812)). Contemporaries had already noticed that British interest rates were higher in war than in peacetime and a chronology of conflict was valuable to determine the safe rate of interest to be used in government calculations. It was also useful to private actors engaged in vital activities such as computing the present value of life insurance policies, arbitrage etc.

Subsequently, the experience of the working classes during the industrialisation phase in the 19th century led to many investigations into the standard and cost of living, and the development of official wage and price statistics (Bowley (1937)). The need to finance two world wars in the 20th century accelerated the development of official national income accounts based on international standards (Stone (1984)). The

hyperinflations of the 1920s (Cagan (1956)) and the monetary interpretation of the Great Depression by Friedman and Schwartz (1963) provoked revived interest in measurement of the money supply and the development of monetary aggregates, a subsequent landmark was Capie and Webber (1985).<sup>3</sup>

With this in mind, this volume provides a reflection on the possibility of constructing a historical scorecard for monetary and financial time series. As a result, while the ultimate goal is indeed the production of usable time series (and many are given in the online country appendices), this volume is more fundamentally concerned with identifying the difficulties such an exercise involves and articulating a methodology, as well as an organisational framework in which the methodology can be embedded. In particular, it reviews the various solutions to the difficulties encountered as well as the extent to which these solutions can be realistically implemented – and at what cost.

Accordingly, the rest of this introductory chapter reviews the key ideas that underpin the HMFS project. In particular, it outlines the epistemological problems involved in such an endeavour and surveys the reasons for the launch of HMFS at this particular juncture. It also reviews the history of the project and discusses the organisational template that has been adopted.

## b. Historical monetary and financial statistics: the state of play

There have been two dominant approaches to the gathering of historical data. On the one hand, country-specific initiatives have focused on the collection of a wide range of macroeconomic and financial time series for a particular country or geographical area. On the other hand, there have been compilations of cross-country data sets, which typically focus on a narrower set of series but which allow broader international comparisons.

Pioneering examples of a country-based approach include the book of centenary historical statistics published by the Bank of Japan (Bank of Japan (1966)) and successive editions of *Historical statistics of the United States*, culminating in the millennial edition of 2006 (Carter et al (2006)). For Britain and Ireland, Brian Mitchell's volumes on British historical statistics with various collaborators also stand out. They were themselves made possible by landmark achievements such as Deane and Cole (1967).<sup>4</sup> Subsequent cases include the work of the Scandinavian central banks (ie Danmarks Nationalbank, Central Bank of Norway and Sveriges Riksbank) and the Bank of Italy. These latter projects provide access to multiple spreadsheets together with documentation in books and papers.<sup>5</sup> A modern instance is the Bank of England's "A millennium of macroeconomic data" spreadsheet, which has produced a single data set with access to (multiple) historical series and spliced (headline) data up to the present day constructed by Bank of England researchers. In its approach, the Bank

<sup>3</sup> Bordo (1986). See eg the collection of cross-country studies focused on the long-term behaviour of velocity (Bordo and Jonung (1987, 1990)).

<sup>4</sup> See Mitchell and Deane (1962), Mitchell and Jones (1971) and Mitchell (1988).

<sup>5</sup> See Abildgren (2017), Edvinsson et al (2010, 2014) and Eitrheim et al (2004, 2007). For Italy, the pioneering contribution was De Mattia (1968); it was extended by the historical series of the Bank of Italy (Collana Storica della Banca d'Italia – [www.bancaditalia.it/pubblicazioni/collana-storica/index.html](http://www.bancaditalia.it/pubblicazioni/collana-storica/index.html)), published in several separate volumes (Cotula et al (2003)). Details on the documentation and data are available online: [www.bancaditalia.it/statistiche/tematiche/stat-storiche/index.html](http://www.bancaditalia.it/statistiche/tematiche/stat-storiche/index.html).

of England was more reliant on references to original data sources for the construction details.<sup>6</sup>

These projects offer valuable material because of the way they link data, data sources, and the institutional and economic significance of alternative series. By design however, country-specific approaches tend to be idiosyncratic. The Bank of England and Danmarks Nationalbank have primarily relied on individuals at those central banks with a research interest in developing historical statistics. The Bank of Italy, Central Bank of Norway and Sveriges Riksbank have typically commissioned periodic, large-scale projects involving collaboration between central bank and academic researchers. International comparability is not usually within their remit.

Early illustrations of cross-country initiatives include the pioneering work on international historical national statistics by Angus Maddison,<sup>7</sup> the broad-based data sets on European historical statistics (Mitchell (1975)), and data sets for international interest rates (Homer and Sylla (2005)). In turn, such efforts – especially those concerning long-run prices and income dynamics – go back to previous attempts aimed at assembling large cross-country data sets. A landmark that generated many subsequent spillovers was the International Scientific Committee on Price History created in 1929 by William Beveridge and Edwin Francis Gay, the result of a five-year grant by the Rockefeller Foundation. However, except for the work of Homer and Sylla, and the fiscal series in Mitchell, these data sets provide limited macro-financial information.

More recently, the financial globalisation of the 1990s and 2000s and newfound concerns about long-run debt dynamics have boosted interest in historical precedents and led researchers to complement data on real variables with data on finance. A pioneering effort at documenting financial data for the so-called first globalisation wave (1880–1913) is the Global Finance database assembled by Flandreau and Zumer (2004).<sup>8</sup> It provides, in part, data collated from existing sources, but more importantly it is an entirely new attempt to organise a database that assembles novel time series for government debt and interest rates, for 18 sovereign countries and 15 colonies, from primary and secondary archival sources.

The Global Finance database stands out by enabling a focus on comparability and the development of a uniform cross-country methodology. It was the first

<sup>6</sup> See Thomas and Dimsdale (2017). The spreadsheet (available at [www.bankofengland.co.uk/-/media/boe/files/statistics/research-datasets/a-millennium-of-macroeconomic-data-for-the-uk.xls](http://www.bankofengland.co.uk/-/media/boe/files/statistics/research-datasets/a-millennium-of-macroeconomic-data-for-the-uk.xls)) is a collection of official and research data organised into annual, quarterly, monthly, weekly and daily sections. Each section contains a “headline series” sheet containing a set of continuous time series that link various historical components together, plus more detailed worksheets, which show how these time series are calculated together with alternative historical series that can be used. As a general principle, the annual series are typically based on separate and more reliable sources than higher-frequency series and cover a wider range of data. The calculations underlying joined-up series are made clear in the spreadsheet, allowing users to create their own series based on different assumptions. At present, users need to consult underlying sources for detailed information on the construction of the component historical data series used, although additional documentation of the methods and choices made is currently underway.

<sup>7</sup> Maddison (2006). The data are maintained and revised by the Maddison Project, hosted by the University of Groningen and available here: [www.rug.nl/ggdc/historicaldevelopment/maddison/](http://www.rug.nl/ggdc/historicaldevelopment/maddison/). A related project for the post-1950 period is the Penn World Table, on the most recent version see Feenstra et al (2015).

<sup>8</sup> See Flandreau and Zumer (2004) for detail. The data are available at <https://eh.net/database/global-finance/>. See Accominotti et al (2011) for a discussion of the colonial data, which were added subsequently.

contribution that demonstrated the possibility of producing consistent historical time series for public debt going far back in time. Its discussion of government borrowing costs also provides some of the background for Chapter 2 on interest rates. Although its data have been collated in subsequent work (for instance Reinhart and Rogoff (2009), Eichengreen et al (2019)), no real attention was paid to the methodological pitfalls it identified, and in fact the more recent contributions splice together disjointed methodologies.<sup>9</sup> Following the subprime crisis, the Global Finance database became the basis of the Macrohistory Lab database (Jordà et al (2016, 2017)). The Macrohistory Lab collated macro-financial historical series from second-hand sources for 14 countries going back to 1870.<sup>10</sup>

These historical efforts also had parallels with more focused developments and interest by statisticians engaging with contemporary issues, and who sought to expand the data horizon. Early analytical work includes a volume based on the joint conference on Real Estate Indicators and Financial Stability organised by the International Monetary Fund (IMF) and the Bank for International Settlements (BIS) in Washington, DC, in 2003 (BIS (2005)). Its importance was demonstrated by the Great Financial Crisis of 2008–09, and led the BIS to start publishing a detailed data set on property prices. The BIS database currently covers residential property prices for 58 advanced and emerging economies with a starting date of 1970. Another international house price database worth mentioning is maintained by the Federal Reserve Bank of Dallas and documented in Mack et al (2011). They have collected data from national public sources and the database contains quarterly data for 22 countries going back to 1975.<sup>11</sup>

### c. Synthetic countries, synthetic objects, and the methodology of historical monetary and financial statistics

#### The problem

Moving back to the historical series, many economists and policymakers who consume these series are generally unaware of the underlying hypotheses on which some of the data are based. They tend to believe that the historical data has been obtained through a process similar to the one described in the previous paragraph. But in fact, this is not true. There may be significant methodological and, as a result, epistemological difficulties that undermine the use of such series as instruments to address certain questions. As researchers involved in this field know only too well, a large number of the historical time series cross-sectional databases that currently circulate were originally produced by dedicated scholars who did carefully document their sources and, in particular, warned about limits. More often than not, however, the caveats were lost in compilation. Out of mere repetition, some series have established themselves as the final word on the subject despite the expressly stated caveats of their originators.

But the current use of historical data by economists faces an additional challenge, namely doubts about the comparability of series produced in specific national contexts. Today, problems of international comparability of macroeconomic series

<sup>9</sup> Another second-hand data project of note is the crisis data set for 21 countries by Bordo et al (2001).

<sup>10</sup> It also collated an international series of house prices (documented in Knoll et al (2017)).

<sup>11</sup> Available at [www.bis.org/statistics/pp.htm](http://www.bis.org/statistics/pp.htm); [www.dallasfed.org/institute/houseprice](http://www.dallasfed.org/institute/houseprice).

are dealt with by international agreements negotiated under the auspices of international organisations such as Eurostat, the BIS or the IMF. Before the recent age of international statistical harmonisation, however, no such agency was in charge of ensuring international comparability, although efforts began early on to establish an international conversation. International statistical congresses started being gathered in 1853, notably at the instigation of British statisticians and actuaries. An International Statistical Institute was founded in 1885. The League of Nations made extensive efforts to try and inspire economic data standardisation. But despite this, historical statistics are the territory of the idiosyncratic. Individual series are typically born out of individual historians' forays in an age of national self-sufficiency, the product of teamwork by scholars specialising in the quantitative history of a given country. Only in a second stage, as just explained, have they been collated in cross-country databases – with very few projects having international sensemaking and comparability at their heart.

Against this backdrop, we show that the best way to address the problem of comparability is to bring the primary source back in. We believe that it is essential to overcome the tendency to create long threads of references where the link to the original material is lost. Linking data explicitly to sources and providing documentation that fills in the dots between those sources and the processed product also addresses the question of national idiosyncrasies and time breaks. Giving proper room to the microeconomic and institutional evidence that underpins macro-financial series is the necessary prerequisite to ensuring proper meaning, and secure interspatial and intertemporal consistency. For instance, what transactions underlie money market interest rates as quoted in, say, contemporary newspapers? How are bank balance sheets used to calculate credit aggregates drawn up and collated, and which institutions are considered to be banks in the first place? What is the price of a house and how have these prices been recorded? If the primary material is itself already a statistic calculated by contemporaries, can we retrace what kind of original material they used and how they decided to treat it?

When following this principle, it becomes clear that the publication of macro-financial historical data has to be informed by solid methodological interventions. It is the main goal of this paper to outline that this conversation has not yet taken place. We argue that opening up the debate requires a combination of skills. On the one hand, the talents of the historian are needed. On the other, the competence of the statistician is required. To use a neologism, what is needed are "statistorians" fluent in the languages of both history and statistics (Box I.1).<sup>12</sup> The statistorian is properly equipped to handle two of the biggest challenges that affect long-run historical macro-financial series: shifting spaces and changing objects of study. Through their agency, it will become possible to clarify the terms of the conversation between historian and statistician. First, outlining the steps involved in the construction of data enables readers to retrace the logic and to falsify it, if they wish. Second, this will also enable producers (we think in particular of producers of historical statistics, who are hereby invited to join the project) to contribute extensions and improvements, and suggest alternatives that might be better suited to the specific research question at hand and so on.

<sup>12</sup> In fact, the term is older than the HMFS project, as Federico Barbiellini-Amidei found. The term "statistorian" can be traced back to freelance baseball researcher L Robert Davids, who used it to speak of persons interested in the compilation of historical data about that sport (Phillips (2019)).

## The need for “statistorians”

*Vincent Bignon and Marc Flandreau*

Economists, statisticians and historians take different perspectives when studying the evolution of macro-financial magnitudes. Empirical economists or quantitative economic historians start from a specific question (for example, “what variables cause financial crises or banking instability?”). They construct empirical tests to disentangle fundamental determinants from spurious correlations, which they run on panel data sets collated from secondary historical sources or downloaded from pre-existing spreadsheets. The validity of the results and inferences hinge on the quality of the input. When it comes to using the past to shed light on present-day policy debates, consistent and representative historical data sets are thus necessary. To address this requirement, the economist needs the “statistorian”, a hybrid scholar who, as Janus, combines two sides: on the one hand skills borrowed from the best practices of the statistician; and on the other, the skills of the historian. It is the goal of this box to spell out in what those skills consist.

The methodology of statisticians and historians have much in common when it comes to constructing financial series. Unlike the top-down approach taken by the economist, they both start with a definition of financial instruments and prices, eg interest rates or credit, as used in *a given society at a given point in time*. Relying on the source material that has survived, in the case of the historian, and information obtained from regulation and surveys, in the case of the statistician, they are concerned with the representativeness of the aggregates so constructed. Both use a variety of methods and complement their data with qualitative information to account for what is not directly observable. In the case of the historian, use of archives and thorough knowledge of historical context and institutional details allow the interpolation of information that is missing in sources. In the case of the statistician, thorough knowledge of the methods for statistical surveys, including sample selection, help achieve the same goal. In other words, statisticians and historians know very well that the veracity of data is not a given, that they are representations of reality. As a result, they are well aware that the concept of reality embodied in the data has to be questioned and checked against many pieces of information.

Also shared by statisticians and historians is their awareness of possible limitations hanging over the comparison of statistical measures over time and space. This goes a long way to understanding why many historians have focused their studies on short periods of time or on single countries only. It also explains why statisticians tend to shy away from linking two series constructed with different methodologies. Further, it contributes to an understanding of why it is often difficult to find “old” series on official statistical websites, even if those series had been produced by the very same official agencies. This is the case even though statisticians working today have at their disposal techniques like the rebasing of series to ensure the (local) intertemporal consistency of data in the face of, say, the emergence of new financial instruments produced by market participants or evolving definitions of the underlying concept to be measured.

The crucial difference between statisticians and historians relates to the social and institutional context in which their data production is performed. Today, problems of international or temporal comparability of statistical series are taken care of by international organisations such as Eurostat, the BIS or the IMF. These organisations provide forums to discuss good practice and further elaborate the resulting manuals. National statisticians are regularly involved in those debates. They are aware of the gaps in the international methodologies when applied to their specific country cases and try to minimise blatant inconsistencies through cooperation across organisations. The innovation of HMFS is to serve as a less formal yet similar platform for international exchange and mutual learning, providing a productive environment for the future statistorians of money and finance.

## Synthetic countries

One key challenge to which this volume pays close attention is how to deal with the changing borders of nation states. We offer to resolve it by considering synthetic countries. As historian Benedict Anderson famously emphasised, nations and nation states are fictions that have been “imagined” in a variety of ways over the ages (Anderson (1983)). According to Anderson, the production of the nation state involved the creation of narratives that helped “naturalise” it by contributing to the

creation of imagined communities. As he shows, this was true of the colonising state and colonised state alike. In pursuing a project such as HMFS, we find ourselves in a position that is not so different: we “imagine” nations as entities that exist in a meaningful way through space and time, and we assume that these entities can also be the object of statistical projections which can ultimately assist with the falsification of economic hypotheses.

Unsurprisingly, this approach finds its justification (naturalisation) in the historical process through which the modern territorial state was made. As described by political scientist Eric Helleiner, the 19th century witnessed a process of “territorialisation of money” in many countries (Helleiner (2003)). Dominant banks of issue covered the country with a fine grid of branches – or, as in Britain, they became the bankers’ bank – and a uniform means of payment was given high-powered money status. This was a political process, which ultimately contributed to the making of the modern states themselves. We must never underestimate the fact that this process was slow. For instance, as shown by Hugh Rockoff, it was only in the 1930s that regional interest rate discrepancies in the United States became sufficiently small to warrant the characterisation of the United States as a “monetary union” (Rockoff (2000)). At the most basic level, this implies that, contrary to today, the statistician also has to ascertain whether geographically localised financial markets are representative of the whole “country” in which they are situated.

One proof of the importance of the statistical narrative as part of the set of narratives that made the nation state is the fact that we find it harnessed by individuals who challenged existing national orders. A good example is provided by the early 20th century work of the Hungarian nationalist economist, Friedrich von Föllner. In the heyday of Hungarian separatism, he reconstructed the balance of payments of an independent Hungary, as if it were not part of the Habsburg Empire (Austria-Hungary). His goal was to underscore the feasibility of Hungary’s exit from the monetary union (Föllner (1908), Flandreau (2006b)).

The rise of nations in fact explains why reconstructing long-run macro-financial time series is a feasible task. Something we may want to call the “macroeconomic nation state” did emerge progressively in the modern era, and documenting its record becomes possible. This process, turned into statistics, can shed light on a variety of modern, policy-relevant issues. It is thus a legitimate and honourable task. Yet in doing so, we should not forget that nothing is natural or self-evident in the categories we have inherited. Indeed, when we construct a set of countries spanning a given historical time period we actually make two powerful assumptions. First, we assume that in some fundamental way, the number in each of the cells in the table is “comparable” to the number in another cell in the table. Second, we also assume that each entity observed at date  $t$  was the successor of the same entity at date  $t-1$ . In other words, we are creating synthetic nations.

This raises its own problems. A typical case, suggested by the previous reference to Hungary, is provided by the history of “Austria.” This “country” moved from being – like many other polities at the time and after – an amalgamation of territories ruled by the Habsburgs (with a core that was increasingly integrated over the 18th and 19th centuries) to becoming, in 1867, “Austria-Hungary,” representing one half of a dual-monarchy still ruled by the Habsburgs. After its disintegration in 1918, Austria-Hungary in turn became a collection of “successor states” that included Austria, Hungary and Czechoslovakia, and there were more border adjustments than we have space to discuss. In 1938, Austria was absorbed into the German Nazi Reich until the

end of the Second World War and after this it became again a sovereign state. The territory of 21st century Austria is a fraction of that of Austria in the age of the Treaty of Vienna (1815). Given this, how is one going to tabulate "Austria" in any macro-historical work?

This problem is far from isolated and it is bound to grow significantly, as one becomes more ambitious in terms of how far back in time one wants to go and how many countries one wants to cover. For instance, it is widely recognised that in the early modern era and until well into the 19th century, in many parts of the world, the relevant economic entity was the city rather than the nation state (Jacobs (1985)). Helleiner's nationalisation of money, in fact, was a conquest by central governments of monetary and financial spaces originally claimed by cities (Epstein (2000), Flandreau et al (2009)). As historians and political scientists know very well, the question of why nation states emerged as a dominant political form in the 18th, 19th and 20th century remains open. Neither is the nation state the end of history. With the introduction of the euro and the foundation of the European Central Bank, the European Union created a supranational currency, which brings up new quandaries. The Eurosystem balance sheet can only be analysed at the economic and monetary union level, while the banking sector still remains, despite banking union, organised along national lines.

An even more important problem with the "national" definition of borders, already apparent from the previous discussion, is that political and economic spaces need not coincide. Political borders are, of course, important as they often define the space of sovereignty in which government regulations apply. Along with the growing power of the modern state, the 19th and 20th centuries have arguably witnessed strong efforts by states to transform their politically defined territories into unified economic units. But there is historically no unidirectional process which is achieved at some particular point in time. Rather, processes of integration and disintegration continuously create new, now local, now transnational, spaces of economic activity. Financial markets offer prime examples. Both historically and today, more often than not, stock exchange indices are not a measure of the financing costs for the host economy but for the set of regional or global actors that have access to that particular market. Viewed from the shareholders' angle, stock indices might not reflect the returns for national investors but those of a potentially highly international group of security holders.

It may be believed that some unitary states that have not encountered political transformations over time on the scale of the Austrian empire should provide a better incarnation of the abstraction which we try to follow. Indeed, initial historical data collection efforts were focused on countries displaying a certain degree of geographical continuity such as, say, Britain or Japan, both being islands, perhaps not coincidentally. Yet "Britain" is hardly a straightforward case either, as a closer look will immediately reveal. Starting with the early modern period only, there was the union of Scotland and England in 1707, ratified by both parliaments. The union with Ireland in 1800 and the separation of Southern Ireland in 1922 were further turning points. Scottish independence, if it ever materialises, could provide the next one. Each event has led to disruption in "British" or "UK" historical data, including changes in the geographical coverage of the official fiscal and trade data, raising arduous problems of splicing at each turning point.

The previous paragraph has steered clear from that additional conceptual complication in viewing Britain as a unitary state, namely the British Empire. A specific



case of transnational financial linkages are imperial connections. Because of the ties between finance and politics, empires tend to create a special kind of transnational home bias which is bound to shape (distort?) historical monetary and financial statistics (Flandreau (2006a)). As De Cecco (1975) has argued, it is impossible to conceive of the depth of the sterling money market without understanding that the market for sterling was an imperial one. Conversely, British colonies borrowing in the metropolitan market were, to a point, more akin to the municipalities of a unified country. Bearing witness to this important problem and taking into account the significance of imperial subjection has significantly changed the results from macro-historical work – for instance assessing the determinants of macro-financial variables on credit risk, which can be shown to be determined by political bonds and legal constraints, rather than macro-financial “fundamentals” (Accominotti et al (2011)).

### Synthetic objects

A second problem which we propose to address is that economic objects evolve over time. It is not only geography that is in flux in macro-financial history but also the objects of observation. A piece of land, a physical asset that seems to be the incarnation of stability, is, in economic terms, a bundle of property rights (and obligations) subject to change (Lamoreaux (2011)). The customs and laws circumscribing the uses to which the land can be put by its owner and the possibilities creditors have to seize and dispose of it, if posted as collateral, will evolve over time. We argue that this requires identifying synthetic objects.

In this respect, continuity of designations can be misleading and while occasionally different labels refer to similar objects, more often than not the same name denotes different objects over time. To begin with, central banks were not always called central banks but instead, in the 18th and 19th century, “banks of issue” and several banks of issue could exist at once in the same country (although eventually one took over the rest). Following the same line of thought, another example of spurious continuity is provided by the central bank policy rate. For most countries there is indeed a rate coming under the heading of the official or central bank “discount rate.” Very often it is the short-term rate that is available for the longest stretch of time on a regular basis, in the case of the Bank of England going back to the year 1694 (Clapham (1944)). Yet, historically, the term referred to very different animals (Jobst and Ugolini (2016)). It could mean a rate used in a discount operation for the lucky customers who were given access to the discount window (Bank of England before the 1850s); a cost ceiling at which unlimited liquidity was provided (Bank of England after the 1850s); an officially posted maximum, while significant operations were transacted at lower rates (Reichsbank after 1880s), a subsidised rate subject to discount limits providing floor to market rates (Bundesbank post-Second World War); or a below market rate in a facility facing stigma (the Federal Reserve System 1960s to 2003).

As a result, financial prices have very different meanings in their respective environments. Arbitrage is a powerful assumption that allows the postulation that local prices convey information pertaining to the entire macroeconomy from a potentially very small set of information. But arbitrage presumes integration, taking us back to the problem alluded to earlier. In other words, the synthesising of a price is related to the synthesising of countries. Only after tedious exertions can researchers be sure that in place of an ill-defined “discount rate”, they are dealing with well distilled synthetic policy rates, consistent and comparable across time and space.

While historical financial information can come from private sources, very often it is the result of government regulation that requires the recording of prices and balance sheets. Regulation is helpful to statisticians because it usually follows explicit definitions and strives for consistency. Again, however, the modern researcher should not be misled by a naïve reading of categories. A case in point is bank credit. As argued in Chapter 3, historically (as well as today) banking services were provided by a range of institutions. For a long time, with a few exceptions, these institutions operated under general commercial law. The 19th century saw an acceleration in the differentiation of legal categories, which became subject to targeted regulation (Grossman (2010)). As a consequence, available information on the legally regulated banking sector refers only to a subset of the total banking sector defined according to its economic functions. There is no reason to assume that the unregulated banking sector was small or that its size was stable relative to the regulated banks. In fact, regulation leads to a variant of “Goodhart’s law”, in which the use of data for policy (in this case regulatory action on banks) results in changes in the way that data are generated.

### Back to the future: using historical time series

The previous discussion suggests that statisticians find themselves in a position quite like Fellner’s, who tried to construct, through the use of balance of payments numbers, a synthetic Hungary for time periods when it did not “exist”. Likewise, the building of historical financial data sets is really about constructing synthetic nations and synthetic concepts that span the historical universe. It is fundamental to bear this in mind. The problem it raises is not anachronism. The most situated historical discourse has to assume something “anachronistic” in order to proceed. Rather, the concrete question to address is what is gained and lost by assuming that these synthetic historical nations, and magnitudes (whether prices or quantities), are useful objects to think with in the first place. *To answer that, we need to evaluate the balance between the benefits, in terms of additional insights from creating such synthetic objects, and the costs, in terms of knowledge distortions.*

A good illustration of this issue relates to the identification of the origin of times for the construction of particular data sets and series or, more generally, the question of periodisation. What constitutes a nation depends on the question asked: if we were writing, say, a military history of the world, certain dates would be relevant but these dates may or may not be relevant if we write a history of botany. Given the focus on financial policymaking, the criterion above suggests that, at first blush, we should look for information on whether a sufficiently integrated and delineated monetary and financial space existed during the period under consideration, so that modern questions can be asked in a plausible fashion.

Within this general template, an array of possible criteria will be considered. In some cases, the correct starting point in time will be given by the creation of a national payment system, sometimes itself associated with the creation of the bank of issue. In other cases, it will coincide with the adoption of a national banking law. In general, money supply, money markets and credit markets etc were shaped by the advent of central banks or national banking systems. Of course, before this was the case, some kind of (synthetic) money supply, credit variables or interest rates might still be constructed. The exercise can be fascinating, informative and rewarding. But it is also more perilous. In the end, the “ambiguity” of the past in the light of the present should never be held against it. Rather, taking history into account makes for more interesting insights and a better understanding of the present as well.

## d. Setting the agenda

HMFS has the objective of constructing international series to support debate on policy by provoking new ideas, offering different perspectives and adding historical depth. Unlike in earlier cross-country initiatives the focus is not on assembling series per se. The ultimate goal is not to collate, but to produce.

### Objectives and strategy

To that end, HMFS seeks to create momentum through what we describe as the production algorithm of the individual series that ultimately enter cross-country data sets. Put differently, the objective is not to certify specific data sets by including them in a spreadsheet but to organise a transparent and opposable process for their collection, transformation, documentation, publication and update.

This is the challenge. To address it, HMFS follows two lines of attack.

On the one hand, it presents macro-financial data from the countries in the network. One critical aspect of these data is their accessibility, both in literal terms and in terms of methodological integrity. In other words, the data offered through the HMFS framework should clarify origin and limitations. This should also enable users to think “against” the data (the methodological chapters below provides examples). In particular, by providing as much primary material as possible along with the final product, HMFS will promote critical interventions.<sup>13</sup> The country chapters in the follow-up BIS paper provide a first step in this direction; the initial values of the production algorithm.

On the other hand, HMFS aims to organise an expert forum in which policymakers and academics can interact in order to foster a critical understanding of the data underlying macroeconomic history and how they inform policy questions. In particular, it is hoped that HMFS will provide orientation and guidance to researchers inside and outside central banks, inspiring efforts to generate higher quality data.

Both approaches reinforce each other. Data production and update keeps bringing up questions of broader relevance, while policy discussions feed back into the data construction process. The approach should ultimately foster the production of high-quality international statistics with a satisfactory level of epistemological solidity.

### Starting points

While the HMFS project is about an approach, not a specific set of data, a choice had to be made regarding which macro-financial areas to cover in this debut publication. Three types of series were selected: credit, interest rates and housing prices. Why these three series to start with?

One reason is that taken together, these data inform the policy conversation on the secular pattern of real interest rates and the role of leverage in asset price booms (Borio and Lowe (2002), Borio et al (2017)). Another is that the initial series will serve as cases in point to consider the broader issues which the HMFS seeks to resolve. Credit is a quantity, much like monetary aggregates. Correspondingly, a key question is the conceptual delineation of what should be considered credit and what should

<sup>13</sup> This is also consistent with the imperatives of reliability, serviceability and accessibility, as defined by BIS (2003) and IMF (2012).

not. By contrast, interest rates are prices, the prices of assets. Discussing “benchmark” interest rates as we do will serve to illustrate the specific questions raised by collecting and interpreting price data for individual assets. This includes identifying relevant assets and ensuring that the price series do not suffer from unwelcome distortions. Finally, housing prices, much like equity prices, constitute a third type of statistic or measure, ie indices. Again, the construction of indices raises its own set of questions, in particular the selection of prices to be included, the choice of a weighting scheme and the methods used to render the index comparable over time.

## Two volumes as first outcome

Its peculiar nature has imparted to the HMFS project an original design. It is defined through its working dynamics. It was launched as a cluster of central banks hosting significant expertise on the subject of historical statistics, along with academic experts. The group first met in 2016 to identify the perimeter of its focus. Heterogeneity displayed by the countries involved was discussed in subsequent meetings and this served to articulate a number of principles. This methodological volume is the first outcome of this dialogue between sources, data and critical assessment. As the chapters underline, the various fields explored at this stage vary substantially in terms of completeness. In particular, a strong message from the section on credit data is that existing material is not only very heterogeneous across countries, but it also requires significant supplementary efforts in order to produce satisfactory data. At the same time, it shows there are strong reasons to be hopeful and demonstrates the scope for cross-pollination.

The methodological volume is followed by a collection of essays on the countries united in the network in which we attempt to map out the present “state of the art”. It is well understood that this is only a first iteration and that, in particular, some of the series provided there will require subsequent revisions. Nonetheless, compared with previous material, they stand out not only because they contain some original material, but also because, for the first time, they are thoroughly documented. With the exception of Flandreau and Zumer (2004) this is the first international study to provide an integrated perspective that includes contextualising documentation along with a discussion of the reasons for the methods used. Not least, the following chapters flag the core problematic aspects of the series, so that future research can proceed in full knowledge of that uncertainty. The next step will thus be the creation and maintenance of a website that includes a dynamic database, to be continuously improved and enhanced.

## e. The HMFS going forward

The template provided here will facilitate the expansion of the network to new series and new countries. It will permit us to carry on the task of putting together what is currently known. Doing this is bound to bring to the fore questions that were not considered at an earlier stage. More importantly, the approach will provide a mechanism whereby the initial input can be revised through focused studies. In other words, additional research, which we call for, will modify the approach of the core group, leading to updates, corrections, and to the identification of new methodological problems and solutions. Put differently, this study sets out the basis for *a framework for aggregation and international convergence*.

The point that we would like to emphasise is that HMFS is not tied to any particular method, beyond the creation of an empirically driven and conceptually focused research environment. The seemingly simple endeavour of producing long-run time series that are robust and comparable is bound to encounter significant conceptual and empirical hurdles. Their identification, the clear articulation of methods and principles to address them, and finally, the honest recognition of the associated limitations represent the three conditions for such a project to be both successful and useful. Accordingly, our goal is to create a space in which primary material can be transformed into data in a manner that is both transparent and open-ended. From that vantage point, this study can be read as a first attempt to summarise the state of knowledge and, as a result, to help other researchers to emulate and, at the same time, question the logic adopted in the database. This should allow the further integration of national statistical approaches into an emerging international framework. In other words, the study should be read as much as a milestone as a call for contributions. It is an open invitation.

## References

Abildgren, K (2017): "A chart & data book on the monetary and financial history of Denmark", *Working Paper*, May.

Accominotti, O, M Flandreau and R Rezzik (2011): "The spread of empire: Clio and the measurement of colonial borrowing costs", *The Economic History Review*, vol 64, no 2, pp 385–407.

Anderson, B (1983): *Imagined communities: reflections on the origin and spread of nationalism*, London, Verso.

Baily, F (1812): *The description and use of a new chart of history, exhibiting the most material revolutions that have taken place in the principal empires, kingdoms and states, from the earliest authentic records, to the commencement of the present year*, London, J Johnson & Company.

Bank for International Settlements (BIS) (2003): "Guide to the international financial statistics", *BIS Papers*, no 14, February.

——— (2005): "Real estate indicators and financial stability", proceedings of a joint conference organised by the BIS and the IMF in Washington, DC, 27–28 October 2003, *BIS Papers*, no 21, April.

Bank of Japan (1966): *Hundred-year statistics of the Japanese economy*, Statistics Department, Bank of Japan.

Bordo, M (1986): "Explorations in monetary history: a survey of the literature", *NBER Working Papers*, no 1821.

Bordo, M, B Eichengreen, D Klingebiel, M Martinez-Peria and A Rose (2001): "Is the crisis problem growing more severe?", *Economic Policy*, vol 16, no 32, pp 53–82.

Bordo, M and L Jonung (1987): *The long-run behavior of the velocity of circulation: the international evidence*, Cambridge University Press.

——— (1990): "The long-run behavior of the velocity of circulation: the institutional approach revisited", *Journal of Policy Modeling*, vol 12, no 2, summer, pp 165–97.

Borio, C, P Disyatat, M Juselius and P Rungcharoenkitkul (2017): "Why so low for so long? A long-term view of real interest rates", *BIS Working Papers*, no 685.

- Borio, C and P Lowe (2002): "Asset prices, financial and monetary stability: exploring the nexus", *BIS Working Papers*, no 114.
- Bowley, A (1937): *Wages and income in the United Kingdom since 1860*, Cambridge University Press.
- Cagan, P (1956): "The monetary dynamics of hyperinflation", in M Friedman (ed), *Studies in the quantity theory of money*, Chicago Press, University of Chicago, pp 25–117.
- Capie, F and A Webber (1985): *A monetary history of the United Kingdom, 1870–1982*, London: George Allen and Unwin.
- Carter, S, S Gartner, M Haines, A Olmstead, R Sutch and G Wright (eds) (2006): *The historical statistics of the United States: millennial edition*, Cambridge University Press.
- Clapham, J (1944): *The Bank of England, volume 1 1694–1797*, Cambridge University Press.
- Cotula F, M De Cecco and G Toniolo (eds) (2003): *La Banca d'Italia. Sintesi della ricerca storica 1893–1960*, Collana Storica della Banca d'Italia, Saggi e ricerche, vol IV, Roma-Bari, Laterza.
- Deane, P and W Cole (1967): *British economic growth, 1688–1959: trends and structure*, second edition, Cambridge University Press.
- De Cecco, M (1975): *Money and empire: the international gold standard, 1890–1914*, Oxford, Blackwell.
- De Mattia, R (1968): *I Bilanci degli istituti di emissione italiani dal 1845 al 1936: altre serie storiche di interesse monetario e fonti*, Rome: Banca d'Italia.
- Eichengreen, B (2012): "Economic history and economic policy", *Journal of Economic History*, vol 72, no 2, pp 289–307.
- (2015): *Hall of mirrors: the Great Depression, the Great Recession, and the Uses – and Misuses – of History*, Oxford University Press.
- Eichengreen, B, A El-Ganainy, R Esteves and K Mitchener (2019): "Public debt through the ages," *NBER Working Papers*, no 25494
- Edvinsson, R, T Jacobson and D Waldenström (2010): *Exchange rates, prices, and wages, 1277–2008. Historical monetary and financial statistics for Sweden*, Sveriges Riksbank and Ekerlids Förlag.
- (2014): *House prices, stock returns, national accounts, and the Riksbank balance sheet, 1620–2012. Historical monetary and financial statistics for Sweden*, Sveriges Riksbank and Ekerlids Förlag.
- Eitrheim, Ø, J Klovland and J Qvigstad (eds) (2004): "Historical monetary statistics for Norway 1819–2003", *Norges Bank Occasional Paper*, no 35.
- (2007): "Historical monetary statistics for Norway – Part II", *Norges Bank Occasional Paper*, no 38.
- Epstein, S (2000): *Freedom and growth: the rise of states and markets in Europe, 1300–1750*, London, Routledge.
- Feenstra, R, R Inklaar and M Timmer (2015): "The next generation of the Penn World Table", *American Economic Review*, vol 105, no 10, pp 3150–82, [www.ggdc.net/pwt](http://www.ggdc.net/pwt).

- Fellner, F (1908): *Die Zahlungsbilanz Ungarns. Ein Beitrag zur Lehre von der internationalen Zahlungsbilanz im Allgemeinen*, Wien, Deuticke.
- Flandreau, M (2006a): "Home biases, nineteenth century style", *Journal of the European Economic Association*, vol 4, no 2/3, pp 634–43.
- (2006b): "The logic of compromise: monetary bargaining in Austria-Hungary, 1867–1913", *European Review of Economic History*, vol 10, no 1, pp 3–33.
- Flandreau, M, C Jobst, C Galimard and P Nogués-Marco (2009): "Monetary geography in the mid-18th century", *Cambridge Journal of Regions, Economy and Society*, pp 1–23.
- Flandreau, M and F Zumer (2004): *The making of global finance 1880–1913*, Paris, OECD.
- Friedman, M and A Schwartz (1963): *A monetary history of the United States, 1867–1960*. Princeton University Press.
- Grossman, R (2010): *Unsettled account: the evolution of banking in the industrialised world since 1800*, Princeton University Press.
- Guinnane, T (2002): "Delegated monitors, large and small: Germany's banking system, 1800–1914", *Journal of Economic Literature*, vol 40, no 1, pp 73–124.
- Helleiner, E (2003): *The making of national money: territorial currencies in historical perspective*, Cornell University Press.
- Homer, S and R Sylla (2005): *A history of interest rates*, fourth edition, John Wiley and Sons.
- International Monetary Fund (IMF) (2012): *Data quality assessment framework*, <https://dsbb.imf.org/dqrs/DQAF> (last accessed 21 December 2021).
- Jacobs, J (1985): *Cities and the wealth of nations*, New York, Random House.
- Jobst, C and S Ugolini (2016): "The coevolution of money markets and monetary policy, 1815–2008" in M Bordo et al (eds), *Central banks at a crossroads*, Cambridge University Press, pp 145–94.
- Jordà, O, M Schularick and A Taylor (2016): Jordà-Schularick-Taylor macrohistory database, technical report: [www.macrohistory.net/data](http://www.macrohistory.net/data). JSTDatasetR2, release 2, November.
- (2017): "Macrofinancial history and the new business cycle facts", in M Eichenbaum and J Parker (eds), *NBER macroeconomics annual 2016*, vol 31, University of Chicago Press.
- Knoll, K, M Schularick and T Steger (2017): "No price like home: global house prices, 1870–2012", *American Economic Review*, vol 107, no 2, pp 331–53.
- Lamoreaux, N (2011): "The mystery of property rights: a US perspective", *The Journal of Economic History*, vol 71, no 2, pp 275–306.
- Mack, A, E Martínez-García and V Grossman (2011): "A cross-country quarterly database of real house prices: a methodological note", Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute, working paper, no 99, revised January 2017.
- Maddison, A (2006): *The world economy: volume 1: a millennial perspective and volume 2: historical statistics*, Paris, OECD.

Mitchell, B (1975): *European historical statistics 1750–1970*, London Palgrave Macmillan.

——— (1988): *British historical statistics*, Cambridge University Press.

Mitchell, B and P Deane (1962): *Abstract of British historical statistics*, Cambridge University Press.

Mitchell, B and H Jones (1971): *Second abstract of British historical statistics*, Cambridge University Press.

Phillips, C (2019): *Scouting and scoring: how we know what we know about baseball*, Princeton University Press.

Priestley, J (1788): *Lectures on history and general policy*, Dublin, P Byrne.

Reinhart, C and K Rogoff (2009): *This time is different: eight centuries of financial folly*, Princeton University Press.

Rockoff, H (2000): "How long did it take the United States to become an optimal currency area?", *NBER Historical Working Papers*, no 124, April.

Stone, R (1984): *The accounts of society*, Nobel Memorial Lecture, 8 December.

Thomas, R and N Dimsdale (2017): *A millennium of macroeconomic data*, Bank of England OBRA dataset, [www.bankofengland.co.uk/research/Pages/datasets/default.aspx](http://www.bankofengland.co.uk/research/Pages/datasets/default.aspx).



## 2. Historical data on interest rates<sup>14</sup>

Marc Flandreau, University of Pennsylvania and Wharton School<sup>15</sup>

Clemens Jobst, University of Vienna and Central Bank of the Republic of Austria<sup>16</sup>

The level of interest rates is a key macroeconomic and financial magnitude. It informs us about the price of time, the balance between the needs of investors and those of consumers and producers, as well as their perception of risk – or rather of the confidence and trust that permeate an economy and society. In any economy there are as many interest rates as there are individual assets. Macro-financial research is primarily interested in the general level of interest rates. This gives a special role to certain assets sometimes called “safe” or “benchmark”. In practice, they are fixed income instruments that can be valued without expensive and prolonged analysis – they are sometimes called “information insensitive” – that can be traded easily on the market and, ideally, entail limited inherent default risk.<sup>17</sup> As a result, their pricing (the determination of their intrinsic yield) captures many of the underlying factors of the economy at large, as opposed to the idiosyncratic aspects affecting individual securities. When we are interested in measuring specific risks, then the price of these safe assets are again the relevant instruments. They serve as a standard against which various risk premia incorporated in the prices of other assets can then be gauged.

Historically, safe assets have been manufactured by various entities (typically, but not only, by banks and governments) and their structuring cannot be separated from the infrastructures (political, economic, legal and social) in which they were created. That means interest rates cannot be gauged without an understanding of the “markets” in which these assets were traded. In general, reaching such an understanding requires historical investigations that typically extend beyond the narrow question at hand when researchers look for a good indicator of, say, long term interest rates. While it is the goal of HMFS to cater for such a need, it is important to recognise that behind any “clean” series, a complete research infrastructure has typically been put together in order to be able, precisely, to address the question at hand in a meaningful way. While the availability and ostensible intelligibility of the series may give consumers the impression that it can be used in any way they want, it is important to keep in mind the underlying infrastructure, which in many cases may severely restrict its practical use.

In this following chapter, we discuss candidate instruments at both the long and the short end of the yield curve. We also trace their respective historical evolutions and the typical characteristics that users need to be aware of. The focus is on those features of assets that can introduce idiosyncrasies, which, if not taken into account

<sup>14</sup> We thank Vincent Bignon, Mike Bordo, Claudio Borio, Øyvind Eitrheim, Kim Oosterlinck, Ryland Thomas as well as Mark Carlson and Federico Barbiellini-Amidei for their feedback on earlier drafts.

<sup>15</sup> Marc Flandreau is Howard Marks Professor of Economic History at the University of Pennsylvania, Department of History and the Wharton School, and is a CEPR Research Fellow. He is a member of the steering committee for the Historical Monetary and Financial Statistics project at the Bank for International Settlements.

<sup>16</sup> Clemens Jobst is Professor of Economic and Social History at the University of Vienna, a CEPR Research Fellow, and serves as academic advisor to the Austrian central bank. All views expressed here are his own and do not necessarily reflect those of the Central Bank of the Republic of Austria or the Eurosystem.

<sup>17</sup> See Gorton (2017) for a discussion.

properly, may result in misleading inferences being drawn. The key principle is that, in order to achieve the required level of understanding, any series of interest rates has to build on the “pedigree” of the relevant financial products, which has to be meticulously reconstructed and documented. We conclude that the investment in this information infrastructure, if sometimes costly, yields significant returns (no pun intended) in terms of our understanding of financial systems past and present.

## 2.1 Some general considerations

Fixed income securities comprise a wide range of risks. To understand the nature and determinants of these risks, the literature relies on the concept of a risk-free rate, which serves as a starting point for all expected return models. The difference between actual returns and the risk-free rate – the risk premium – is then interpreted as compensation to investors for bearing an asset’s risks and is used to measure investors’ time and risk preferences.

This explains the importance of securing a time series for the “risk-free rate”. To qualify as risk-free, an asset has to meet several conditions. In particular, there can be no risk of default associated with its cash flows and no reinvestment risk. In the real world, such assets are not available. What exists instead is more or less distant approximations to the ideal. To identify risk-free rates, the modern empirical literature typically relies on government bonds of “credible” countries such as the United States, as they are assumed to display a smaller default risk than the bonds of other, default-prone countries (Fabozzi (2001)).

However, recent literature has provided evidence that yields of some government bonds might be misleadingly low and might actually be below the risk-free rate. The reason is that, because government bonds also qualify as safe assets, they provide additional benefits such as being easily traded by uninformed investors, used as collateral or perform other roles similar to that of money. Because of these additional returns – termed “convenience yields” – investors are typically willing to pay higher prices for these bonds, effectively lowering the yield (Krishnamurthy and Vissing-Jorgensen (2012), Nagel (2016)). This is a structural feature, independent of the postulated cyclical property of the yield curve to slope downward in anticipation of recessions (Bordo and Haubrich (2008)).

Historians have repeatedly spotted this phenomenon, especially at the core of the global financial system. A discussion by Fratianni and Spinelli (2006) about the Genoa capital market in the early modern era identifies strikingly low yields (below 2%) for government bonds during the entire 17th century. Predominant amongst these were “luoghi”, the Genoese Republic funding instrument issued by the state’s bank and treasurer, the Casa San Giorgio. As Fratianni and Spinelli (2006) point out, luoghi were widely owned by Genoese and by foreigners, and they were a favoured form of collateral used by bankers, individual borrowers and tax collectors.<sup>18</sup> Likewise, in the last decade of the 19th century, the nominal yield on British government perpetual bonds (“consols”) started to fall towards zero, with the slope of the yield curve exhibiting a structural downward sloping pattern. This too was possibly a reflection of the large demand for a safe long-term security, which could be used as collateral. One interpretation is that parliamentary regulations over trustee portfolio choices and restrictions upon the composition of “sinking funds” (used to amortise

<sup>18</sup> Drawing on Sieveking (1906) and Heers (1961). On Casa di San Giorgio, see also Felloni (2010).

debts) created a captive market for British consols, pushing long-term yields below short-term interest rates and inducing the observed pattern.<sup>19</sup>

A related concept in the finance literature is the benchmark or reference yield. Benchmark yields are returns on assets, thought to be representative for a broader class of instruments: some specific government bond representing all government debt or government debt itself as a benchmark for the general financing conditions of companies and households in the economy. While serving a similar purpose in empirical finance, the concept of a benchmark asset differs from the risk-free asset, because a benchmark asset might well carry risk. In fact, some benchmarks *need to carry risk* to be useful. An example was Libor, which served as a proxy for funding costs of banks with rates set at Libor plus  $x$  basis points. Benchmark status can be tied to yield, liquidity or a high information content for the market as a whole (Dunne et al (2007), Yuan (2005)). Note however, that benchmark assets, when recognised as such by market participants, tend to exhibit superior liquidity and might thereby earn a convenience yield, acting to push their yields down (Chen et al (2007)).

When looking for assets and their yields to construct historically significant low-risk benchmark series, it is useful to keep in mind both the criteria of theoretical finance as well as the caveats of empirical research. Risk-free rates are elusive, both today and in the past. Like any empirical research, historical research has to be pragmatic but it also has to be mindful of the limitations that can be encountered. This renders the ultimate goal difficult to reach in some cases. A good starting point is a focus on rates “considered high grade by contemporary standards” (Homer and Sylla (1991)). This initial question will inform the subsequent stages of the research. The key is to be transparent about what is “high grade”, what are “contemporary standards” and how both change over time. From there some general considerations for the production of low-risk interest rate series can be derived. Below we discuss five key considerations in the production of risk-free and low-risk benchmark interest rate series.

### The selection of assets

It is important to note that what constitutes the high(est) grade of security is usually historically determined. In modern “advanced” economies, government debt is usually considered to be of the lowest risk, followed by high-quality corporate debt. This idea is closely tied to the notion of a “sovereign ceiling”, a controversial concept formerly used by rating agencies in an era of rampant capital controls. According to this view, the highest grade (and thus the lowest yield) is enjoyed by securities issued by sovereign governments. According to this hypothesis, sovereign debt is always the lowest risk in its own country, presumably because the sovereign is capable of expropriating other assets to repay its own debts. This notion still informs some empirical work today, though in today’s world it is well known that some corporations enjoy better ratings and better borrowing terms than their governments.<sup>20</sup> Still, reputational spillovers ensure that sovereign ratings remain a significant determinant of the credit rating assigned to corporations (Borensztein et al (2013)), and goes some way to explaining the persistence of this view. The sovereign’s rating may not provide

<sup>19</sup> For a recent discussion of legal restrictions over trustee investment and references, see Chavaz and Flandreau (2017).

<sup>20</sup> This is the case of Energias de Portugal for instance (Almeida et al (2017)).

a ceiling to the rating received by other domestic borrowers, but variations of the former would nonetheless inform variations of the latter.

History presents us with many situations in which the highest grade security was not provided by the debts issued by the sovereign. Instead, large chartered companies, cities and merchants competed against the “sovereign” for best standing. Dickson (1967) provides a survey of the English situation in the 18th century with special reference to the stock issued by large chartered companies such as the East India Company and the Bank of England. Epstein (2000) discusses the case of early modern municipal borrowers.<sup>21</sup> Flandreau et al (2009) discuss the case of merchants in 18th century France. In all three cases, the economy’s credit ceiling is not provided directly by the sovereign.

The problem of the credit “pecking order” is not only that it is difficult to work out, but that it also changes over time. As stated, during the early modern period, the credit quality of propertied individuals, corporations and municipalities towered over that of the central government for extended periods of time in several countries. Until the mid-19th century, Prussian bonds secured on agricultural lands fetched higher prices than government bonds for a similar nominal rate of interest (Meitzen (1871)). Similarly, while US government bonds were considered to be a benchmark in 19th century in the United States, they faded in importance behind top-notch corporate debt in the 20th century. As Wooldridge (2001) argues, in the post-Second World War period, market participants in the United States and abroad referred to bonds issued by first-class corporations like AT&T as the key benchmarks, rather than US government bonds.

As a result, regardless of their ability to approximate a risk-free rate there is a direct interest in collecting long-term series on certain specific assets. This interest also arises because the question “what is the benchmark rate for such or such country?” cannot be answered without due diligence. Such series provide a useful first step in the evaluation process and enable future scholars to discover new rates which actually “beat” those series previously thought to provide a benchmark.<sup>22</sup> The rates paid by the government are a good example. Even if, in a given historical economy, these rates did not constitute the lower bound to borrowing costs, their collection is still useful.

In some cases, as we demonstrate below, even the price of a highly idiosyncratic government bond (eg one denominated in some specific foreign currency or including unusual clauses) might be insightful. A general recommendation when analysing risk-free rates is therefore to collect and document the price of (and yield on) central government debt as a very important series; but also to document all characteristics that possibly qualify its use as a benchmark rate and complement it with alternative high-grade private assets, whenever these carry lower yields. Also, note that what is considered representative varies according to maturity. As we will see, historically, government debt has tended to provide a benchmark for long-term rates for much longer than it has for short-term rates.

<sup>21</sup> See also Stasavage (2007) for a careful treatment and use of Epstein’s data.

<sup>22</sup> Schmelzing (2020).

## Individual yields versus averages or indices

A second related question concerns the use of **individual yields versus averages or indices**. While in the case of equity or real estate prices, (careful) averaging produces a broader and more representative measure of the underlying trends, this is not the case with low-risk benchmarks. By definition, there can only be a few benchmarks, and so mixing benchmark securities with other assets in an index clouds the picture. From contemporary testimony, for instance, it is known that 19th century financial observers typically looked at a single price for each country, like the British 2½% consols or the 3% French rentes. The governments had issued other debt instruments, but this debt was certainly less liquid and potentially carried further idiosyncratic features. In most cases, it is therefore better to rely on a few representative assets, carefully chosen and discussed. Nonetheless, there might be cases in which the creation of indices is unavoidable. Under such circumstances they must be carefully documented and, in particular, the primary data used in the construction must be transparent and accessible. In other contexts, however, composite series can make sense, eg when constructing a measure of average financing costs or the average spread between investment-grade and speculative-grade bonds. In these cases again, the need for proper documentation and access to the primary data cannot be emphasised enough.

## Maturity of assets

A third dimension of interest concerns the **maturity of assets**. The modern finance literature considers a whole range of interest rate maturities along the yield curve. Typically, in historical data, the best information we have is on very short maturity instruments (between three and six months) and very long-term instruments (such as bonds with a maturity of more than 10 years, sometimes perpetual), although the details vary across countries. These maturities feature prominently in the macro-financial historical literature, in part because short maturities have informed research on money and monetary policy, while long-term maturities have served to assess creditworthiness and debt sustainability. By contrast, medium-term maturities have been less frequently collected, used and discussed.

A related aspect is that an instrument with a long maturity at issuance, if it is not a genuine perpetuity, will mechanically become shorter term as time progresses and redemption approaches. As will be discussed below, this complicates the calculation, interpretation and comparison of yields both internationally and over time. An important point to bear in mind is that, in most countries, the issuance of long-term instruments was, historically, not sufficiently regular to allow for actual calculation of (for example) consistent 10-year yield series based on bonds with a remaining maturity of 10 years, as is typically done nowadays. With respect to the issuance of intermediate maturities Allen (2019) and Garbade (2007) claim, in respect of the United States, that these became available in the 1970s when government debt managers shifted to the regular issuance of relatively short-term bonds, compared with the longer maturities that had prevailed earlier such as the Liberty Bonds issued during and immediately after the First World War and which had a maturity of 20–25 years. However, it is our impression that, historically, short to intermediate maturities were not as uncommon for government debt as their relative paucity in extant studies implies. The US government, for instance, issued Treasury notes in the 1920s that had maturities of between one and five years (Moody (1922), pp 480–2). Similar instruments show up in histories of international banking as vehicles for short-term

funding, whether syndicated or not (see Gille (1965) and (1967) for illustrations). This means that there are still untapped stores of data waiting to be exploited.

### Institutional environment

A fourth issue that pervades the study of long- and short-term markets alike is the importance of the **institutional environment** that may explain the liquidity or illiquidity of assets. Quite frequently, market outcomes were impacted by official institutions in charge of intervening to regulate the price of the assets on the market, including by buying and selling, as well as laws limiting arbitrage. Illiquidity may also arise in otherwise very liquid capital markets if assets are held by certain types of investor (such as pensioners), or buy-and-hold institutional investors (such as insurance companies) or subject to extensive buyback provisions (as in the case of sinking funds). Chavaz and Flandreau (2017) find that, for British colonies, the liquidity premia (or illiquidity discount) stood at around 50 basis points. Moreover, illiquidity may cause idiosyncratic behaviour, reducing the information content of the yields. Such distortions may cause the yields calculated from such instruments to be very poor indicators of the background interest rates we are looking for.

A direct consequence of the issues discussed so far is the **(non-)availability of adequate assets**. Under the best possible circumstances, the chosen historical instruments will exhibit some more or less worrying idiosyncratic features. But under the worst-case scenario, it is possible that no proper series exist. Understanding why this was the case ought to be a primary task at this point, and having done this, it becomes possible to determine whether or not some alternative could be conceived. Although it may be useful in the absence of a proper series to provide a substitute, this can only be done with extreme care and a transparent discussion, when the series is introduced, of the possible problems and issues this may raise. Ideally a proof of concept will be provided, demonstrating, say, the overlap between the alternative and the actual measure during a previous or subsequent period when both are available. The point to bear in mind is that switching to a “lower” quality series is likely to create a structural problem, compromising the very significance of the measurement exercise. If the substitute simply does not measure the phenomenon one is interested in, then it is not a substitute at all.

### Sources and treatment of data

Lastly, some comments on **sources and the post-collection treatment of data** are in order. These may sound trivial, but it is our experience that they have defeated more than one researcher. The reader familiar with such pitfalls may skip the following and join us at the next section.

Financial market prices and interest rates are often reported in several sources. Which source is selected, and why, should be duly reported. It might be useful to compare alternative sources to identify and understand discrepancies. Secondary sources that compile prices, such as journals, yearbooks, and private or government agencies, might follow editorial or other policies which make them an unsuitable source for some purposes.<sup>23</sup> A typical trap, into which many published series have fallen in the past, is the case in which the original source reports the same price that was quoted in the previous chronological entry, because there were no transactions,

<sup>23</sup> See eg the critique of *The Investor's Monthly Manual* as a source for stock market capitalisation, in Hannah (2018).

or no transactions of which the source was aware. As a result, a flat price series is not evidence of a quiet market, but an echo of reporting practices. Such assumptions, if used, should be highlighted and documented, so that users can make their own judgment about their appropriateness. Even (or especially) for price listings, some news outlets might have an incentive to distort facts or even report untrue prices – news outlets may also be vehicles of influence rather than unbiased vehicles of information or evidence.

Given that financial prices are often available on a high frequency, it is also crucial to clarify the date for which the data are collected (for instance, a decision must be made in the case of, say, monthly data regarding which day of the month we choose). If the only available observations are mixed (eg the available data contain a set of observations on a month-end basis and another set of observations on a monthly average basis) it is best to document each subseries separately and then create an additional joined-up series on a more consistent basis. For example, a two-period average of the month-end data is taken and combined with the observations on a monthly average basis. Break points – when you must switch from one asset to another – require particular attention. Problems of splicing pervade the construction of historical data sets. In general, the best way to go about it is to implement the switchover not necessarily when the “former” series ends, but rather to use all the information over the period when the old and new series coexist. Even a short overlapping period may highlight useful details (such as differences in level, volatility and trend movements), which help inform the appropriate splice point and the method of extrapolation. Experimentation with different assumptions provides a good test of the quality of the splicing and helps to avoid “locking” in level differences that are inappropriate in earlier or later periods.

The fact that there is not always a one-to-one correspondence between (reported) prices and (desired) yields means that publications of historical financial time series should always give both yields (calculated as to the best knowledge of the researcher) as well as the original price data. If a series must be transformed, the original series as well as the transformed series should be reported simultaneously and the transformation procedure carefully described. It is essential that future researchers can disagree with the construction and provide an alternative.

## 2.2 Long-term interest rates

Based on the general considerations above, we will now discuss long- and short-term rates in more detail. Given that long-term yields usually have to be calculated by researchers themselves based on historical price information, the emphasis here will be on the pitfalls in such exercises. In this context, we point to the necessity of understanding the underlying markets. By contrast, short-term rates are typically already available expressed as yields or discount rates in percentage terms.

### Selection of assets

As already indicated, when looking for safe assets that convey an idea of the general financing costs in a historical economy, **central government debt** is often the first choice and we start there. Given that countries typically issue both domestic and foreign currency denominated debt, it is useful to look at both yields, or at least to clearly spell out the currency that a particular yield series refers to. Domestic and foreign currency yields have been shown in previous research to differ significantly, with the spread between bonds denominated in alternative currencies providing

ground for interesting inferences. For instance Flandreau and Oosterlinck (2012) study the relation between alternative denominations of risk-free Indian debts in the 19th century to extract a pure measure of silver depreciation risk.<sup>24</sup> The yield on foreign currency debt measures default risk, while the yield on domestic currency debt measures currency and/or inflation risk, even though it should be borne in mind that on occasions governments have defaulted on domestic debt rather than inflating it away (Reinhart and Rogoff (2009)). There is no reason, therefore, to “prefer” either domestic or foreign currency yields.

As already emphasised, interest rates on alternative high-grade assets are available, and some of these assets may provide greater security than government debt. The key challenge is that, unlike for central government, there are several issuers to consider at any point in time and no systematic argument exists, a priori, regarding which one should be chosen. Nonetheless, several natural candidates can be identified and are briefly discussed below.

A number of historical studies show the importance of **mortgage markets** as a form of credit and provide an invaluable source of interest rate data. A country that has received particular attention in this respect is France, owing to the peculiar institution of the notary, which has been emulated in many Latin countries (Poisson (1985), Hoffman et al (2000)). On the one hand, a problematic aspect of notarial sources is that, by construction, they do not yield a single, well identified data series but rather an array of individual rates, reflecting individual risks. On the other hand, the fact that such credit was backed by land, and had a trustworthy mechanism that facilitated the seizing of collateral, implies that at least some contracts falling under this denomination represented moderate risk. In theory, work could be done to extract high-quality interest rate series from such material, although this has not been undertaken so far. Chaining of individual loans as well as intertemporal comparisons (not to mention international comparisons) need further elaboration. The HMFS project highlights the interest that there would be in undertaking more research in this direction.

Another interesting market providing information on credit is the market for **corporate debt**. Today, indices for interest rates on high-grade corporate loans are routinely generated. Doing this, however, requires an ability to identify high-grade corporate debt, something which can only be done straightforwardly when there are rating agencies at work producing such ratings (Hickman (1958)). In other words, the availability of such material hinges on rating agencies actually operating. Because ratings were developed first in the United States, this material is available for this country for as far back as the early 20th century, but it only comes into existence much later in other countries. In many cases, there is nothing available before the 1970s (Sylla (2002)). Extrapolating on this insight, a synthetic measure of high-grade corporate yields could in theory be reconstructed in many countries for a much longer time period. Using Poor’s handbooks, for instance, it should be possible to construct a high-grade bond price index for US railways going far back into the 19th century (Chandler (1981)). For other countries, measures of creditworthiness could be worked out. The bonds of chartered companies likewise provide an instrument for examining high-grade corporate yields in the early modern period. This is again a topic for future research and may warrant the interest of HMFS researchers.<sup>25</sup>

<sup>24</sup> For an early discussion of silver risk in the context of a monetary regime change, see Calomiris (1992).

<sup>25</sup> See Dickson (1967) for a background discussion and more specifically concerning the East India Company, see Nogués-Marco and Vam Malle (2007) and Chamley (2011).



A third important family of instruments is the **covered bond**. Thanks to the dual protection they offer – investors have a claim against the issuer and a claim over the cover pool in the case of the issuer’s insolvency – covered bonds are considered to be high-grade. The combination of the issuer’s guarantee and pooling should also limit the impact of idiosyncratic risks associated with rates such as those on individual mortgage loans, as described above. Moreover, the early appearance of such bonds has been confirmed, creating the scope for interesting breakthroughs. Such bonds became available in Prussia in the late 18th century (they were known as “Pfandbriefe”) and they gradually spread through Europe during the 19th century (Schulte (1918)). Covered bonds thus provide a possible alternative to government debt prices as a representative risk-free asset. However, covered bonds were not always and everywhere the safe instrument that they are nowadays purported to be, as attested by differences in actual interest rates.<sup>26</sup> This suggests a coexistence of alternative standards of securitisation (Packer et al (2007)). At this stage, more work needs to be done to better understand the legal setups in different countries and their consequences for risk and pricing. Reaching a better understanding of these questions could help extract more information from such asset prices, and provide methods to adjust or harmonise individual rates and extract relevant, time-consistent information. This is a knowledge growth area in which HMFS is particularly interested.

### Historical yield calculations

The remainder of the section discusses the challenges in the calculation of historical yields, both for benchmark and other long-term instruments.

#### *Case 1. The problem of anachronism*

The problem of determining interest rates from the prices of securities typically presents itself as a root extraction problem. The yield must be calculated – indeed, constructed – as a function of the individual bond’s characteristics. For instance, in the case of a perpetual bond, which has attracted some attention in the literature, the yield is obtained by dividing the fixed coupon or dividend by the price of the asset. Perpetual bonds are a relatively recent innovation. For instance, during the early modern era, a more popular instrument was the annuity. It paid a fixed income each period, but unlike the perpetual bond it did so only for a certain number of years. Annuities could also feature amortisation schedules and repayment options, among other trappings. The point is, the calculation of yields is generally not straightforward, raising both technical and methodological issues. In particular, given theoretical advances in financial economics at any point in time, which method should be used when extracting yields? Should we use modern methods (at the risk of anachronism)? But if we try and aim at consistency with contemporary techniques and proficiency, we face another question: how similar are these to the methods which contemporary investors used?

<sup>26</sup> Due to fraudulent behaviour, one German mortgage bank failed in 1900, while bond holders suffered losses at another two (Sattler (2022)). In Austria-Hungary, malpractices in the run-up to the crisis of 1873 led to the enactment of new legislation designed to protect investors (Schulte (1912)). In the United States, several attempts to establish covered bonds failed, underscoring the importance of the legal setup, regulation and supervision for the actual riskiness of covered bonds (Snowden (2010)).

## Pricing terminable annuities: what we did not know they knew

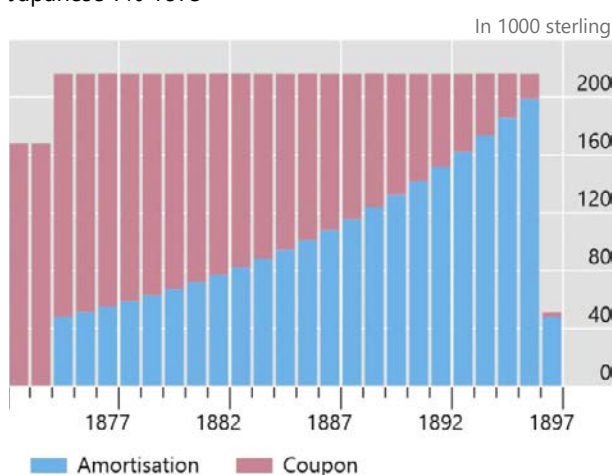
The extraction of the yield from a bond price series entails significant computing issues. Although the concept of present value had been understood since the “commercial revolution” in the late middle ages, thanks to the works of Fibonacci and others (Goetzmann and Rouwenhorst (2005)), the actual computation required the numerical resolution of a demanding equation. This has led authors to suggest that, owing to the lack of computing power, investors must have resorted to approximations, particularly in the case of so-called terminable annuities, ie contracts which guaranteed to the annuitant a stream of income for a finite period and which were sometimes characterised by complex amortisation procedures (eg Mauro et al (2006)).

In this box, we summarise work by Flandreau and Legentilhomme (2022) devoted to the problem of root extraction. They suggest that the computing power already accumulated in the 19th century was considerable. Moreover, they show that contracts were structured in a way that facilitated computations. A case in point was the predominant “accumulative sinking fund” method, whereby reimbursements were increased at the end of each year by the dividend savings on bonds reimbursed during the previous year. As a result, the sum of the coupon and amortisation was essentially constant. Graph II.1 (left-hand panel) illustrates this in the case of the Japanese bond of 1873. This mode of structuring enabled the agent to resolve the yield equation with the help of existing algorithms that had high converging power. Computations were handled by actuaries who contributed powerfully to pricing and whose services were affordable to affluent investors. Economists, statisticians and historians take different perspectives when studying the evolution of macro-financial magnitudes. Empirical economists or quantitative economic historians start from a specific question (for example, “what variables cause financial crises or banking instability?”). They construct empirical tests to disentangle fundamental determinants from spurious correlations, which they run on panel data sets collated from secondary historical sources or downloaded from pre-existing spreadsheets. The validity of the results and inferences hinge on the quality of the input. When it comes to using the past to shed light on present-day policy debates, consistent and representative historical data sets are thus necessary. To address this requirement, the economist needs the “statistorian”, a hybrid scholar who, as Janus, combines two sides: on the one hand skills borrowed from the best practices of the statistician, and on the other the skills of the historian. It is the goal of this box to spell out in what those skills consist.

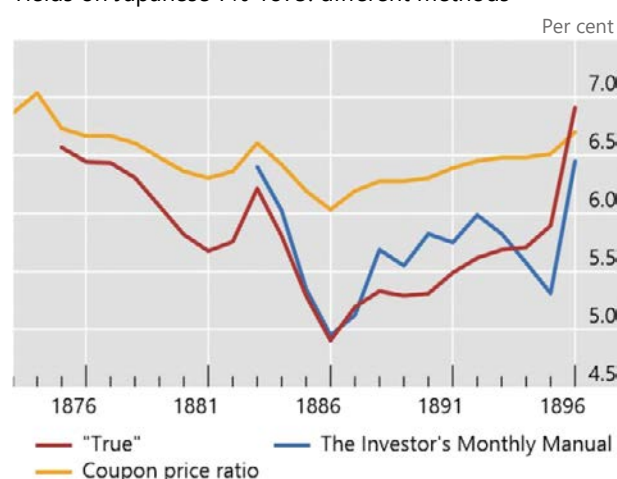
### Pricing terminable securities

Graph II.1

Japanese 7% 1873



Yields on Japanese 7% 1873: different methods



Source: Flandreau and Legentilhomme (2022).

Another takeaway from Flandreau and Legentilhomme is the fact that this sophistication was also made accessible to less wealthy investors through the provision of industrialised versions. A prominent example was the one published monthly by *The Investor's Monthly Manual*. Graph II.1 (right-hand panel) shows the yield on Japanese bonds as calculated by the incorrect application for the coupon/price formula (which Mauro et al (2006) argue, incorrectly, that investors were using), by the "true" mathematical formula that we would use today, and by the cheaper alternative, informed by the cutting-edge actuarial method offered to its readers by *The Economist* through its companion publication, *The Investor's Monthly Manual*. As can be seen, the "true" yield series (the one which would be computed using a computer today) was similar to the yield reported by *The Investor's Monthly Manual*. The last two are similar to one another and convey a different picture to that conveyed by the incorrect coupon/price formula. This provides grounds for our contention that, short of an actual investigation of contemporary methods, modern calculations are to be preferred as the starting point, until an actual investigation is conducted.

As outlined in the introduction, the HMFS project strongly argues for careful consideration of the historical context. It is obviously advisable to garner knowledge and understanding of the financial techniques available at any point in time, as they may have affected the manner in which economic agents priced products. However, such knowledge is difficult to accumulate. Its absence may sometimes constitute a serious roadblock for researchers interested in producing yield series.

A recent contribution by Flandreau and Legentilhomme (2022) offers some relevant evidence. Discussing ostensibly complex reimbursement clauses prevalent in 19th century government bonds traded in London, the authors demonstrate that these clauses were in fact priced according to what we would regard as "modern" criteria. Specifically, pricing took into account maturity, as reflected by calculations circulated in the contemporary press. This is interesting because earlier research suggested that, owing to investor ignorance, investors would have used a "simile" approach to bond pricing. They would have likened the securities they traded to the simpler bonds they understood. However, as Flandreau and Legentilhomme argue, the seemingly complicated features that became popular in many government bonds were popular precisely because they could be solved and precise yields could be constructed. Box II.1 provides a summary discussion, focusing on the extraction of the yield from an 1873 Japanese sterling bond. It illustrates the quality of historical pricing formulae judged by modern standards, as opposed to the assumed naiveté of historical investors.

In other words, the principle we advocate in the case of financial products amounts to a kind of market rationality hypothesis: although the modern observer may not be aware (yet) of the underlying historical techniques, it is safe to assume that some method existed, as demonstrated by the fact that financial contracts were traded and therefore understood. One way to put it is to say that 19th century financiers gave themselves only problems they could solve (to one extent or another). At a first level of approximation, therefore, pending the necessary additional explorations which may qualify the above rule, it might be reasonable to use "modern" methods as an approximation of (yet unknown) historical methods to extract yields from bond prices.

### *Case 2. When the sovereign is not sovereign*

There are situations when debt issuance is subject to restrictions that distort yields for the “benchmark” borrower to the point that they are no longer representative of the credit situation for the economy at large. This will arise, for instance, if credit is rationed. A typical instance is when the borrowing government entity is subject to supranational control. Supranational control might allow the borrower to access capital at a cheap price, but only in small amounts. In such a situation, the measured yield is not informative either as to the conditions in which other borrowers in the economy can access the capital market or, indeed, of the conditions in which the benchmark borrower himself can secure a marginal unit of capital.

Historical examples include the situation of colonies in the British and French empires, their yield’s principal factor being the yield in the mother country (Flandreau (2006), Accominotti et al (2010)). Another illustration of this are the special arrangements that secured certain debts by providing guarantees from foreign powers, as in the case of international control in Greece or the Ottoman Empire in the 19th century (Esteves and Tunçer (2016)), or the case of the Austrian loan of 1923 issued under the auspices of the League of Nations and jointly guaranteed by the United Kingdom, France and five other countries (Flores and Decorzant (2016)).

What can be done in such cases? Options may be limited. In the case of lending restrictions, their very occurrence is indicative of significant information asymmetries when lending to the country in question. It is likely therefore, that no alternative debt instrument without restrictions will be available. Box II.2 takes a look at this question for interwar Austria. In this country during the period 1918–30, no yield on foreign currency central government debt can be calculated. Such instances have to be identified clearly. Here, “missing observations” does not mean that the data have not yet been found or collected. Rather, it is the fact that there are no prices which is indicative of the institutional and economic constraints facing the country.

## When no clean government bond series are available

Sometimes there is no government bond that can serve as a single measure of government financing costs. Examples include when countries access global bond markets for the first time or are coming out of a period of financial trouble and default. A case in point is provided by interwar Austria. Although a large number of Austrian debt titles were traded in financial markets, clean measures for the financing costs of the new republic are difficult to come by.

The first major bond issued by the Republic of Austria in 1923 was under the auspices of the League of Nations. It carried a joint guarantee by the United Kingdom, France and five other countries, and as a result its yields reflected the quality of the guarantors and the guarantee arrangements, rather than an Austrian level of interest and risk. In addition, the League loan included a conversion option at par after the year 1933. With the initial bonds issued at high nominal rates, it soon came to be expected that Austria would invoke the clause and prices became glued at par. It was only in 1930, ie 12 years into the life of the new republic, that Austria managed to float an unguaranteed loan in the international markets. Possible alternatives to these major bonds, like the non-guaranteed bonds floated by cities and regions during the 1920s, suffer from *low volumes and rare or missing price information*.

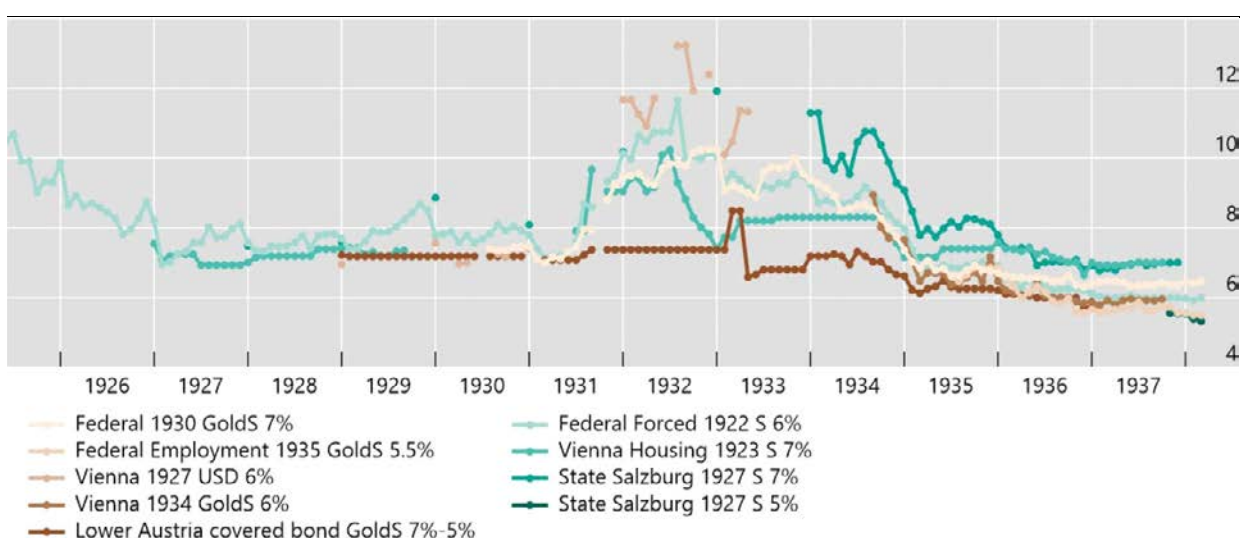
A further issue was *uncertainty on the future conditions of debt service*. With Austria imposing capital controls in the wake of the 1931 financial crisis and the United States leaving gold in 1933, it was not clear whether payments on hard currency loans would be made in schillings and US dollars according to pre-1931 gold parities or not. This uncertainty was not removed until Austrian legislation in 1933 reinstated payment on the schilling tranche at the old parity, while a 1935 ruling by the US Supreme Court, which upheld Roosevelt's cancelling of all gold clauses, allowed Austria to service the US tranche in devalued US dollars. As a result, prices between 1931 and 1935 reflect political and legal uncertainties – some even extraneous to Austria – rather than underlying economic and financial conditions, and even less the marginal financing costs for the Austrian government.

The point is not that it is impossible to get a sense of credit conditions. Rather it is that it is impossible to distil this into one unambiguous and comparable statistic. Graph II.2 shows the yields on selected non-guaranteed bonds issued by the federal government, the regions and the city of Vienna. Some of the bonds were denominated in domestic currency (schillings), some in goldschillings and some in dollars. Long stretches of stale or missing prices (eg Vienna Housing or Vienna 1927 USD) testify to the low liquidity of some of these bonds. Due to their idiosyncratic features, the cross-sectional range of rates is quite considerable. At the same time, the sharp increase in 1931 (financial crisis and default) and the return to more normal conditions in 1934–35 are visible in all series. Note also that, for some time windows, special features of the bonds such as guarantees and currency clauses can be exploited to assess expectations in financial markets concerning eg the future course of relevant macro-financial statistics such as exchange rates (on the 1931 crisis see eg Schubert (1991) and Marcus (2018)).

### Yields to maturity on non-guaranteed bonds

In per cent

Graph II.2



Source: Jobst (forthcoming).

### *Case 3. The problem of embedded options*

Embedded options also deserve emphasis, given their historical prevalence and the fact that they have drawn much attention in the past. Of particular interest is the case of the right for the issuer to convert a given bond at will, after a certain date. Historically, many leading securities and government benchmarks displayed this “redemption option”. This was the case for the majority of the international bonds issued during the 19th century (Flandreau and Zumer (2004)). In fact, even British perpetual consols had such clauses (Harley (1976), Klovland (1994)), as well as various French issuances during the 19th and 20th century (Vaslin (1999), Oosterlinck et al (2016)). One rationale may have been that, *ex ante*, when the bond was issued and exercising the option was unlikely (or unlikely to take place any time soon), the value of the option right (and thus its cost to the issuer) was very small. This enabled the borrower to secure, at a very small price, the opportunity for more substantial savings later. It may also be seen as an incentive mechanism: countries with such clauses in their debt contracts were encouraged to improve their credit. Finally, this may have been a signalling instrument, through self-selection.

Such a redemption/conversion right, if it is embedded in the bond that is being used for a benchmark, is bound to affect the measure of the interest cost. The price of the bond is the sum of the price of the “vanilla” debt asset and the price of the option to redeem. As long as the asset price is far from the strike price, the value of the option is small. But it becomes relevant at some point. The treatment of conversions raises many significant issues, including that they are not random. Being exercised at the discretion of the borrower (the government) conversions result from both debt service minimisation considerations and also from political economy considerations, since a conversion operates a transfer from creditors to tax payers.<sup>27</sup>

When the option is “in the money”, one will observe a flattening out of the yield series (if computed naïvely by ignoring the redemption option). Computed in this way, the yield series becomes a very poor indicator of long-term interest rates.<sup>28</sup> The problem of correcting yields for the effect of options is separate from the discussion of anachronism above. “Modern” option pricing theory does not provide clear guidelines as to what should be done because it only applies to certain types of option and typically requires further assumptions, which are often not warranted in practice. For instance, the Black and Scholes model refers to so-called “European” options (options that can be exercised only on the expiration date), when historical bond conversion options could in most cases be exercised at any time. Moreover, the Black and Scholes model assumes that the price of the underlying assets follows a geometric Brownian motion with constant drift and volatility (Black and Scholes (1973)). Using it in contexts in which the geometric Brownian motion assumption is not supported is a serious stretch. Finally, the extent to which this model even describes what is going on in the market today has been questioned (MacKenzie (2006)).

<sup>27</sup> For an illustration of the fact that conversions are not correctly understood, see for instance Reinhart and Rogoff (2008, Appendix Table 3, p 40) which treats British conversions in the 19th century as an “episode of domestic debt default or restructuring”, when in fact the option to convert was part of the original financial contract.

<sup>28</sup> See Flandreau and Zumer (2004), p 24 for a discussion of this point.

The same applies, of course, to the historical treatment of options as well. As is well known, mathematical approaches to option theory go back in time (Lefèvre (1873), Jovanovic (2006), Hafner (2009)). Likewise, some broad, intuitive features of options were well understood by previous writers. A case in point is the positive association between the underlying volatility of the asset and the price of the option discussed in Pinto (1771). Nogués-Marco and Vam Malle-Sabouret (2007) provide evidence of an awareness among 18th-century investors of features of option pricing that are emphasised in option theory. Yet in the absence of a modern consensus, it is unclear which approach should be preferred.

In the existing literature, we can identify two main ways to tackle the difficulty. One is to model conversion risk using modern option theory, even if its restrictions typically do not hold in the context under study. An illustration is the discussion by Chamley (2011) about conversion risk in 18th century England. Because conversions reduce creditors' income, the optimal strike price is the result of extraneous strategic interactions. Assuming that the Black-Scholes formula applies, the author is then able to extract conversion expectations, which is a creative way of investigating the political process.

In practice, if the focus is on yield extraction, a more prosaic alternative may be preferred. It consists of using the information from another yield series during the period when the value of the conversion option cannot be neglected. A good illustration is the Goschen conversion of British consols in 1888 (Harley (1976), Capie and Webber (1985), Klovland (1994)). For example, Klovland deftly exploits information in alternative safe bond prices to reconstruct a convincing government yield series that controls for conversion risk. This is, in essence, an example of the benchmark approach discussed earlier.

Box II.3 discusses this episode along with two other examples, including the Austrian conversion of 1893 and the Italian rendita conversion in 1906. One takeaway is that bonds with lower coupons than the bond that is "in the money" provide a valid alternative as they are less likely to be profitably converted any time soon, implying that the value of the conversion option can be neglected.

## Dealing with conversion risk: three cases

Conversions are historically frequent events. Graph II.3 (first panel) illustrates the simplest case with the conversion of the Italian rendita in 1906. The conversion had been considered from 1900 onwards, preventing the price of the rendita from rising above 100 and yields from falling below 4% (net of tax). Comparing the series with another bond that was not subject to conversion risk, the 3% gold bond "Vittorio Emmanuele", shows that the yields on it and the 5% bond were almost the same as long as no conversion was in sight (ie around 1896–99). After 1900, however, the yield on the 3% declined while the threat of an imminent conversion prevented any fall of the yield on the 5%. After the conversion was completed, yields again coincided.

Graph II.3 (second panel) shows a similar example, the conversion of the Austrian 5% Notenrente into the 4% Kronenrente in 1893. In this case again, an alternative is available in the form of the 5% (after tax 4.2%) unified debt. The example, however, also exposes the limits of the approach. Note that, while yields on the 5% Notenrente and the 4.2% unified debt were glued together before 1890, this is no longer true for the unified debt and the 4% Kronenrente after 1893. A reason could be that by 1893 interest rates had declined to such a level that investors now expected a conversion of the 4.2% unified debt as well.

Finally, the yield on British consols shown in Graph II.3 (third panel) illustrates the case of a conversion that did not take place. The price of the 2.5% consol was exceptionally high between 1894 and 1899. Had the price persisted at that level, it would have been very likely that the government would convert the bond at its earliest convenience, ie in 1923. Harley (1976) therefore argues that the consol cannot be considered to be a perpetual but rather a bond maturing in 1923. The resulting yield is close to 2%. Based on yields on other securities not subject to the risk of conversion or early repayment, Klovland (1994) counters that investors did not expect that interest rates would stay at their low level until 1923 and thus attached a low probability to a conversion of the 2.5% consols. Instead, he argued that yield calculations should persist with the assumption of a perpetual, meaning that interest rates were not as low as implied by Harley.

The general advice is thus to always be careful, when (i) the price of a bond is close to par and (ii) when splicing two bond series with different coupons (eg the 5% and the 4% Austrian Kronenrente) as, most of time, the earlier, higher coupon bond has already not been representative for a certain period beforehand.

### Yields when bonds are subject to conversion risk

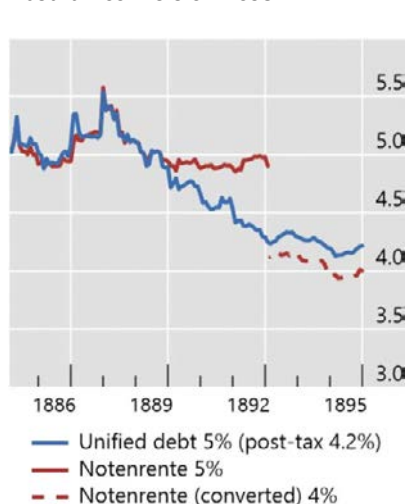
In per cent

Graph II.3

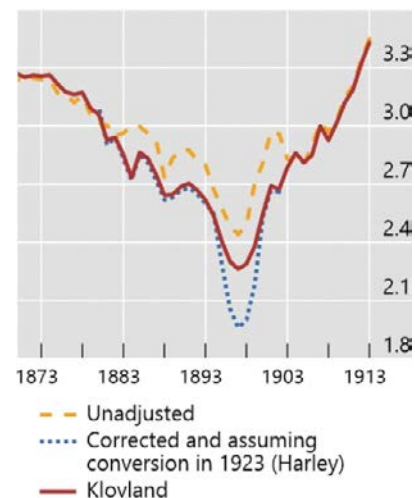
Conversion of the Italian rendita 5% in 1906



Austrian conversion 1893



A conversion that did not happen  
British consols. different estimates



Sources: Flandreau and Zumer (2004); Jobst (forthcoming); Klovland (2004) and sources cited therein.



#### *Case 4. When bonds become special ex post*

Even simple bonds that appear straightforwardly structured *ex ante* can become “special” when external events *that were not part of the initial contract* change the *ex post* payment flows associated with the bond. Such events might include a change in taxation, a currency reform or a default. While such events are relatively straightforward to incorporate as soon as they have taken effect, the challenge is to deal with the period of uncertainty before the change takes place. This can be illuminated in the context of post-default negotiations. For instance, creditors may expect the government to reverse its decision, so that the eventual payout would be better than the one promised. In view of this, what future coupon ought to be used?

An illustration is provided by the US government’s lifting of the gold clauses on all USD-denominated debts in 1933 and the parallel devaluation of the USD (Edwards (2018)). If such a decision is unexpected, tracing the impact on yields is not problematic. But as soon as we look more closely, things complexify. Between 1933, when the gold clauses were lifted, and 1935, when the Supreme Court upheld the government decision, a main source of uncertainty was whether the Supreme Court would revoke the government decision and force a return to payments in gold. Note that this uncertainty affected not only US government debt but *all* foreign debts previously issued in New York and denominated in US dollars (to the extent that creditors could find a way to enforce the Supreme Court decision in respect of foreign debt). An open question is whether the uncertainty was also passed on to the economy or not in terms of more volatile borrowing terms.

Another related situation results from implicit or explicit insurance or indexation which may or may not be spelt out in the debt contract. History provides a number of examples. One is Greece in the early 20th century. It issued a foreign bond whose performance was indexed to the performance of the exchange rate (Flandreau and Zumer (2004)). In this case, while the “bonus” was explicit in the clauses, it is impossible to calculate *ex ante* with full certainty. Another example is provided by Austria in the 1970s, which voluntarily increased the coupon on existing debt to compensate debt holders for the unexpected increase in inflation, even though the bond was not inflation indexed (Kernbauer (2017)).

One important takeaway is that, in practice, bonds often exhibit features (implicit or explicit) that make them more comparable with stocks and complicate the process of making an inference as to the interest rate. Sometimes this arises from conditions in contracts. Often, it is voluntary and occurs despite contracts: for instance, as in the Austrian case above, because governments want to build or maintain reputation. Such decisions may be motivated by political economy considerations, including concerns over the stability of the financial sector. In summary, history suggests that the outlook is more complicated than it might initially appear.

### 2.3 Short-term rates

Short-term rates reflect the opportunity cost of (non-interest bearing) money. Historically, just as in the case of long-term debt, trading in short-term instruments was permitted by the development of innovations that minimised information asymmetries, reducing the need for investors to engage in active research of the underlying characteristics of products, of the markets where they were traded or of the intermediaries involved (Gorton (2017)).

Over time, a variety of instruments emerged, both secured and unsecured, publicly and privately issued, as well as derivative products associated with them. Just as in the case of long-term debt, the resulting innovations were the result of local government, central government or private sector initiatives, though the private sector seems to have played a more important role in short-term debt. The fact that usually there were many relevant debtors makes the question of what constitutes a benchmark much more complicated a priori than in the case of long-term rates where the relevant debtor might just be the government. There are also other differences across countries. Unlike central government debt, which did exist and was actively traded in many political entities for which general statistical information is available, there was significant diversity in money markets across countries and over time.

As a result, there is no clear first-best benchmark and it is necessary to discuss a broad range of possible instruments. Each market is specific, and to make sense of the prices (interest rates) quoted, it is important to figure out its *modus operandi*. Even in the modern period, risk, eligibility and market structure can suddenly matter in a big way, as the divergence between unsecured, secured, overnight index swap and Treasury bill rates during the 2008 crisis well testified (BIS (2008, p 12)). After looking at several important market rates we will return to central bank rates below.

### Unsecured money market rates

As we found for long-term private rates before, one problem is that, by construction, there are as many individual rates as there are private agents, each reflecting individual risk. Historically, this has created a demand for benchmarks, which by aggregating information in the market did help to overcome asymmetric information and allow bilateral transactions to take place. In the case of interbank rates, they enable banks to hedge changes in their funding costs. Interbank benchmark rates (“lbors”) emerged in the 1960s, governed by groups of large banks and have been based, typically, on a poll of participating banks. A recent development has been the involvement of government institutions and regulators following the Libor scandal (Duffie and Stein (2015)) and the move to benchmarks based on actual transactions such as Sofr, €STR and Sonia.

Allowing for a bit of conceptual flexibility, such interbank-type lending goes far back in history. It has existed since time immemorial in the shape of “bills of exchange” that were traded in active money markets in European cities as far back as the middle ages. Bills of exchange were a promise to pay in a given centre at a given time. Trading in such instruments was permitted by the development of a technology (the “accepting” of the bill) that minimised information asymmetries by relying on recognised and identifiable signatures, thus reducing the need for investors to engage in active research of underlying characteristics (De Roover (1953)). Many of these bills originated in international trade, but not exclusively. One initial reason for “exchange markets” to provide the default asset for money markets was usury regulations. The charging of interest was in fact prohibited, leading bankers to develop an alternative technology involving money market “swaps” (bills of exchange), whereby money in the here and now was traded for money abroad in the future. That means if quotations for bills of different maturities are available, then implicit interest rates can be derived from their prices (see Flandreau et al (2009)). Another possibility is to calculate the return from an exchange operation followed at maturity by the converse operation, although since the future “return” exchange rate is not known, the resulting interest rate is uncertain.

While for a long time interest rates were hidden in the exchange rate of foreign bills, in the 19th century formal domestic interest rates started to show up regularly in the financial press, either through verbal comments or in the shape of published tables. In truth, however, we do not really know how such rates, as reported in 19th and early 20th century newspapers, were established. The most likely interpretation is that, in well organised markets, de facto interbank lending existed, and this rate was communicated to correspondents, who could use it to price exchange contracts.

A well known example is *The Economist*, which reported, from the 1860s onwards, what it described as “open market” rates, along with the central bank lending rates for a number of lending centres/countries (Graph A). As the language implies, these open market rates were essentially interbank rates, that is, they are the rates at which an international draft could be drawn competitively in a given centre – as opposed to the “bank rate” which was the rate at which the central bank would take first-grade eligible paper. While it is not possible to *exactly* match the rates reported by *The Economist* with figures for local markets when they are available, the resemblance is generally noticeable. In the absence of centralised price setting – in the absence of an lbor – the journal must have either relied on some aggregation of individual prices or, alternatively, used a correspondent who tracked relevant financial institutions and reported its own, concocted, lbor.

Short-term interest rates in *The Economist*

Graph A

The following are the rates of discount on the principal exchanges of the Continent, showing no alteration beyond an easier tendency at Paris. The rate at St. Petersburg continues heavy :—

	Bank Rate. Per cent.	Open Market. Per cent.
Paris .....	4	3½ ¾
Vienna .....	5	5
Berlin .....	4	3½
Frankfort.....	3	2¾
Amsterdam.....	3	3
Turin .....	5	5
Brussels .....	3	2¾
Hamburg.....	...	2½
St Petersburg.....	6	8

Source: *The Economist*, 29 August 1863, p 965.

In general, sources of which we are aware provide some indications that help understand the nature of what one is dealing with, according to origination, type, quality or eligibility for central bank refinancing operations. For instance, one finds mentions that private short-term interest rates quoted in contemporary sources corresponded to “high-quality paper”, which makes sense as it suggests that contemporary journalists were, like us, concerned with distilling market-wide trends by focusing on instruments for which idiosyncratic risk was smaller.

Another question is maturity. The more sophisticated London market did produce a differentiated set of rates corresponding to differentiated maturities. For instance, in the second half of the 19th century, London market interest rates, as documented by *The Economist*, included rates for 30 to 60 days; three, four and six months; as well as six-month bankers' drafts and trade bills. In other markets less

detail is available, generally because less specialisation existed. In France, for instance, the private interbank rate, known as the “taux d’escompte hors banque” was reported by the French counterpart to *The Economist*; *l’Economiste français*. The rate it reported, essentially the same as the rate documented by *The Economist* for France under “Open Market” in Graph A above, consisted in one rate only, which was applied to any short maturity. In fact, this interbank money market was less important than the repo (“report”) market in which banks lent funds to the stock exchange. This explains the cruder statistical outlook (Dessirier (1929), Flandreau and Sicsic (2003)). For the US, the absence of anything close to a benchmark short-term rate is well documented for the 19th century (Siegel (1992), Goetzmann and Ibbotson (2008)).

For researchers collecting such data, another question to bear in mind is the one that has been relevant to the modern debate over the most appropriate benchmark rate, namely whether the quotes in the financial press were based on actual transactions or rather on a poll of most important market participants on what is the “prevalent rate”. Also, researchers need to secure a solid understanding of why benchmark rates start to appear when they do. Generally, there is a reason, which is worth identifying. Such knowledge helps to better understand the data available by fostering a cleaner grasp of the micro-structure of the underlying market.

A final remark has to do with liquidity. When it comes to short-term debt, liquidity presents itself in a slightly different guise than for long-term debt, in that it is tied to the greater or lesser extent to which a given currency internationalises. The development of a short-term money market is heavily influenced by agglomeration and path dependency. Because some markets took the lead in producing safer forms of the denomination, they also tended to internationalise. In 1900, a merchant in, say, Saint Petersburg seeking to invest in liquidities, would have been able to invest in Russian short-term instruments, but alternatively he could hold bills in London, Paris or Amsterdam. If these latter markets were more liquid, he would have preferred to take advantage of them. As a result, illiquidity in money markets is the (endogenous) result of processes of currency/money market polarisation according to centre/periphery delineations (Flandreau and Jobst (2009)).

The upshot is that the money markets of “peripheral” currencies tend to be shallower, other things being equal. This ensures that interest rates in peripheral markets tend to be structurally higher on account of structurally lower liquidity (Bordo and Flandreau (2003)). The outcome raises difficult questions as to what constitutes a proper benchmark in a shallow market. In a context of financial globalisation, what should be considered the relevant short-term interest rate? Is it the high local rate on an illiquid instrument? Or is it the lower foreign rate on a highly liquid international deposit to which domestic agents have access? Better understanding of these markets and of their operation is a topic for continuing investigations and it warrants the interest of HMFS researchers, especially as the circle of countries it covers starts to broaden.

### Secured money market rates – repo rates

A repurchase agreement (repo) is an agreement to sell securities (referred to as “collateral”) at a given price, coupled with an agreement to repurchase these securities at a pre-specified price at a later date. Such markets have a long history, an early account being Pinto (1771). Repos have existed under a number of names such as “continuations” in the London capital market, “reports” in the Paris capital market etc. A repo is economically similar to a collateralised loan since the securities provide

credit protection in the event that the seller (ie the cash borrower) is unable to repurchase the security. Haircuts and margin calls provide additional layers of protection for the cash lender.

Repo rates play a significant role in modern markets, but research has demonstrated that they have been important since the 18th century (Flandreau and Sicsic (2003), Koudijs and Voth (2016)). From the point of view of HMFS, repo rates are attractive as the implicit collateralisation potentially reduces the idiosyncrasies that affect the interpretation of unsecured money market rates. At the same time, it is critical to understand and document the manner in which they were traded, secured and priced, especially regarding margin calls (“haircuts”). Another feature is the existence of central clearing counterparties. In some historical setups, repos were priced centrally, in others individually (ie “over the counter” or per security) and exhibited variations across securities. A possible explanation is that in such systems the main use of the repo market was the borrowing and lending of specific securities – rather than the borrowing and lending of cash – using (non-specific) securities as collateral. This is mirrored today in the distinction between general collateral (“GC”) and specific repo rates.

Having received only limited attention in the literature in the lead up to the Great Financial Crisis of 2007–09, the “run on repo” at the centre of the crisis has sparked interest among economic historians in earlier manifestations of repo markets. As a by-product, the better understanding of repo markets and their interaction with other segments in the money market could allow us to actually use repo prices as measures for conditions in the broader money market. For HMFS this is a highly promising area of research.

### Rates on short-term government debt (treasury bills)

Treasury bills and other types of short-term government debt enjoy the advantage of being issued by a single debtor of known credit quality. The need to balance the seasonality in tax receipts, payments on debt and other expenditures mean that, even early on, governments had incentives to issue short-term paper alongside long-term bonds. But in some instances, short-term bonds with legal tender status provided a hybrid between government money and government debt.

When they became the reference for short-term rates, government bills were typically issued at short and regular intervals, so that enough paper with a remaining maturity of eg one or three months would be available. To reflect funding conditions in the money market, they also needed to be traded actively or at least to be auctioned off through some competitive procedure. Individual countries’ experiences have been highly idiosyncratic in this respect. There is evidence of an active UK market in the 19th century, and of a Russian one. In the 20th century, the US government started to issue bills regularly in 1929 (Siegel (1992), Garbade (2008)).

In other countries, treasury bills faced a mixed fate, such as for example in Norway, where during the 1940s they briefly provided a representative rate before disappearing and then reappearing only in the mid-1980s, though with limited liquidity until the 1990s (Eitrheim et al (2007)). A final aspect to consider is the possibility that options were incorporated in short-term government debt. These might include the possibility of using them to pay taxes, of cashing them at par, or converting them into long-term government debt, an option for early repayment by the treasury or a possibility to roll them over at maturity. Such features, which existed

far back in time, have the potential to significantly affect the interpretation of prices and the calculation of yields.<sup>29</sup>

### Central bank policy rates

The final short-term rate we consider is the central bank policy rate. For those countries with long-established central banks, it is often one of the first short-term rates to become available on a regular basis. The central bank policy rate – if it is unique – captures the condition of the money market. This is due to the usually tight link between central bank operations and price formation in money markets (Jobst and Ugolini (2016)). Modern central banks, seeking to influence broader economic outcomes, typically operate in liquid segments of the money market. Consequently, they serve as a benchmark for money markets and the economy at large. At the same time, thanks to the central bank's size and its frequent role as marginal provider or absorber of funds, the central bank is a key player and its choice of instruments is a key factor in directing liquidity.

As a result, analysis of central bank operations is doubly useful. First, it points to which markets are relevant, even though at times central banks' active promotion of targeted markets may not prove successful. The failure of the early Federal Reserve to establish a market for bankers' acceptances in the 1930s (James (1995)), or the difficulties faced by some European central banks to set up liquid markets in treasury bills in the 1980s (Forssbæck and Oxelheim (2007)) bear witness to this. At other times central banks fell behind, continuing to intervene in markets that had lost some of their earlier allure. In these cases, much depends on how integrated different segments of the money market continue to be (Brousseau et al (2013)). Over the medium run, however, central bank operations and money market functioning should coincide "spatially".

Secondly, as argued before, the central bank policy rate should be highly informative about money market conditions, because of the central bank's role as the marginal provider of central bank money. To understand how the policy rate relates to market rates, it is essential to document the framework in which the central bank implements its market interventions. As the central bank operates in potentially idiosyncratic instruments (like bills of exchange) understanding the meaning of the central bank rate requires knowledge of these underlying instruments. Central bank archives contain a wealth of information on the kind of assets that are eligible at the central bank and whether there is any evidence of the central bank actually applying the rules.

Central bank policy rates themselves can interact in different ways with market rates. They can lie close to market rates or provide a floor or a ceiling. When a central bank sets not one but several policy rates (eg a *refinancing* and a *deposit facility rate* or differential rates by types of eligible assets or counterparties) discussing which is the marginal rate is crucial. Note that in different countries and in the same country over time, the same term might refer to different concepts. The Bank of England's "Bank Rate" has been a discount rate on bills, a secured lending rate and a rate paid on commercial bank deposits with the central bank, in the course of its long history. Finally, because of non-price measures (such as credit rationing in the presence of usury law and reserve constraints) it is likely that for extended periods of time, there

<sup>29</sup> For an example, see the use of East India bonds to tease out the short-term interest rate in Mirowski and Weiller (1990) and the discussion in Nogués-Marco and Vam Malle-Sabouret (2007).

will be a need to accompany the produced series with notes warning users of the limited informative power of prices in a regime of quantitative constraints. To conclude, while easily accessible (at least in comparison to market rates) and potentially highly informative, central bank rates also require a significant effort in documentation.

### Other relevant rates

The rates discussed so far in this section reflect benchmark rates that can be standardised. There exists a set of other rates for which standardisation appears much more challenging but which are important for the HMFS initiative as they represent the rates faced by many households and companies either as borrowers or savers. Some of them (mortgage rates and corporate bond rates) have already been mentioned above. But we can also think of various *deposit rates* (typically uncovered from the archives of commercial and savings banks), as well as other borrowing rates such as those relevant for consumer loans, pawn rates etc. All may convey highly useful information, especially about the cost of borrowing during financial crises and the influence on house prices. While the higher level of idiosyncrasy compared with benchmark government bond yields or the interbank short-term rate suggests they are more difficult to collect, aggregate or combine, they may be profitably investigated and integrated in international panel studies. So they are an important part of the HMFS initiative. But their economic significance depends on “spreads” – their prices relative to safe rates. So the development and understanding of historical rates on low-risk benchmark rates is the first priority.

## 2.4 Conclusion

The key message of the chapter is that safe (or, more correctly, low-risk) assets and their prices cannot be understood without a clear grasp of how (and by whom) such assets were designed, produced and maintained. The design of a specific historical instrument, in turn, responded to the demands of the economic actors as well as the constraints in the given historical setting, which could have been technological, informational, social, legal or political. As we have seen, the derivation of relevant yields or interest rates hinges on the achievement of a correct understanding of underlying instruments and markets as well as the political, social and legal contexts in which they operated.

The chapter has emphasised two points. First, the need to consider all the many relevant characteristics of financial instruments when calculating and reporting interest rates and yields – what we have termed their “pedigree”. Second, the need to combine the publication of any time series with the bibliographical directions and explanatory comments that enable users to gauge critically the relevance of the selected assets as well as the way in which information was recorded at the time and in what form it has survived. At the heart of the chapter is the critical importance of background research. The creation of a knowledge infrastructure is essential for both producing and understanding any consumable time series.

Two practical conclusions can be drawn at this stage.

First, most (if not all) of the interest rates that may appear easy to collect because, say, they come in ready-made retrospective tables produced by some deceased, yet cooperative, institution or researcher, or were published in newspapers from where they can be retrieved, are generally less straightforward than they appear. Their

documentation and interpretation require diligence, care and a thorough understanding of the institutional setup. In many cases, this task cannot be outsourced. It does require intermediary steps and research papers devoted to discussing context, hurdles and alternative solutions.

Second, the chapter also points to a silver lining. Given that if one is to approximate the ideal, an important investment is necessary anyway, much more of the information available in surviving sources could be exploited. In other words, the relative cost of mining information from the deeper veins is not so high, *especially given the potential for further discoveries*. The case of interest rates illustrates this principle. In the chapter we have listed a number of directions that could profitably be pursued, from mortgages and non-sovereign bonds to medium-term debt, repos and implicit short-term rates. Together with properly derived series on government debt and well documented high-quality short-term rates, such data series would not only provide consumers with greater security but they could also help scholars on the research frontier to create a more encompassing and complex understanding of past (and present) financial systems. In fact, it is the creation of this background research infrastructure, which is the precondition for posing meaningful questions to our past – including those which policymakers are tasked to answer.

## References

- Accominotti, O, M Flandreau, R Rezzik and F Zumer (2010): "Black man's burden, white man's welfare: control, devolution and development in the British Empire, 1880–1914", *European Review of Economic History*, vol 14, no 1, April, pp 47–70.
- Allen, W (2019): *The Bank of England and the government debt. Operations in the gilt-edged market, 1928–1972*, Cambridge University Press.
- Almeida, H, I Cnha, M Ferreira and F Restrepo (2017): "The real effects of credit ratings: the sovereign ceiling channel", *The Journal of Finance*, vol LXXXII, no 1, pp 249–90.
- Bank for International Settlements (BIS) (2008): *Quarterly Review*, December.
- Black, F and M Scholes (1973): "The pricing of options and corporate liabilities", *Journal of Political Economy*, vol 81, no 3, pp 637–54.
- Bordo, M and M Flandreau (2003): "Core, periphery, exchange rate regimes, and globalization", *Globalization in historical perspective*, NBER and Chicago University Press.
- Bordo, M and J Haubrich (2008): "The yield curve as predictor of growth: long-run evidence, 1875–1997", *The Review of Economics and Statistics*, vol 90, no 1, pp 182–5.
- Borensztein, E, K Cowan and P Valenzuela (2013): "Sovereign ceilings 'lite'? The impact of sovereign ratings on corporate ratings", *Journal of Banking & Finance*, vol 37, pp 4014–24.
- Brousseau, V, A Chailloux and A Durré (2013): "Fixing the fixings: what road to a more representative money market benchmark?", *IMF Working Paper*, vol 13, no 131.
- Calomiris, C (1992): "Greenback resumption and silver risk: the economics and politics of monetary regime change in the United States, 1862–1900", *NBER Working Papers*, no 4166.
- Capie, F and A Webber (1985): *A monetary history of the United Kingdom, 1870–1982: data, sources, methods*, London, Allen & Unwin.



- Chamley, C (2011): "Interest reductions in the politico-financial nexus of eighteenth-century England", *The Journal of Economic History*, vol 71, no 3, pp 555-89.
- Chandler, A (1981): *Henry Varnum Poor: business editor, analyst, and reformer*, New York, Arno Press.
- Chavaz, M and M Flandreau (2017): "'High & dry': the liquidity and credit of colonial and foreign government debt and the London Stock Exchange (1880-1910)", *The Journal of Economic History*, vol 77, no 3, pp 653-91.
- Chen, L, D Lesmond, and J Wei (2007): "Corporate yield spreads and bond liquidity", *The Journal of Finance*, vol 62, no 1, pp 119-49.
- De Roover, R (1953): *L'évolution de la lettre de change, 14e-18e siècles*, Paris, Armand Colin.
- Dessirier, J (1929): "Nouveaux essais de statistiques sur les mouvements boursiers", *Journal de la Société de Statistique de Paris*, vol 71, no 2, pp 35-94.
- Dickson, P (1967): *The financial revolution in England: a study in the development of public credit 1688-1756*, London, Macmillan.
- Duffie, D and J Stein (2015): "Reforming Libor and other financial market benchmarks", *Journal of Economic Perspectives*, vol 29, no 2, pp 191-212.
- Dunne, P, M Moore and R Portes (2007): "Benchmark status in fixed-income asset markets", *Journal of Business Finance & Accounting*, vol 34, no 9, pp 1615-34.
- Edwards, S (2018): *American default: the untold story of FDR, the Supreme Court, and the battle over gold*, Princeton University Press.
- Eitrheim, Ø, J Klovland, and J Qvigstad (2007): "Historical monetary statistics for Norway, Part II", *Norges Bank Occasional Paper*, no 38.
- Epstein, S (2000): *Freedom and growth: the rise of states and markets in Europe, 1300-1750*, London, Routledge.
- Esteves, R and A Tunçer (2016): "Feeling the blues: moral hazard and debt dilution in Eurobonds before 1914", *Journal of International Money and Finance*, vol 65, pp 46-68.
- Fabozzi, F (2001): "The structure of interest rates", in F Fabozzi (ed), *The handbook of fixed income securities*, sixth edition, McGraw Hill, pp 131-54.
- Felloni, G (2010): "A profile of Genoa's Casa di San Giorgio, a turning point in the history of credit", *Rivista di storia economica*, vol 3, pp 335-46.
- Flandreau, M (2006): "Home biases, nineteenth century style", *Journal of the European Economic Association*, vol 4, no 2/3, pp 634-43.
- Flandreau, M and C Jobst (2009): "The empirics of international currencies: network externalities, history and persistence", *Economic Journal*, vol 119, no 537, pp 643-64.
- Flandreau, M, C Galimard, C Jobst and P Nogués-Marco (2009): "The bell jar: commercial interest rates between two revolutions 1688-1789", in J Atack and L Neal (eds), *The origins and development of financial markets and institutions*, Cambridge University Press, pp 161-208.
- Flandreau, M and G Legentilhomme (2022): "Cyberpunk Victoria: the credibility of computers and the first digital revolution (1848-1883)", *Economic History Review*, vol 75, no 4, pp 1083-119.

- Flandreau, M and K Oosterlinck (2012): "Was the emergence of the international gold standard expected? Evidence from Indian Government securities", *Journal of Monetary Economics*, vol 59, no 7, November, pp 649–69.
- Flandreau, M and P Sicsic (2003): "Crédits à la spéculation et marché monétaire: le marché des reports en France de 1875 à 1914", in O Feiertag and M Margairaz (eds), *Politiques et pratiques des banques d'émission en Europe (XVIIe-XXe siècles)*, Paris, Albin Michel, pp 197–222.
- Flandreau, M and F Zumer (2004): *Making of global finance*, Paris, OECD.
- Flores, J and Y Decorzant (2016): "Going multilateral? Financial markets' access and the League of Nations loans, 1923–8", *Economic History Review*, vol 69, no 2, pp 653–78.
- Forssbæck, J and L Oxelheim (2007): "The interplay between money market development and changes in monetary policy operations in small European countries, 1980–2000", in D Mayes and J Toporowski (eds), *Open market operations and financial markets*, London, Routledge, pp 120–52.
- Fratianni, M and F Spinelli (2006): "Did Genoa and Venice kick a financial revolution in the quattrocento?", *Oesterreichische Nationalbank Working Paper*, no 112.
- Garbade, K (2007): "The emergence of "regular and predictable" as a treasury debt management strategy", *Federal Reserve Bank of New York Economic Policy Review*, March, pp 53–71.
- (2008): "Why the US treasury began auctioning treasury bills in 1929", *Federal Reserve Bank of New York Economic Policy Review*, July, pp 31–47.
- Gille, B (1965): *Histoire de la maison Rothschild*, vol 1, Geneva, Droz.
- (1967): *Histoire de la maison Rothschild*, vol 2, Geneva, Droz.
- Goetzmann, W and R Ibbotson (2008): "History and the equity risk premium", in R Mehra (ed), *Handbook of the equity risk premium*, Amsterdam, Boston, Elsevier, pp 515–29.
- Goetzmann, W and K Rouwenhorst (2005): *The Origins of value: the financial innovations that created modern capital markets*, Oxford University Press.
- Gorton, G (2017): "The history and economics of safe assets", *Annual Review of Economics*, vol 9, no 1, pp 547–86.
- Hafner, W (ed) (2009): *Vinzenz Bronzin's option pricing models: exposition and appraisal*, Berlin, Springer.
- Hannah, L (2018): "The London Stock Exchange, 1869–1929: new statistics for old?", *Economic History Review*, vol 71, no 4, pp 1349–56.
- Harley, K (1976): "Goschen's conversion of the national debt and the yield on consols", *Economic History Review*, vol 29, no 1, pp 101–6.
- Heers, J (1961): *Gênes au XVe Siècle*, Paris, SEVPEN.
- Hickman, B (1958): *Corporate bond quality and investor experience*, Princeton, NBER.
- Hoffman, P, G Postel-Vinay and J-L Rosenthal (2000): *Priceless markets, the political economy of credit in Paris*, University of Chicago Press.
- Homer, S and R Sylla (1991): *A history of interest rates*, third edition, New Brunswick, Rutgers University Press.

- James, J (1995): "The rise and fall of the commercial paper market, 1900–1930", in M Bordo and R Sylla (eds), *Anglo-American financial systems*, New York, Irwin, pp 219–60.
- Jobst, C (forthcoming): Historical credit and interest rate series for Austria, 1816–1998.
- Jobst, C and S Ugolini (2016): "The coevolution of money markets and monetary policy, 1815–2008", in M Bordo et al (eds), *Central banks at a crossroads: what can we learn from history?*, Cambridge University Press, pp 145–94.
- Jovanovic, F (2006): "Economic instruments and theory in the construction of Henri Lefèvre's science of the stock market", in G Poitras (ed), *Pioneers of financial economics*, vol 1, Cheltenham, Edward Elgar, pp 169–90.
- Kernbauer, H (2017): *Österreichische Geld- und Währungspolitik von 1969 bis 1998*, Wien, Oesterreichische Nationalbank.
- Klovland, J (1994): "Pitfalls in the estimation of the yield on British consols, 1850–1914", *The Journal of Economic History*, vol 54, no 1, March, pp 164–87.
- Koudijs, P and H Voth (2016): "Leverage and beliefs: personal experience and risk taking in margin lending", *American Economic Review*, vol 106, no 11, pp 3367–400.
- Krishnamurthy, A and A Vissing-Jorgensen (2012): "The aggregate demand for treasury debt", *Journal of Political Economy*, vol 120, pp 233–67.
- Lefèvre, H (1873): "Physiologie et mécanique sociales", *Journal des Actuaire Français*, vol 2.
- MacKenzie, D (2006): *An engine, not a camera: how financial models shape markets*, MIT Press.
- Marcus, N (2018): *Austrian reconstruction and the collapse of global finance, 1921–1931*, Harvard University Press.
- Mauro, P, N Sussman, and Y Yafeh (2006): *Emerging markets and financial globalization: sovereign bond spreads in 1870–1913 and today*, Oxford University Press.
- Meitzen, A (1871): *Der Boden und die landwirthschaftlichen Verhältnisse des Preussischen Staates*, vol 3, Berlin, Paul Parey.
- Mirowski, P and K Weiller (1990): "Rates of interest in 18th century England", *Explorations in Economic History*, vol 27, pp 1–28.
- Moody, J (1922): *Government and municipal investments, Moody's analyses of investments and security ratings books*, New York, Moody's Investors Service.
- Nagel, S (2016): "The liquidity premium of near-money assets", *Quarterly Journal of Economics*, vol 131, no 4, pp 1927–71.
- Nogués-Marco, P and C Vam Malle-Sabouret (2007): "East India bonds, 1718–1763: early exotic derivatives and London market efficiency", *European Review of Economic History*, vol 11, no 3, pp 367–94.
- Oosterlinck, K, L Ureche-Rangau and J-M Vaslin (2016): *Politics and the conversion of French sovereign bonds*, Université libre de Bruxelles, mimeo.
- Packer, F, R Stever and C Upper (2007): "The covered bond market", *BIS Quarterly Review*, September, pp 43–55.

Pinto, I (1771): "Tableau ou exposé de ce qu'on appelle le commerce ou plutôt le jeu d'actions en Hollande", *Traité de la Circulation et du Crédit*, Amsterdam, Marc Michel Rey.

Poisson, J-P (1985): *Notaires et société; travaux d'histoire et de sociologie notariale*, Paris, Economica.

Reinhart, C and K Rogoff (2008): "The forgotten history of domestic debt", *NBER Working Papers*, no 13946.

——— (2009): *This time is different: eight centuries of financial folly*, Princeton University Press.

Sattler, F (2022): "German mortgage banks between safety thinking and speculation fever: real estate finance during the building boom of the late 19th century", *Jahrbuch für Wirtschaftsgeschichte/Economic History Yearbook*, vol 63, no 1, pp 105–34.

Schmelzing, P (2020): "Eight centuries of global real interest rates, R-G, and the "suprasecular" decline, 1311–2018", *Bank of England Staff Working Paper*, no 845.

Schubert, A (1991): *The Credit-Anstalt crisis of 1931*, Cambridge University Press.

Schulte, F (1912): *Die Bodenkreditinstitute der Österreichisch-ungarischen Monarchie, 1841 bis 1910*, München, Duncker & Humblot.

——— (1918): *Die Hypothekenbanken*, München, Duncker & Humblot.

Siegel, J (1992): "The real rate of interest from 1800–1990: a study of the US and the UK", *Journal of Monetary Economics*, vol 29, no 2, pp 227–52.

Sieveking, H (1906): *Studio sulle Finanze Genovesi nel Medioevo e in particolare sulla Casa di S. Giorgio*, Atti della Società Ligure di Storia Patria, Genoa, Tipografia della Gioventù.

Snowden, K (2010): "Covered farm mortgage bonds in the United States during the late nineteenth century", *Journal of Economic History*, vol 70, no 4, pp 783–812.

Stasavage, D (2007): "Cities, constitutions, and sovereign borrowing in Europe, 1274–1785", *International Organization*, vol 61, no 3, July, pp 489–525.

Sylla, R (2002): "An historical primer on the business of credit rating", in R Levich, G Majnoni and C Reinhart (eds), *Ratings, rating agencies and the global financial system*, Boston, Springer, pp 19–40.

Vaslin, J (1999): *Le marché des rentes françaises au XIXème siècle et la crédibilité financière de l'état*, Université d'Orléans, unpublished dissertation.

Wooldridge, P (2001): "The emergence of new benchmark yield curves", *BIS Quarterly Review*, December, pp 48–57.

Yuan, K (2005): "The liquidity service of benchmark securities", *Journal of the European Economic Association*, vol 3, pp 1156–80.

### 3. Historical data on credit<sup>30</sup>

Vincent Bignon, Bank of France and Université de Aix-Marseille

Marc Flandreau, University of Pennsylvania and the Wharton School

In its most general definition, credit is the lending of resources – be they commodities or money – to an agent who commits to reimbursement in the future. Due to this intertemporal nature of credit, the lender needs to examine the default risk of the counterparty, and, if needed, enforce repayment.<sup>31</sup> Inside a polity, be it an empire like Babylon or a nation state today (Goetzman (2016)), the legal system in which credit is made has a strong influence on the design of the contractual forms of credit instruments as it is the legal system that ensures that debtors will honour their obligations in due time. In particular, and depending on the legal system in which it is embedded, credit will take different forms, as agents may prefer intermediated or disintermediated forms of borrowing.

Against this background, economists, statisticians and historians take different perspectives when studying the evolution of credit over time. A statistician and a historian – the *statistorian* discussed in the introduction – is interested in drawing up historical credit aggregates. Statistorians will start with a definition of the credit instruments used *in a given society at a given point in time*.<sup>32</sup> This definition involves a judgment on the inclusion or not of the various instruments available, typically depending on the contractual form they take. The act of constructing credit aggregates then consists in the adding up of the values associated with the various credit instruments used in that place and time. The series then tracks modifications in the composition of credit as the popularity of some instruments fades away while other new ones emerge.

In contrast to this bottom-up approach, the point of departure for an empirical economist or a quantitative economic historian is almost the opposite. Starting from a specific question (for example, “what variables cause financial crises or banking instability?”) they construct an empirical test to disentangle fundamental determinants from spurious correlations. A key question is whether conclusions or inferences from the data set are specific to the countries and time period covered by the data or whether they apply more broadly. In particular, because financial (in)stability is endogenous to the type of financial regulation and the policies implemented by public institutions such as central banks (Calomiris and Haber (2014, Bignon et al (2012))), a natural question that arises is the extent to which results are still valid in different institutional environments.

A number of academic studies were published in the 2000s that tested the robustness “over the very long run” of claims such as the relation between the level of government indebtedness or credit growth and the failure of banks or the

<sup>30</sup> We are grateful to the members of HMFS, and in particular to Claudio Borio for generous comments. Detailed comments by Clemens Jobst proved vital in producing this final draft. We also acknowledge important input from external readers and contributors, in particular, Michael Bordo, Patrice Baubeau, Federico Barbiellini-Amidei and Pilar Nogués-Marco.

<sup>31</sup> According to the *Oxford Dictionary of Economics* (Hashimzade et al (2017)) credit is “the system by which goods or services are provided in return for deferred rather than immediate payment”.

<sup>32</sup> Because credit aggregates are usually produced in central banks, the tradition is to restrict attention to claims denominated in money and which has to be reimbursed in money as well.

incidence of financial crises. In order to include more data points, the authors of those studies left the quiet waters of “ready to use” modern international data sets to chart a new course within historical data. But in the process of attempting to assemble historical data sets – building on data constructed long before the age of international harmonisation – they reached those very shores that are familiar to the *statistorian*, in which one runs the risk of wrecking one’s ship on the reefs of measurement issues.

In a further twist of the plot, modern quantitative historical studies (almost always) identify credit with *bank* credit, ie with loans made by institutions that lend money by writing debt contracts that are funded with other people’s money. This may have been inspired by the historical context in which early statistics were constructed. This was after the Second World War, and bank credit reigned supreme in many parts of the world. Yet, as historians knew very well, this had not always been the general case. One need not look further than the years before the First World War to find evidence of a period when other financial intermediaries were active in supplying credit, therefore creating mismeasurement of credit aggregates if contemporary definitions of credit were applied retrospectively. Ultimately, the problems tied to swapping definitions of “credit” and “banks” are explained by changes in the institutional, regulatory and political environment.

Such issues are especially challenging for the study of secular patterns. Any attempt at testing important claims on financial relations with policy-relevant messages by using long-run data runs into the risk of using incomparable data. Long-term shifts in the relation between, say, the credit-to-GDP ratio and financial crises may signal shifts in the threshold of unsustainable indebtedness, but they may also reflect secular changes in the construction of credit aggregates. For example, if the share of banks in the credit market varies over time, as it has for many countries since the 19th century and afterwards, restricting attention to banks may create the spurious impression that the credit-to-GDP ratio has changed when, in fact, the change reflects alterations in the structure of the credit market only. What is more, episodes of credit expansion are typically characterised by structural change, with new forms of lending emerging and concentrating speculative pressures. The consequence is that, absent any methodology to harmonise series across time and countries, the quest for empirical corroboration may well end up returning spurious relations. What appear to be mere measurement issues have a bearing on policy recommendations.

In this chapter, we dwell on these important questions to engage the question of credit measurement in the super long run. We aim to provide building blocks that the *statistorian* can use to correct the loopholes in existing series by mobilising the most representative information available. To that end, we rely on examples drawn from the historical corpus to spell out three types of methodological difficulties faced in the construction of long time series of credit aggregates.

The first section explains why, although this approach is frequently pursued, it is rarely sufficient to rely on publicly available information which originated in reporting activity by supervisory or regulatory bodies or of financial market participants. Though easy to collect, those pieces of information focus on well identified credit markets and, as a result, are likely biased either by changes in the perimeter of the credit market in question – itself reflecting the changing contours of government regulation – or by the very nature of voluntary public disclosure and the incentives that underpin it.

The second section discusses the problem of what we call the intensive margin. It refers to the challenge of obtaining information for the entire population of – in principle – identical financial institutions or instruments (or for a representative sample of this population). We provide an example that shows that the biases from uncounted banks can be huge, even for countries and periods for which one may imagine that matters were settled.

The third section describes the potential biases in credit aggregates coming from not paying enough attention to the extensive margin of credit. As the subprime crisis reminded us (but this is hardly novel), credit is also provided by agents other than banks (in this case, the infamous “shadow banks”) and this may generate financial instability. We therefore explore how wide we have to cast our net, ie what are the financial institutions and instruments that we must consider as providers of credit and how their credit can be measured. Here again the issue, first and foremost, boiled down to a question of sources. The result is to provide a list of credit providers that have existed in history, allowing *statisticians* to check against the national historical literature whether some of them were active in the period they consider.

### 3.1 The lamppost syndrome: pitfalls in the construction of historical credit aggregates

Constructing a bank credit aggregate requires access to quantitative information on the total amount of credit contracts granted to non-financial agents. Two main sources are easily accessible to historians: public reports published by state authorities on regulated banks and the voluntary disclosure of financial information made by banks listed on the stock market. Those two easily accessible sources are low-hanging fruit that have been used extensively by historians compiling national credit aggregates. For reasons discussed below, the likelihood of these yielding consistent credit aggregates across time and space are low.

Current historical reconstructions of credit statistics rely on the legal definition of a bank rather than on an economic definition. However, the conditions for holding a banking licence have changed over time, and, historically, often no licence was required at all. Licensing is linked to privileges, eg different tax treatment, the possibility of incorporating as a joint-stock bank, having access to deposit insurance etc. But in some periods and places, entry into banking was essentially unsupervised, as banks could be created under general commercial law. This was the case, for instance, in most European countries before the First World War, and in some even until the 1940s. By contrast, in other places such as the United States, government regulation of banks is as old as the republic.

Government regulation and supervision of banks is not a universal phenomenon across history. According to the conventional view, bank regulation is a 20th century innovation that appeared in the 1930s and 1940s (Goodhart (2010)). Two causes have been advanced for this push: first, the rise of banking instability during the Great Depression and second, the needs of fiscal authorities who wanted access to bank deposits to finance the holding of increasing public debt levels during the First World War and in order to manage it after the war. Following this view, regulation has gradually broadened the scope of banks within its purview. Bank supervision is interpreted as resulting from the fact that banks generate a peculiar form of financial instability by transforming short-term *liquid* liabilities into long-term *illiquid* assets (Fama (1980), Goodhart (1987)). In other words, banks are special not because they

are providers of credit per se but because the credit they provide is associated with maturity transformation.

An alternative approach is provided by Toniolo and White (2015). They suggest that modern forms of bank supervision are sui generis. Bank supervision that began spreading in the 1930s in an age of expanding capital controls was a form of financial supervision under financial repression. However, the previous regime was characterised by the supervision of a subset of banks only – specifically, of banks endowed with the right to issue banknotes, known as “banks of issue”, which were the predecessors of central banks. Van Dillen (1964) suggests that bank supervision in fact started with the creation of this new type of bank in Europe around the 16th century. The rationale for this discretionary oversight, was that, as underlined by the Italian statesman Cavour, these peculiar banks posed a “disruptive threat of crises that made “useless the most stringent [legal] precautions”” (Toniolo and White (2015)).

Accordingly, regulation did not originate historically in the monitoring of the asset side of banks’ balance sheet (as under modern prudential arrangements) but in the peculiarities of the liability side. The invention of banknotes had itself built on existing credit and payment instruments known as bills of exchange (against which the bank of issue issued its notes) but with one (major) difference. While bills of exchange could be used in payment, they did not eliminate the residual liability of the issuer and endorser, creating a system of checks and balances. On the contrary, banknotes were anonymous and memory-free claims on the issuer (Kocherlakota (1998), Segura (2022)). They were accepted in payment by other banks and by non-financial agents, and typically also by government treasuries. Because anonymity increased the risk of fraud, banknote issuance ends up being supervised by public authorities. As a minimum, entry was subject to the securing of a public licence, while continued operation involved abiding by prudential standards.

The rationale, in the case of the United States, for instance, where the comptroller of the currency became responsible for “national banks” in 1863, was the potential moral hazard risk created by the opportunity to issue banknotes in an irresponsible manner. Under this regime, the ostensible concern was to contain the behaviour of so-called wildcat bankers who would issue banknotes to finance doubtful credit claims (Greenfield and Rockoff (1995)). Likewise, banks of issue benefiting from a monopoly (and which were in fact proto-central banks) operated subject to set proportions between the volume of banknotes which they circulated and their capital and/or reserves.

In the end, depending on the regime of banknote issuance – whether decentralised issuance by several banks or a monopoly regime granted to a central bank – different regimes of supervision and data provision emerged. In countries where free banking (including the United States, Chile, Colombia, Mexico and Sweden) or multiple banks of issue predominated, balance sheet items were automatically produced because banks were subject to disclosure standards. In those countries, the archives of the regulators therefore contain the balance sheets of regulated banks. This includes most American countries and some European ones, see Friedman and Schwartz (1970) for the United States; Briones Rojas (2004) for Chile; Hernandez-Gamarra (2001), Acuña-Mantilla and Alvarez (2014) for Colombia; Maurer (2002) for Mexico; Ögren (2006) for Sweden. In countries where one central bank enjoyed a monopoly of issue, the rest of the banking system went “unregulated” and hence no supervisory information is available for them. A “mixed” situation is



provided by Meiji-era Japan since it combined government supervision of the banking system *and* a central bank.

Against this background, the question is to determine the extent to which supervised and thus data-generating banks were representative of broader trends in the economy. The representativeness of the credit series constructed using banks of issue hinges on the breakdown between banknotes and deposits. When agents prefer banknotes (such as in France), then banks of issue secured a competitive advantage, making the data pertaining to their balance sheet more “representative.” This is in contrast to countries where agents preferred holding coins or deposits. The United Kingdom offers a case in point. Peel’s Act of 1844 restricted the issue of banknotes by the Bank of England and placed severe limits on its ability to extend credit, making the Bank of England’s circulation poorly representative of British credit at large.<sup>33</sup> To complicate matters, within each country, the preference for banknotes changed over time, thus implying changes in the market share of note-issuing banks (Germany or Spain during the 19th century are obvious examples, see Burhop (2011) and Martín-Aceña et al (2014)).

In other words, in looking for credit data in those places in which regulators operated, we are at risk of being misled. From a comparative point of view, the general rule is that different regulatory regimes produce very different outlooks for how credit aggregates ought to be computed, because they were really focusing on (and created) different things. Apparent disparities may not, therefore, reflect structural differences across countries. Instead, they are an artefact of the viewpoint adopted and of the manner in which agents responded in kind by adjusting their portfolios. This suggests that it is necessary to transcend such difficulties if we are to produce relevant credit data.

To complicate matters further, the regulation of banks responds to the market response. The result is a never-ending game of markets adjusting to regulation through innovation and authorities catching up with banking innovation (the current controversy on new digital payment solutions bears witness to this). Authorities have shifted supervisory work, sometimes centralising, sometimes decentralising, sometimes unifying and sometimes segmenting. The canonical case is in the United States (White (1983, 2013)), but it is hardly unique. Spain, to name but one other country, is also a good example (Martín-Aceña et al (2014)). As they engaged with this supervisory work, regulatory authorities – whether they are central banks, the treasury or agencies reporting to the government or parliament – not only defined the scope of banking but they also defined the “measurable” part of banking, something that we will explore further in Section 3.3.

In countries that historically did not closely regulate entry into banking such as the United Kingdom, Germany or France, a key challenge for the post-Second World War *statistorian* is to identify proper sources for balance sheet data. One source of information arises in the shape of balance sheets published as a result of statutory rules on, say, joint-stock incorporations. This led to audited annual reports that were reproduced in newspapers and are sometimes found in archival repositories. Such banks often shared the characteristic of being listed on the stock market, and the publication was made either because publication was mandatory or – in cases where disclosure was voluntary – because investors rewarded transparency. The

<sup>33</sup> See Capie and Webber (1985). Under Peel’s Act, the Bank could issue £14 million without backing after which every note issued had to be backed one for one by gold holdings.

documentation tends to be more extensively available for the principal institutions, both because their size warranted disclosure, or because their significance ensured coverage in the media. In other words, the availability of balance sheets (and thus, credit data) is biased in favour of big joint-stock banks – leaving out many banks, including those privately incorporated (thus eschewing joint-stock requirements) and those not large enough to be eligible for public listings. As shown in the next section, the portion of the credit market that goes undocumented can be enormous and its size can be variable over time.

The main takeaway from this discussion is that the different banking systems rested on different disclosure principles, and as a result, for the long historical period before the era of universal supervision, historians find themselves very much in the position of the drunken person looking for their key under a lamppost. It follows that available estimation of credit aggregates have to be checked against the fact that all statistical (re)construction is historically situated. Accordingly, the quality and scope of existing credit aggregates can only be gauged by taking into account the context. And because the lampposts are situated in different places depending on the country under study, comparability is never warranted and should never be presumed.

### 3.2 Intensive margin and representative samples: do we measure bank credit adequately?

Today, constructing credit aggregates requires using either information on the amount of credit provided by the aggregate population of credit providers or, when this is not implementable, using a representative sample of that population. As discussed in the previous section, both regulation and voluntary public disclosure produce samples that are not necessarily representative of either the broader population of banks or of their credit granting activity. Regulation (and the absence thereof) implies different reporting requirements and practices. Voluntary public disclosure of balance sheets is influenced by the bank's business model (as stated above, for instance, if it is listed on the market). It may well be the case therefore that in many countries, regulated banks and banks listed on the stock market were not representative in any meaningful sense.

To get a sense of the problem we first explore the "dark matter". We discuss which types of financial institution may have been left out of regulatory activity, or may not have had incentives for voluntary public disclosure of their bank balance sheets. On the basis of a few cases studies, we then turn to sketching out the importance of what is left unknown, in order to illustrate the extent of the problem. Finally, we provide indications about how the quality of existing credit aggregates can be tested and, if need be, corrected.

We identify four types of credit lender that have typically been left out of the spotlight.

- First, the typical unknown is the (small or large) unsupervised private banks, which were, for example, organised as partnerships and were not subject to any form of regulation or oversight. Information on those banks was in some instances so rare that when, during the investigation of Wall Street banking in 1933, a rough summary of JP Morgan's titanic balance sheet was extracted from Jack Morgan – by Ferdinand Pecora, the Chief Counsel to the United States Senate Committee on Banking and Currency – this was hailed as a triumph (Chernow (1990)). Yet such banks were both prominent in the main financial

centres and in some cases they were important financiers of the local economy. Even if there are economies of scope in banking that lead to the creation of banking conglomerates with many branches spread across a given national territory, small banks may be better (at some points in history) at screening and monitoring borrowers (Nishimura (1995), Stein (2002)). What is more, private banks sometimes do quite different things from listed joint-stock banks. For instance, as was the case in the age of the “commercial revolution” and remained the case to some extent in the interwar period in Britain, private banks specialised in specific activities such as the financing of world trade. In other words, the omission of certain institutions is not innocuous due to underlying structural features. In a deep way the problem is reminiscent of the one discussed in Chapter 2 on interest rates: one would really need a theoretical model to interpolate the missing parts of the credit nexus.

- A second type of bank that has received limited attention in most previous studies is credit unions, credit cooperatives and more generally community lenders. Credit unions often emerged because of the existence of strict bank regulation that made it difficult for banks to expand in some places or because of the difficulty experienced in enforcing repayment for certain types of creditor within the realm of the legal system. Akin to the current “microcredit” associations that are pervasive in emerging market economies, credit unions started by being local in order to exploit social and peer pressure to improve on legal enforcement techniques. While research suggests that in the United States for instance, financial repression stood in the way of the development of banking in African-American communities (Badaran (2019)), community lenders thrived in other places, and their role was especially important, for instance, in 19th century rural Germany (Suesse and Wolf (2020)). More broadly, in pre-Second World War Europe, the growth of farmers’ credit cooperatives was bolstered by unionisation of the local associations, and the phenomenon became sizable.
- A third omission concerns savings banks. Here the question at hand is the function they fulfilled. In several countries such as Spain, savings banks were allowed to lend to the private sector (Martin-Aceña (2014)). Lehmann-Hasemayer and Wahl (2021) have recently documented the importance of savings banks for industrialisation in Germany. In other countries, however, the inclusion of savings banks in credit aggregates may be a case of double-accounting. In France for instance, savings banks were instruments to channel funds to a government-owned bank, which then lent to the government or the private sector (Priouret (1966)). In this case the addition of savings banks makes no sense.
- Finally, there were institutions that operated below any regulatory or statistical radar because of the manner in which they were incorporated, for instance, as trusts. To take but one example, in Canada, responsibility for banking supervision was assigned to the federal government but provincially incorporated mortgage and loan companies soon started accepting deposits, turning themselves into “near-banks” (Kyer (2017)). While the history of these institutions goes far back into the 19th century, the national statistical office of Canada only started to publish (fragmentary) data in the 1960s, at the moment when large parts of the sector voluntarily accepted oversight in exchange for becoming part of federal deposit insurance.<sup>34</sup>

<sup>34</sup> See country chapter on Canada in the accompanying BIS Paper.

Turning now to specific countries, it is possible to get a sense of the omissions. One striking aspect is their variable character from one country to the next, themselves a result of political, regulatory and historical idiosyncrasies.

Germany provides a good illustration of the bias induced by focusing solely on banks under the lamppost. For the German Reich (post-1871), no official banking statistics were published. Publication began under the Weimar Republic in 1923. Before the 1930s, only mortgage banks and note-issuing bank(s) fell within the purview of regulation. Savings banks, which were subject to specific regulation (Guinnane (2002)), started to be included in the data of the Reich's statistics office in 1900. As a result, the lending institutions left outside included regional government banks, joint-stock banks and all privately incorporated banks. Burhop (2011) and Krieghoff (2013) estimated that in 1900, only a third of the German banking system was under federal regulation, while a public authority owned another third (without being supervised by the federal authority) and the last third was both private and unregulated. Guinnane (2002) estimates that, in 1913, there were 1,221 private banks (to be compared with the 160 incorporated banks). For this last type of bank, the aggregate balance sheets published in Deutsche Bundesbank (1976) draw mostly on contemporary compilations undertaken by newspapers (notably *Der Deutsche Oekonomist*) and various economic federations. The *Oekonomist* relied for its series on the published balance sheets of the larger banks, while information on all joint-stock banks is only available from a parliamentary enquiry undertaken in 1908–09.

In Austria, the creation of joint-stock banks was regulated by a general incorporation law in 1862, while banks organised as single proprietorships or partnerships operated under standard commercial law and escaped scrutiny. It is only after 1924 that the law required private banks to apply for a licence (Resch (2018)). Still, there were no reporting requirements. Special legislation on banking was only introduced with the annexation of Austria in 1938. Austria kept the German banking law after 1945, while the current legal framework was enacted in 1979 (Jobst and Kernbauer (2016)). As a result, it was mainly commercial information providers that compiled and published bank balance sheets. Based on the annual reports that joint-stock banks were required by law to send to the statistical office, a regular biannual publication was started in 1882. However, data were not harmonised, differed in the level of disaggregation and sometimes were just lacking.<sup>35</sup> Similar compilations existed for savings banks, while data on credit cooperatives were rarely published, as in Germany, due to the large number of cooperatives and their small individual size. Banks organised as partnerships were covered neither in private handbooks nor in official statistics. The first survey of the entire banking sector including private banks dates from a parliamentary enquiry in 1921 (Resch (2018)).

In the United Kingdom, the role of private banks (from country banks to discount houses) is shrouded from view (Pressnell (1956)). They were initially unregulated and did not automatically publish balance sheets. According to Sheppard (1971), published statistics exclude the UK business of overseas banks, the assets of investment trusts, unit trusts, private trust and superannuation funds, and the accepting houses as well as most of the merchant banks. Yet he estimated that those institutions accounted for 50% of the evaluation of the included institutions' assets.

<sup>35</sup> See the introduction to KK Statistische Zentralkommission (1885).

Including those institutions would yield its own issues such as the separation of claims on the financial sector versus the non-financial sector.<sup>36</sup>

In France, the balance sheets of banks published as part of their state supervision are available only for the Bank of France as a note-issuing bank, for the state-owned Caisse des dépôts et consignation and for the savings banks (which did not lend money to non-financial agents). All other banks were unregulated and were not subject to a special chartering procedure between 1867 and 1941. A survey of all banks operating in 1938 is available. However, a handful of big joint-stock national deposit banks started publishing balance sheets in financial newspapers in the 1860s. Except for those banks, the paucity of information is quite complete for the remaining couple of thousand banks which were active during the 19th century, except for a few dozen regional banks starting in the 1900s for which information is available in contemporary books such as Kaufman (1914).

In Italy, until the Banking Law of 1926, regulatory and disclosure requirements differed across bank categories, being more stringent for savings banks and mortgage banks than for joint-stock banks and cooperative banks, which were subject to the general commercial law (Natoli et al (2016)). However, all were required to file balance sheets either to the District Court or the Ministry of Agriculture, Industry and Commerce, following a prearranged schedule (Cotula et al (1996)). The Law of 1926 also granted supervisory powers to the Bank of Italy, including the possibility of running inspections. This enabled the provision of standardised schedules which banks had to file on a regular basis with the central bank (Molteni and Pellegrino (2021)). The new regulation produced a marked improvement of banking statistics, which were further refined after the banking laws of 1936–38.

In Spain, entry into the banking sector and the activities of banks were unregulated until the 1920s except for note-issuing banks, and the disclosure requirements were lax. This complicates the ex post construction of bank credit aggregates. But some surveys are available. Savings banks were unregulated, state-sponsored banks owned and run by private shareholders were established (and supervised) in the 1920s to direct credit to some specific sectors (Martín-Aceña (2012)).

This list could be continued with many other similar cases, for example in Canada, Mexico or Switzerland etc. They would invariably show that the frontiers of regulation or voluntary public disclosure were not only changing over time but also varied importantly across countries. The complexity of this issue is increased by the fact that in all of those cases, qualitative historical sources suggest that the market share of each type of bank evolved over time, sometimes very significantly, see for example Kaufman (1914) for France, or Burhop (2011) for Germany. Though early *statistorians* have been aware of this issue, most of the time reconstructions of the total of credit assumed a proportion between the credit market share of included banks (those with a publicly disclosed balance sheet) and the credit market share of omitted banks. A not small corollary of this is that, because all the evidence points to an increase in coverage over time, existing studies under-estimate aggregate bank credit for earlier years, producing a distorted image of long-run trends in aggregate leveraging.

<sup>36</sup> For recent explorations of the British dark matter and the promise held out by archival research see Brunt (2006), on country banks; Temin and Voth (2006) for the material contained in the papers of Hoare's Bank, a London bank; Bignon et al (2012) and Accominotti (2012) for the archive of the Bank of England as a source for studying merchant banks.

## Sources

The country examples discussed so far have already pointed to a number of sources that can be used for periods when official/administrative statistics based on regulatory information are not (yet) available. A standard source to build credit aggregates are (private) compilations of published balance sheet data. These compilations have typically been undertaken by leading financial newspapers like *The Economist* in Britain, the *Deutsche Oekonomist* in Germany, and in France, *l'Économiste français* and *Le Rentier*, or by financial yearbooks like the Austrian *Compass*, the French *Annuaire Desfossés*, or the Italian *Bollettino ufficiale delle società per azioni*, which have already been used in research. As explained above, these sources typically cover a subset only of all banks.

Against this backdrop, a useful element pertains to institutional demographics. If detailed information is available for subsamples, simple counts of banking firms can be used to extrapolate aggregates. A potent source for demographics consists in commercial yearbooks, trade lists and address books, and sometimes the archives of rating agencies. These are available for many countries and cities, and feature lists of firms classified by profession. They tend to be much more comprehensive than the compilations cited above when it comes to identifying individual entities. An example for France is the "Didot-Bottin", a prominent commercial address directory, published from the 1830s onwards.<sup>37</sup> While typically not giving information on the balance sheet, registers may contain additional information that can help assess the size of the banking house, eg by giving the numbers and names of partners, business correspondents, subsidiaries or branch offices, or, in the case of Germany, the number of Prokuristen (officers authorised to sign on behalf of the firm).<sup>38</sup> Note that in most cases, address books were published by private editors and contained information deemed important for users, but again, the information is not necessarily complete. Omissions should, however, be limited, as long as editors strove to cover the most important firms and, on the other hand, firms looking for clients had an interest in being included themselves. An alternative but similar source is provided by legal registers of firms. In many countries, companies have to register in a public ledger to exist. Such lists have the advantage of being public and complete, but some care has to be taken as those companies were not necessarily active in business and may only have been empty shells.

Another source for reconstructing demographics is the discount window of central banks. The Bank of England's Discount Office kept rating books, discounter ledgers and acceptor ledgers; and eventually required acceptors to submit annual reports to the Bank (Flandreau and Ugolini (2013, 2014), Accominotti (2012)). However, the two-tiered nature of the British banking system ensured that institutions having access to the discount window of the Bank of England were, by construction, not representative. Other central banks such as the Bank of France or the Austro-Hungarian Bank, by contrast, kept more indiscriminate links with the banking system, making their archives highly informative (Avaro and Bignon (2019), Jobst and Rieder (2023)). Box III.1 illustrates how information on the French banking system in the 19th and 20th centuries can be teased out from the archive of the Bank of France. Alternative sources for building up an image of demographics are the archives of

<sup>37</sup> See eg Firmin Didot (1875).

<sup>38</sup> For a use of this proxy, see Resch (2018).

private banks, which collected intelligence on the economy at large. Here, a famous case is the *Crédit Lyonnais* also mentioned in Box III.1 (Flandreau (2003a,b)).

Credit unions, but also building societies and savings banks, were often organised in associations. For the individual institutions, which were often small in size, associations provided various services including legal counsel, model statutes, training and lobbying. In turn, associations collected information on their members, including, sometimes, their balance sheets. In other cases, individual institutions relied on the services of a head institute, which, among others, could reallocate funds between member banks and provide access to money and capital markets. For Germany and Austria, the (fragmentary) information available on credit cooperatives comes to a large degree from such associations (Deutsche Bundesbank (1976), KK Statistische Zentralkommission (1901)).

The last source we mention is government-commissioned surveys. Historically, surveys were often undertaken at the behest of parliament or of the executive in the aftermath of financial crises, in response to perceived tensions in the banking sector or when political criticism was levelled against the banking system. Examples include the German *Banken-Enquêtes* of 1908–09 and 1933, following the crises of 1907 and 1931; the Austrian *Bankkommission* of 1921, after the break-up of the Habsburg monarchy; and the French survey of 1938 (Laufenburger (1940)). While surveys may thus reflect quite specific historical situations, they can, depending on their design, have the significant advantage of covering the full universe of banking institutions. Such investigations also provide opportunities for gathering evidence not otherwise available in standard sources. While, by definition, surveys are only snapshots, they are a powerful yardstick by which to benchmark series.

### Representative samples

For some research questions, it might not seem to be necessary to calculate the balance sheet or credit aggregate for the entire population, as growth rates for a subsample of banks is sufficient. The argument is only valid if selection into the group of reporting banks was random. There are, however, good reasons to believe that this was not the case. Banks behave strategically, optimising the level of disclosure by moving from private partnership to public listing and back again, or by seeking refinancing. Emphatically, the population of banks showing up at the discount window or taking advantage of the various support mechanisms created – for instance, in the aftermath of the crisis of 2008 – is *not* random, owing to the fear of stigma. In the same manner, cosy regulators decide to ask for regulatory information from some, but not all, banking institutions.

One complication is the transformation of the representative sample as banks hop in and out of a given category, leading to delicate splicing issues. In Germany and Austria, most modern universal banks were founded by private bankers in the second half of the 19th century. The geographic expansion of these banks around 1900 was often effected through the acquisition of private banks and their conversion into branch offices (Michel (1976), Riesser (1911)). Likewise, many of the 50 joint-stock banks founded in Spain between 1900 and 1914 were transformed old commercial houses or resulted from the incorporation of leading private banking firms (Martín-Aceña (2012)). The same movement has been observed in France during the period (Kaufman (1914)). The key takeaway is that the observed growth of joint-stock banking is thus, to an (unknown) part, due to a shift between corporate forms within the banking sector, not aggregate growth per se.

## Dark matter in banking: a case study

France provides a case study to illustrate the census problem associated with constructing credit aggregates using publicly available information on banks. As this box shows, this approach leaves in the shadows many banks that did not publish a balance sheet, or whose published balance sheets have been lost.

In the historical statistics literature, three competing series exist covering the pre-1941 period:

- Saint Marc (1983) relied on a group of four prominent banks in operation in 1938. They were so prominent at this date that they represented 50% of the total credit, at a time (1938) when comprehensive banking and credit statistics were first produced for the economy at large. Before 1938, comprehensive data did not exist but the annual balance sheets of the four banks were available. To adjust for the excluded banks, she multiplied by two the series for total credit of the four large banks going back to 1870 (since the omitted granted as much credit as the measured ones in 1938). She does not really explain what she did for the earlier part of the series when only three out of the four banks were in operation, but we may imagine that a similar ratio was constructed.
- Patat and Luftala (1986) used essentially the same method (benchmarking with 1938), but to adjust for the banks left out, they computed a correction coefficient from what they describe as a “census” of all the banks that had published a balance sheet at various points in time. Their coefficient is the share of credit granted by the four prominent banks as a percentage of the credit granted by all the banks present in the “census”. The result is a series that starts in 1910 and ends in 1938, with correction coefficients computed for 1913, 1921, 1926, 1931, 1935, 1936, and 1937. They finally interpolate the series between these benchmark years. Patat and Luftala do not detail which banks they considered.
- Before 1941, the research department of the largest bank of the time, Crédit Lyonnais (one of the banks used in the benchmarking exercised described above) made a huge effort in collecting data on its competitors (Flandreau 2003a,b). The card index it produced has been lost but some of the aggregates that it constructed at the time have survived in its archive. Baubeau et al (2021) rely on this source to produce a series for total credit. The resulting bank credit series, centred on the interwar period, assumes that the banks not counted by Crédit Lyonnais were too small to influence the level of the aggregate credit.

All three studies make assumptions about the possibility of ignoring the banks that did not disclose their balance sheet (and thus are not observable). Owing to the auditing reports of the Bank of France (which examined all banks and occasionally documented their balance sheets) it is possible to show that, in 1898 for instance, when there were approximately 2,000 banks in operation in the country, the lack of a published balance sheet did not necessarily mean that one was dealing with a small bank (See Avaro and Bignon (2019) for a discussion). Using contemporaneous commercial almanacs (which documented the existence of the banks but not their data) and the archive of the Bank of France (which documented random balance sheets), one can construct an estimate of the “dark matter”. For instance, on the assumption that the size of the balance sheet of a branch of the three largest banks was similar to the size of the balance sheet of a bank operating only one branch (an assumption warranted by the balance sheets available in 1898), one finds that, at that date the three largest banks represented about 18% of the credit market. Using this number to correct the credit series leads to an estimate for aggregate credit in 1898 which is three times larger than Saint Marc’s estimate, the only currently available estimate for this date.

Bank credit to the private sector as a percentage of GDP in France, various dates Table III.1

	1898	1920	1929	1938
<b>Saint Marc (1983)</b>	11.0%	14.8%	20%	15.7%
<b>Patat and Luftala (1986)</b>	NA	13.6%	27.6%	15.7%
<b>Baubeau et al (2018)</b>	NA	18.5%	23.1%	15.7%
<b>Alternative estimate in 1898</b>	31.3%	NA	NA	15.7%

Table 1 converts these numbers into credit-to-GDP ratios. The estimates suggest very different conclusions for the secular evolution of the credit-to-GDP ratio.

To illustrate the extent of the problem to be dealt with, Box III.1 discusses the case of France. Available French credit series for French “credit” before 1941 go back to the work of Saint Marc (1983), revised by Patat and Luftala (1986), and these



numbers have been used in recent macro-historical research. The box demonstrates that these credit aggregates do require significant revisions, leading to a reassessment of secular patterns. In the case the box considers, this arises because the largest banks, operating through a national network of branches, experienced an increase of their share in the credit market, from 20% in 1900 to 50% in 1938, at the expense of smaller banks which had been omitted from earlier studies. In other words, what may have looked like an increase in credit-to-GDP ratio, was in fact a substitution of credit from small banks to credit from large banks. As shown in Table III.1 in Box III.1, the potential correction is large.

As the box also demonstrates, the demographic information provided by commercial almanacs and the archives of central banks (here the archive of the Bank of France) are in fact an appropriate source for exploring such questions and may help construct significantly more accurate aggregates. There is no reason to believe that France is an outlier and so important revisions in country-wide historical bank credit aggregates are in order.

### 3.3 Institutions that are not “banks” but act like banks

The previous two sections have discussed the pitfalls of measuring credit when information on *banking* is fragmentary. But bank credit is not the only relevant credit aggregate, as emphasised by the pioneering study of Goldsmith (1958) on US financing of non-financial agents between 1900 and 1951.

Credit can be intermediated by banks, sometimes by other intermediaries or it can be directly granted through organised or decentralised financial markets. Financial markets have changed over time, and as a result, some countries went from being bank-based to being market-based and vice versa (Rajan and Zingales (2003)). That back and forth over time or across countries may be explained by the relative cost of arm’s length versus relationship-based finance and by the difficulty for each credit provider of relying on the legal system to enforce the reimbursement of credit claims (Allen and Gale (2001)). It may also be explained by the political economy of finance (Rajan and Zingales (2003)) as the frontier between banks and non-banks is a political one (Calomiris and Haber (2014)).

In this section we argue that proper credit series ought to account for this. After the subprime crisis, the term “shadow banking” emerged. It suggested that there was something “shady” and thus illegitimate about parallel forms of banking that had developed before the crisis and were perceived as having contributed to it. However, in theoretical approaches, banks are agents that issue liquid liabilities – used as means of payment – in order to finance the activity of other agents (Tobin (1963), Gorton and Pennachi (1990)).<sup>39</sup> This is a reminder that shadow banking is just banking. Disregarding non-bank credit in the construction of credit aggregates is therefore problematic since it overlooks sources of fluctuations that have the potential to shed light on financial crises. Historically, examples of non-bank banking are legion. We

<sup>39</sup> Some have defined banks by focusing on their capacity to screen and monitor the credit risk of the borrowers (Diamond (1984)). Others have noticed that the very temporal nature of credit creates enforcement risks, something that differentiates credit and money (Hawtrey (1919), Kocherlakota (1998)). Specific institutions, such as registration etc, ensure that banking and money stand apart from one another. A common understanding of the risk associated with banks is that they are exposed to the risk of bank runs created by the fact that banks offer liquid deposits to investors to fund less liquid investments (Goodhart (1987), Diamond and Dybvig (1983)).

give a list below. They may have played variable roles over time and countries – and not necessarily a destabilising one – but they have been sufficiently important to attract the attention of previous researchers who have typically connected them to crises. We identify five significant cases, which sometimes overlap with one another.

A first significant case consists in non-bank investment vehicles, in the shape of money market funds or investment trusts, investing in debt whether securitised or not. These constitute genuine anticipations of modern shadow banking. Not incidentally, these creatures have typically been discussed previously against the backdrop of financial crises. A well known example is the growth in Britain of so-called “finance companies” during the 1850s and 1860s. These were joint-stock companies which raised funds in the money market and originated short term credit (typically, trade credit), which they then kept on their balance sheets. Some ended up locking up their funds in operations whose maturity did not match their short-term funding structure, one famous case being Overend Gurney, which lent its name to the panic of 1866 (Xenos (1869), Flandreau and Ugolini (2014)). Another example is provided by the creation of “investment trusts” (also known as “financial trusts”) sponsored by investment banks and which served to park the securities which these banks originated. The creation of investment trusts in the United Kingdom in the late 1880s is conventionally associated with the Baring crisis (Anonymous (1892)).

Likewise, from the beginning of the 20th century, US investment banks were routinely criticised for originating loans and then parking them in investment trusts and insurance companies which they controlled. Investment trusts featured in the crisis of 1907 in the US and led to the investigation of the “money trust” (Brandeis (1914), Moen and Tallman (1992)). They also featured in other episodes of financial stress. The tight regulation that followed the Great Depression resulted in a clearer delineation of lines, but after the deregulation of the 1980s and 1990s, the phenomenon became pervasive again leading to the generalisation of the “originate and distribute” model in the 2000s.

A second significant case in which non-banks played a critical role is provided by the growth of the short-term credit market instrument known as bills of exchange. Inherited from commercial fairs of the European middle ages (De Roover (1953), Van der Wee (1978)), a bill was a transferable form of cheque. With this instrument, merchants extended credit to one another, and this credit was then securitised by the intervention of reputable endorsers, and circulated. The non-banking origins of this form of credit are illustrated by the terminology used to describe the originators of prime bills. In London, they were known as “*merchant* bankers”, who rarely incorporated until the 20th century (Chapman (1984)). This bears witness to the fact that leading international merchants had a comparative advantage in the screening of the credit risk associated with a specific commercial borrower, and as a result “grew to become bankers” rather than being bankers to begin with.<sup>40</sup> This form of credit retained critical importance until the 20th century. Its informational content as a macro-financial indicator is attested by attempts by early students of the business cycle to relate business cycles, crises and fluctuations in the supply of bills to one another (Juglar (1856)). The need to manage the circulation of bills became one rationale for the creation of central banks (Bignon and Flandreau (2018)) and drove the consolidation of “modern” central bank policymaking (Wood (1939)).

<sup>40</sup> Gorton (2020) describes the historical process of origination and distribution of bills in Britain as private money production without banks.

A third case is provided by companies that finance their customer purchases with debt contracts. Credit is then created outside the banking system. Yet this form of credit does require the same kind of screening and monitoring that is required in banking, and in practice, the agents who extend such credit act as banks. The early history of the US economy, which Flandreau and Geisler Mesevage (2014) describe as an instance of separation of information and lending, offers extensive examples of this situation. In the 19th century, the business of “dry goods” rested exclusively on non-bank credit. New York wholesalers extended loans to their distributors scattered across the territory. The need to screen and monitor distributors provided the foundation for a separation of monitoring and lending. They bought monitoring services from so-called mercantile agencies – the forerunners of modern rating agencies – which undertook to sell credit updates. Later, the mercantile agencies found themselves capable of providing high-quality credit data and generating important business cycle indicators.

A fourth case is of peer-to-peer credit markets. Here, the most famous historical example is that of the role of notaries as intermediaries in mortgage credit (although their role was not limited to real estate). In many countries, including France, Mexico, Portugal, Spain, Switzerland and the Netherlands, notaries fulfilled an important intermediation function until well into the 19th century by matching borrowers and lenders (Hoffman et al (2019), Levy (2012), Costa et al (2017), Peña-Mir (2020), Roth-Lochner (1997), Gelderblom et al (2018)). In these countries, mortgages were not usually dealt with by banks but by notaries who acted as brokers between lenders and borrowers.

Last but not least, a fifth case concerns bonds at large. We have already mentioned the issue of bonds intermediated by banks but some issuers have been able to tap the market directly. A case can be made for the inclusion of bonds at large as part of overall credit aggregates, whether they are held by banks, shadow banks, pension funds or insurance companies, or by the public. Indeed, they often substitute for bank credit. At the very least, studies of the secular evolution of the respective role of banking and alternative forms of credit should be commissioned.<sup>41</sup> For instance, in France, when the role of notaries started to fade away in the late 19th century, they were replaced by special purpose finance companies that financed themselves by issuing shares and debentures in the market (Lescure (1980)). In Germany and Austria, the financing of residential and commercial mortgages relied at an early stage on the issuance of Pfandbriefe (they were discussed in Chapter 2 on interest rates).

In summary, the discussion shows that much remains to be done if we are to provide reliable credit data, though many promising avenues exist.<sup>42</sup> There are also a number of promising sources to explore. Examples include the following.

First, to the extent that they were incorporated and as a result, in some countries, subjected to disclosure standards, historical “shadow banks” – such as the British credit or finance companies of the 19th century and the investment trusts of the late

<sup>41</sup> Series of bonds financings have been constructed for a number of countries since the study on the financing of US corporations by Hickman (1953).

<sup>42</sup> A significant question is that of the most adequate way to tackle the resulting data collection effort, as different perspectives may be adopted. For instance, one may attempt to collect data on the balance-sheet of historical shadow banks (such as the investment trusts) or instead aggregate data on the kind of investments in which these shadow banks participated. Different strategies will meet with different challenges.

19th and 20th centuries – have left balance sheets. This could provide a basis for reconstructing their demographics and activity.

Second, bills of exchange were often taxed by a stamp duty, sometimes under the penalty that otherwise contract enforcement would be limited. Such sources have been used by a number of authors in the past in order to reconstruct the volume of bills originated during a year (Roulleau (1914), Nishimura (1971)).<sup>43</sup> The approach could be extended and generalised.

Third, the archives of notaries are another type of source with which to reconstruct more complete credit aggregates, especially with respect to real estate, although we would also flag the role of joint-stock companies in this sector.

Fourth and finally, the revenues of bond and equity shares were often taxed, thus providing a source to reconstruct the total volume of securities floated in the public market. This type of source has already been used to reconstruct capital flows during the period from 1870 to 1914 (Esteves (2007, 2011)). Combining this information with more adequate bank credit data offers promising avenues to identify secular trends in credit.

### 3.4 Conclusion

The study of broad credit aggregates can illuminate the causes of economic crises. Existing historical credit data, however, suffer from a number of flaws which we have reviewed here. First, credit data rest on an exercise in bank demographics. The quality of previous achievements may be questionable and is up for a reassessment.

Second, although country regulators have contributed to produce either primary or secondary material of great interest, they have tended to see the world from their own window. Different regulatory traditions actually induce cross-sectional biases by the fact that they focus on different items.

Last, the great lesson of the subprime crisis (that shadow banking matters) leads to the realisation that historically, not one but several forms of shadow banking have always existed. It is not realistic to call for the construction of reliable credit aggregates, if we are not prepared to tackle the preliminary step of reconstructing these critical components of the credit system. The alternative is what we have described as the lamppost syndrome.

The conclusion is that we are still in the shadows (no pun intended) when it comes to the measurement of the secular evolution of credit. Unlike for interest rates, where a relatively clear picture is emerging, unlike for real estate prices for which decisive elements are fast emerging, we are still largely (and surprisingly) in terra incognita. Not so much because the territory has never been explored, but because much remains to be discovered as we grapple with the question “what is credit?” At heart, we suggest that what is most needed at this stage are careful country studies and comparative work on specific banking institutions.

<sup>43</sup> Converting this into outstanding credit amounts raises questions pertaining to average maturities and has to be dealt with in a piecemeal way.

## References

- Accominotti, O (2012): "London merchant banks, the central European panic and the sterling crisis of 1931", *The Journal of Economic History*, vol 72, no 1, pp 1–43.
- Acuña-Mantilla, K and A Álvarez (2014): "De la moneda metálica al billete de banco en Medellín y Bogotá (1871–1885): complementariedad y sustitución de medios de pago en un régimen de banca", *Revista Tiempo & Economía*, vol 1, no 1, pp 77–106.
- Allen, F and D Gale (2001): *Comparing financial systems*, MIT Press.
- Anonymous (1892): "The art and mystery of financial trusts", *The Investors Review*, no 1, pp 25–55.
- Avaro, M and V Bignon (2019): "At your service! Liquidity provision and risk management in 19th century France", *CEPR Discussion Paper*, no 13556.
- Baubeau, P, E Monnet, A Riva and S Ungaro (2021): "Flight-to-safety and the credit crunch: a new history of the banking crisis in France during the Great Depression", *Economic History Review*, vol 74, no 1, pp 223–50.
- Bignon, V, M Flandreau and S Ugolini (2012): "Bagehot for beginners: the making of lending of last resort operations in the mid-19th century", *Economic History Review*, vol 65, no 2, pp 580–608.
- Bignon, V and M Flandreau (2018): "The other way: a narrative history of the Bank of France", in R Edvinsson, T Jacobson and D Waldenström (eds), *Sveriges Riksbank and the history of central banking*, pp 206–41.
- Brandeis, L (1914): *Other people's money, and how the bankers use it*, New York, Frederick A Stokes.
- Briones Rojas, I (2004): *Banque libre: de l'idée à la réalité : le cas du Chili, 1860–1898*, Paris, Institut d'Etudes Politiques, unpublished dissertation.
- Brunt, L (2006): "Rediscovering risk: country banks as venture capital firms in the first industrial revolution", *The Journal of Economic History*, vol 66, no 1, pp 74–102.
- Burhop, C (2011): *Wirtschaftsgeschichte des Kaiserreichs 1871–1918*, Göttingen, Vandenhoeck & Ruprecht.
- Calomiris, C and S Haber (2014): *Fragile by design: the political origins of banking crises and scarce credit*, Princeton University Press.
- Capie, F and A Webber (1985): *A monetary history of the United Kingdom, 1870–1982: data, sources and methods*, vol 1, London, George Allen and Unwin.
- Chapman, S (1984): *The rise of merchant banking*, London, George Allen & Unwin.
- Chernow, R (1990): *The house of Morgan*, New York, Atlantic Monthly Press.
- Costa, L, M Rocha and P Brito (2017): "The alchemy of gold: interest rates, money stock, and credit in eighteenth-century Lisbon", *Economic History Review*, vol 71, no 4, November, pp 1147–72.
- Cotula, F, T Raganelli, V Sannucci, S Alieri and E Cerrito (eds) (1996): *I bilanci delle aziende di credito 1890–1936*, Collana storica della Banca d'Italia, Statistiche, III, Rome-Bari, Laterza.
- de Roover, R (1953): *L'evolution de la lettre de change, XVe–XVIIIe siècles*, Paris, Librairie Armand Colin.

- Deutsche Bundesbank (ed) (1976): *Deutsches Geld- und Bankwesen in Zahlen 1876–1975*, Frankfurt/Main, Fritz Knapp.
- Diamond, D (1984): "Financial intermediation and delegated monitoring", *Review of Economic Studies*, vol 51, no 3, pp 393–414.
- Diamond, D and P Dybvig (1983): "Bank runs, deposit insurance, and liquidity", *Journal of Political Economy*, vol 91, no 3, June, pp 401–19.
- Esteves, R (2007): *Between imperialism and capitalism: European capital exports before 1914*, mimeo.
- (2011): "The belle époque of international finance: French capital exports, 1880–1914", *Oxford Department of Economics Discussion Paper*, no 534, June.
- Fama, E (1980): "Banking in a theory of finance", *Journal of Monetary Economics*, vol 6, no 1, pp 39–57.
- Firmin-Didot (ed) (1875): *Annuaire-almanach du commerce, de l'industrie, de la magistrature et de l'administration, ou almanach des 500,000 addresses de Paris, des départemens & des pays étrangers*, Paris, Firmin Didot.
- Flandreau, M (2003a): "Audits, accounts, and assessments: coping with sovereign risk under the international gold standard, 1871–1913", in M Flandreau, C-L Holtfrerich and H James (eds), *International financial history in the twentieth century: system and anarchy*, Cambridge University Press.
- (2003b): "Le service des études financières sous Henri Germain (1871–1905) : une macro-économie d'acteurs", in B Desjardins (ed), *Le Crédit Lyonnais (1863–1986), Études historiques*, Geneva, Librairie Droz, pp 271–301.
- Flandreau, M and G Geisler Mesevage (2014): "The separation of information and lending and the rise of rating agencies in the USA (1841–1907)", *Scandinavian Economic History Review*, vol 62, no 3, pp 213–42.
- Flandreau, M and S Ugolini (2013): "Where it all began: lending of last resort at the Bank of England monitoring during the Overend-Gurney panic of 1866", in M Bordo and W Roberds (eds), *The origins, history, and future of the Federal Reserve: a return to Jekyll Island*, Cambridge University Press, 25 March, pp 113–61.
- (2014): "The crisis of 1866", in N Dimsdale and A Hotson (eds), *British financial crises since 1825*, Oxford University Press, pp 76–93.
- Friedman, M, and A Schwartz (1970): *Monetary statistics of the United States: estimates, sources, methods*, New York, National Bureau of Economic Research.
- Gelderblom, O, J Jonker and M Hup (2018): "Public functions, private markets: credit registration by aldermen and notaries in the Low Countries, 1500–1800", in M Lorenzini, C Lorandini and D'M Coffman (eds), *Financing in Europe – evolution, coexistence and complementarity of lending practices from the middle ages to modern times*, London, Palgrave Macmillan, pp 163–94.
- Goetzmann, W (2016): *Money changes everything: how finance made civilization possible*, Princeton University Press.
- Goldsmith, R (1958): *Financial intermediaries in the American economy since 1900*, Princeton, NBER Books, January.
- Goodhart, C (1987): "Why do banks need a central bank?", *Oxford Economic Papers*, vol 39, no 1, March, pp 75–89.

- (2010): "The changing role of central banks", *BIS Working Papers*, no 326.
- Gorton, G (2020): "Private money production without banks", *NBER Working Papers*, no 2663, [www.nber.org/papers/w26663](http://www.nber.org/papers/w26663).
- Gorton, G and G Pennachi (1990): "Financial intermediaries and liquidity creation", *Journal of Finance*, vol 45, no 1, pp 49–71.
- Greenfield, R and H Rockoff (1995): "Gresham's law in nineteenth century America", *Journal of Money, Credit, and Banking*, vol 27, no 4, pp 1086–98.
- Guinnane, T (2002): "Delegated monitors, large and small: Germany's banking system, 1800–1914", *Journal of Economic Literature*, vol 40, no 1, March, pp 73–124.
- Hashimzade, N, G Myles and J Black (2017): *A dictionary of economics*, fifth edition, Oxford University Press.
- Hawtrey, R (1919): *Currency and credit*, London, Longmans, Green and Company.
- Hernandez-Gamarra, A (2001): *A monetary history of Colombia*, Bogota, Villegas Editores.
- Hickman, W (1953): *The volume of corporate bond financing since 1900*, Princeton University Press and NBER.
- Hoffman, P, G Postel-Vinay and J-L Rosenthal (2019): *Dark matter credit: the development of peer-to-peer lending and banking in France*, Princeton University Press.
- Jobst, C and H Kernbauer (2016): *The quest for stable money: central banking in Austria, 1816–2016*, Frankfurt, Campus.
- Jobst, C and K Rieder (2023): "Supervision without regulation: discount limits at the Austro-Hungarian Bank, 1909–13", forthcoming in *Economic History Review*.
- Juglar, C (1856 [1889]): *Des Crises Commerciales et de leur retour périodique en France, en Angleterre et aux États-Unis*, Paris, Guillaumin.
- Kaufman, E (1914): *La Banque en France*, Paris, Giard et Brière.
- Kocherlakota, N (1998): "Money is memory", *Journal of Economic Theory*, vol 81, no 2, pp 232–51.
- Krieghoff, N (2013): *Banking regulation in a federal system: lessons from American and German banking history*, London School of Economics, unpublished thesis.
- Kyer, C (2017): *Next best to world class: the people and events that have shaped the Canada Deposit Insurance Corporation, 1967–2017*, Ottawa, Canada Deposit Insurance Corporation (CDIC).
- KK Statistische Zentralkommission (1885): "Statistik der Banken in den im Reichsrathe vertretenen Königreiche und Länder für die Jahre 1882 und 1883", *Österreichische Statistik*, vol 9, no 2.
- (1901): "Statistik der registrierten Erwerbs- und Wirtschaftsgenossenschaften in den im Reichsrathe vertretenen Königreiche und Länder für das Jahr 1898", *Österreichische Statistik*, vol 57, no 1.
- Laufenburger, H (1940): *Les banques françaises*, Paris, Sirey.
- Lehmann-Hasemeyer, S and F Wahl (2021): "The German bank-growth nexus revisited: savings banks and economic growth in Prussia", *Economic History Review*, vol 74, no 1, pp 204–22.

Lescure, M (1980): *Les sociétés immobilières en France au XIXe siècle: contribution à l'histoire de la mise en valeur du sol urbain en économie capitaliste*, Paris, Publications de la Sorbonne.

Levy, J (2012): *The making of a market: credit, henequen and notaries in Yucatan 1850–1900*, Philadelphia, Penn State University Press.

Martín-Aceña, P (2012): "The Spanish banking system from 1900 to 1975", in J Malo de Molina and P Martín-Aceña (eds), *The Spanish financial system*, London, Palgrave Macmillan, pp 99–143.

Martín-Aceña, P, M Pons and C Betrán (2014): "150 years of financial regulation in Spain: what can we learn?" *The Journal of European Economic History*, vol 43, no 1–2, pp 35–81.

Maurer, N (2002): *The power and the money: the Mexican financial system, 1876–1932*, Stanford, Stanford University Press.

Michel, B (1976): *Banques & banquiers en Autriche au debut du 20e siecle*, Paris, Presses de la Fondation Nationale des Sciences Politiques.

Moen, J and E Tallman (1992): "The bank panic of 1907: the role of trust companies", *The Journal of Economic History*, vol 52, no 3, pp 611–30.

Molteni, M and D Pellegrino (2021): "Lessons from the early establishment of banking supervision in Italy (1926–1936)", Bank of Italy, *Economic History Working Papers*, no 48.

Natoli, S, P Piselli, I Triglia and F Vercelli (2016): "Historical archive of credit in Italy", Bank of Italy, *Economic History Working Papers*, no 36.

Nishimura, S (1995): "The French provincial banks, the Banque de France, and bill finance, 1890–1913", *The Economic History Review*, vol 48, no 3, pp 536–54.

——— (1971): *The decline of inland bills of exchange in the London money market 1855–1913*, Cambridge University Press.

Ögren, A (2006): "Free or central banking? Liquidity and financial deepening in Sweden, 1834–1913", *Explorations in Economic History*, vol 43, no 1, pp 64–93.

Patat, J-P and M Lutfalla (1986): *Histoire monétaire de la France au 20ème siècle*, Paris, Economica.

Peña-Mir, J (2020): "How well were property rights protected in early modern Spain? The case of the public mortgage registry in Malaga", *Revista de Historia Económica / Journal of Iberian and Latin American Economic History*, pp 1–31.

Pressnell, L (1956): *Country banking in the industrial revolution*, New York, Oxford University Press.

Priouret, R (1966): *La Caisse des dépôts, cent cinquante ans d'histoire financière*, Paris, PUF.

Rajan, R and L Zingales (2003): "The great reversals: the politics of financial development in the twentieth century", *Journal of Financial Economics*, vol 69, no 1, pp 5–50.

Resch, A (2018): "Österreichische Privatbanken von 1919 bis 1945", in P Eigner, A Resch and H Falschlehner (eds), *Geschichte der österreichischen Privatbanken*, Wiesbaden, Springer, pp 81–211.



- Riesser, J (1911): *The German great banks and their concentration*, National Monetary Commission Doc 593, Washington, Government Printing Office.
- Roth-Lochner, B (1997): *De la banque à l'étude: une histoire institutionnelle, professionnelle et sociale du notariat genevois sous l'Ancien Régime*, Geneva, Droz.
- Roulleau, G (1914): *Le règlement par effets de commerce en France et à l'étranger*, Paris, Dubreuil.
- Saint Marc, M (1983): *Histoire monétaire de la France: 1800–1980*, Paris, Presses Universitaires de France-PUF.
- Segura, S (2022): "Making money amnesiac", forthcoming in *Capitalism, A Journal of History and Economics*.
- Sheppard, D (1971): *The growth and role of UK financial institutions 1880–1962*, London, Methuen.
- Stein, J (2002): "Information production and capital allocation: decentralized versus hierarchical firms", *Journal of Finance*, vol 57, no 5, pp 1891–921.
- Suesse, M and N Wolf (2020): "Rural transformation, inequality and the origins of micro-finance", *Journal of Development Economics*, vol 143, March.
- Temin, P and H-J Voth (2006): "Banking as an emerging technology: Hoare's Bank, 1702–1742", *Financial History Review*, vol 13, no 2, pp 149–78.
- Tobin, J (1963): "Commercial banks as creators of 'money'", in D Carson (ed), *Banking and Monetary Studies*, US Treasury, Richard D Irwin, pp 408–19.
- Toniolo, G and E White (2015): "The evolution of the financial stability mandate: from its origins to the present day", *NBER Working Papers*, no 20844.
- Van der Wee, H (1978): "Monetary, credit and banking system", in E Rich and C Wilson (eds), *Cambridge Economic History of Europe*, vol. 5, pp 290–313.
- Van Dillen, J (1964): *History of the principal public banks*, London, Franck Cass & Co.
- White, E (1983): *The regulation and reform of the American banking system, 1900–1929*, Princeton University Press.
- (2013): ""To establish a more effective supervision of banking: how the birth of the Fed altered bank supervision", in M Bordo and W Roberds (eds), *The origins, history, and future of the Federal Reserve: a return to Jekyll Island*, Studies in Macroeconomic History, Cambridge University Press, pp 7–54.
- Wood, E (1939): *English theories of central banking control 1819–1858 with some account of contemporary procedure*, Harvard University Press.
- Xenos, S (1869): *Depredations; or, Overend, Gurney, & Co, and the Greek and Oriental Steam Navigation Company*, London, S Xenos.

## 4. Historical data on real estate prices<sup>44</sup>

Øyvind Eitrheim, Central Bank of Norway<sup>45</sup>

Clemens Jobst, University of Vienna and Central Bank of the Republic of Austria<sup>46</sup>

### 4.1 Introduction and overview

Real estate accounts for a significant share of private sector wealth and has historically been an important source of collateral for lending. Accordingly, changes in valuation can significantly affect aggregate demand and financial stability, and there is significant interest in understanding the determinants of house price movements. In particular, it is important to understand the reasons why prices deviate from fundamental values and, if so, how fast they might adjust and what economic and financial risks this creates.

Today, property prices are monitored via indices produced according to particular quality criteria, which were developed in the early 2000s by a joint collaboration of central banks and statistical offices, and coordinated by the Bank for International Settlements (BIS) and the International Monetary Fund (IMF) (BIS (2005)). A central principle is that house price indices (HPIs) have to take into account changes in the quality and composition of the housing stock. For most of the countries, quality-adjusted indices started to be produced approximately 20 to 30 years ago. In 1989, the BIS started collecting property price series for research purposes, which it made publicly accessible alongside detailed metadata from 2010 onwards (Scatigna and Szemere (2014)). Similar data sets, although with varying coverage, are maintained at the Organisation for Economic Co-operation and Development (OECD) and the Federal Reserve Bank of Dallas (Mack et al (2011)). With a few exceptions, most of the series in these data sets date back to the 1970s.

But spotting and understanding financial cycles and assessing the returns and risks of housing investments require price data for much longer periods and the demand for historical series has increased accordingly. A first approach has been to assemble and splice existing series from secondary sources. Knoll et al (2017) have managed to construct a panel for 14 advanced economies, covering, depending on the country, the past 100 to 150 years. However, many of these indices, as the excellent documentation accompanying Knoll et al (2017) makes clear, do not live up to today's standards. In fact, more than half of the 57 series used in splicing the 14 country indices are simple mean or median purchase prices without any quality adjustment, while those trying to account for quality mostly rely on some form of mix-adjustment, the least sophisticated method available.<sup>47</sup>

<sup>44</sup> We would like to thank the members of HMFS, and in particular Marc Flandreau, Jan F Qvigstad and Ryland Thomas, for generous comments and input. Warm thanks also go to André Kallåk Anundsen and Lisa Reiakvam.

<sup>45</sup> Øyvind Eitrheim is a director at the General Secretariat, Central Bank of Norway.

<sup>46</sup> Clemens Jobst is Professor of Economic and Social History at the University of Vienna, a CEPR Research Fellow, and serves as academic advisor to the Austrian central bank. All views expressed here are his own and do not necessarily reflect those of the Central Bank of the Republic of Austria or the Eurosystem.

<sup>47</sup> Of the 57 series, eight follow a repeat sale methodology, 12 mix-adjusted, six hedonic, four sale price appraisal ratio, 29 mean/median and three other (some series combine two methods) (Knoll et al

Another approach has been to go back to the archives. The publication in 1997 of the Herengracht index (Eichholtz (1997)), covering 350 years of housing prices for Amsterdam, marked the beginning of a flurry of new historical housing price indices. Some of them were constructed in the context of central bank-led HMFS projects and explicitly aimed at extending existing series further back in time.<sup>48</sup> Other indices were set up by researchers interested in particular historical periods, or, simply, because some promising primary material was available.<sup>49</sup> Most of the new indices cover the late 19th and 20th centuries, and all control for quality through sale price appraisal ratio (SPAR) or hedonic models.

Some historians have taken an even longer view. Land has been sitting there from day three of Genesis, and it arguably provides not only the oldest but historically the most important asset.<sup>50</sup> For significant periods of time, land has been traded in markets, and prices in these markets have reacted to changes in supply and demand. There is evidence related to Babylonia (Jursa (2014)) and markets in ancient Rome were linked to the financial system (Temin (2001)). As Tacitus, Dio and Suetonius recount, a credit crunch in 33 AD, brought about by an interest rate cap imposed by Emperor Tiberius, caused a significant decline in the price of land as wealthy senators scrambled for liquid funds (Goetzmann (2016)). For medieval Italy there are records of land sales as early as the eighth and ninth centuries, probably in continuation of Roman practice, and evidence mounts for 11th-century Milan, Bergamo and Tuscany. In Holland and Flanders, market-intermediated transactions can be documented from the 13th century onwards and for south-east England somewhat later.<sup>51</sup> Pervasive and well functioning factor markets are seen as one of the explanatory elements in the debates on the “great divergence” and the “little divergence” (Van Zanden (2009)). But factor markets were not exclusive to north-western Europe, as Cerman (2008) has argued for medieval central and east-central Europe. Evidence in Faroqhi (1987) for the 16th century Ottoman Empire can be interpreted in a similar vein, as can study of post-16th century China by Pomeranz (2000), and Saito (2009) on Tokugawa Japan, in which markets for land defied restrictive official legislation (Matsubara (2017)).<sup>52</sup>

(2017, pp 141–2)). On the relative merits of different quality adjustment methods see Section 4.2 below.

<sup>48</sup> Eitrheim and Erlandsen (2004, 2005) cover four Norwegian cities from the mid-1800s onwards. Bohlin (2014) covers Gothenburg, and Söderberg et al (2014) Stockholm (both 1875 onwards). These efforts are still ongoing. Austria has made plans to construct quality-adjusted HPIs starting in 1900, and Italy plans to go back to 1904.

<sup>49</sup> See eg Samy (2015) for London (1895–1939); Friggit (2008) for Paris (1840 onwards); Edvinsson et al (2021) for Stockholm (1818–2018); Raff et al (2013) for Beijing (1644–1840); Deeter et al (2017) for Dublin (1708–1949); Keely and Lyons (2020) for Dublin (1945 to date); and Nicholas and Scherbina (2013) for New York (1920–1939).

<sup>50</sup> The use of the term “asset” is somewhat anachronistic and care has to be taken to understand the legal framework within which land and houses were organised in specific historical settings and which might significantly limit the uses to which land could be put. However, as argued here, markets for land and therefore the use of land as an asset can be widely observed throughout history.

<sup>51</sup> For a survey of the literature on early land markets in Italy and the Netherlands see Bavel (2016).

<sup>52</sup> With the existence and efficiency of markets as the central point of interest, this literature has focused on the share of market-intermediated transactions in total property transfers, the organisation of markets, price formation, volumes and characteristics of property traded, the participation of different social groups as well as the impact of these markets on the rest of the economy and society. Interest in the evolution of price levels per se, on the other hand, has been more limited, with a focus on rental values as measure of agricultural productivity, distribution of income and price indices used to deflate nominal wages. Long-run rent series are available for the usual suspects in early modern

Against this backdrop, the inclusion of real estate prices in the HMFS project is motivated by the close link we see between real estate and financial technology. Again, this link is old and property has long been used as a means to secure debt. Athenians pledged land as ὑποθήκη (or hypotheke, literally something placed under; pledged) and they placed stones called *horoi* on it to mark the encumbrance. By the late 15th century, land registers in the Low Countries provided legal guarantees to creditors and allowed borrowers to access lenders outside their personal networks, creating active markets in mortgages (Bochove et al (2015)). Ito (2013) surveys political controversies on the creation of land registries in 17th century Britain.<sup>53</sup> Bellhouse (2017) shows the role played by land rental/purchase contract templates in the rise of “dematerialised” life insurance in 18th century Britain.<sup>54</sup>

Land transactions produced written records and, because of their importance as proof of ownership, these records were stored and are available far back in time, creating as yet not fully tapped opportunities for data harvesting. Historically, there are many, many prices for land and houses sold. The key issue is how to transform this raw material into statistics that measure what we want to measure. The challenge for the construction of historical property price indices is the sheer amount of transactions, combined with the idiosyncrasies of each individual object transacted. Recent methodological advances in statistics provide a menu of techniques that help overcome the latter issue by controlling for quality. In parallel, large-scale digitisation has rendered more sources more easily accessible. Computational advances, such as OCR and AI, (at least) promise automation in the gathering of standardised information and allow the organisation of much larger databases. At the same time, this more readily available information from periods further back in time creates demand for historians able to assess the context, meaning and implications of the data collected.

Accordingly, today we have quite a good understanding of what needs to be done as well as of the primary material to do it with. In the context of this chapter we will therefore specifically focus on the questions that come across when working with historical property prices. In particular, we survey difficulties encountered when constructing historical price series and methods available to address them. As for other data elaborated under the auspices of HMFS, the process is open-ended and focused on transparency. Transparency is also key for our goal of achieving historical HPIs that lend themselves more easily to international comparisons and interpretations. We distinguish three levels of documentation:

Firstly, the provision of detailed evidence of the primary data sources and the historical context in which they were produced including the markets they related to (in particular, information on the geographical scope, types of properties traded and agents active in this market as well as the representativeness of the source for the broader property market).

Secondly, detailed accounts and explanations of the splicing procedures and potential break adjustments, which are required when constructing composite historical HPIs spanning a long time period.

European history: Amsterdam, Antwerp, Brussels, Bruges, Ghent and Milan, among others (see references cited in Drelichman and González Aguado (2014)).

<sup>53</sup> See Clark and Clark (2001).

<sup>54</sup> On this point, see also the chapter on interest rates in this volume.

Thirdly, a discussion of the methods used in constructing the series for analysing the research question at hand, enabling interested readers to fully interpret, discuss, criticise and improve upon the approach.<sup>55</sup>

Before concluding the introduction, a short note is in order about what this chapter is not going to cover. First, the calculation of historical rent indices, although a crucial piece of information on property markets, comes along with its own specific challenges and merits separate treatment. Second, despite their historical importance, land prices will also only be mentioned in passing, as well as prices for commercial property. In the conclusion, we will, however, mention all three as promising fields of future interest in the context of HMFS.

The rest of the chapter is structured as follows. The next section sketches the principal methods available, their data requirements, and specific advantages and disadvantages in the context of historical research. Section 4.3 looks at the historical evolution of property markets and primary sources. Section 4.4 concludes.

## 4.2 Methods

Property prices have been of interest and have been monitored ever since markets existed. Moving from observations of individual sales to an index is tricky and it is only relatively recently that methods have been developed to systematically separate out changes in the composition and quality of property from the general trend in prices.<sup>56</sup> Interest in aggregate measures increased as the boom-bust episodes in Japanese commercial property in the 1980s and Scandinavia in the 1990s pointed observers to the interaction between property markets, financial stability and the macroeconomy. The BIS started to pay closer attention to developments in property markets in the late 1980s and early 1990s (BIS (1990)) and it was soon recognised that there were numerous problems involved in collecting and maintaining reliable and comparable data sets for the prices of houses and commercial property. This was even before we observed major fluctuations in prices in asset markets of the late 1990s and early part of the 2000s, which of course spurred further requests for analysis of financial stability issues and a need for better data on asset prices, including those for residential and commercial properties.

In an effort comparable with the earlier creation and standardisation of data on national accounting, balance of payments and consumer prices at the international level, the BIS joined forces with the IMF in 2003, recognising that in almost no country did property price statistics seem to meet the criteria of data integrity, methodological soundness, accuracy, reliability and serviceability.<sup>57</sup> Some countries

<sup>55</sup> For example, in their online appendix, Knoll et al (2017) score both levels two and three. Thus it seems to be feasible to trace the information about house price data and splicing procedures for many of the 14 countries back to their respective origins.

<sup>56</sup> For two early cases of repeat-sales indices see Duon (1943) and Grebler et al (1956).

<sup>57</sup> See the IMF's Data Quality Assessment Framework (DQAF). David Dodge, when Governor of the Bank of Canada, complained in a speech to the Conference of European Statisticians in June 2003: "Given that the investment in housing represents a big chunk of household spending, and that for most people their homes represent their most valuable asset, it is surprising that in many countries there are no comprehensive, quality-adjusted data on housing prices or rents." Mervyn King, when Governor of the Bank of England, is said to have complained about "statistical fog" when he was reviewing the trend in house prices in the United Kingdom. It was unclear whether UK house price inflation had come to a turning point or not.

were, at the time, starting to construct their first house price indices, while other countries already had several different house price indices available. These were, however, based on different coverage (geographical area, dwelling type, new versus existing dwellings), different data sources (ask prices, value appraisals, final sale prices), and different construction methods. For some countries, there were only crude statistical measures of mean or median house prices available, without any explicit control for quality differences (dwellings that differ in type, size, amenities and geographical location). Other countries used one of the many methods which are available today to adjust for such quality differences, eg stratified mix-adjustment methods, hedonic models, repeat sale models, hybrid models and SPAR models.

A comprehensive overview of the data situation on housing prices was published in BIS (2005), with recommendations for future improvements.<sup>58</sup> This work has been followed up through the publication of international standards on HPI methodology, in particular Eurostat (2013), *Handbook on residential property price indices* (RPPI handbook) and Eurostat (2017), *Technical manual on owner-occupied housing and house price indices* (OOH manual). During the past years the overall availability of house price indices has greatly improved.<sup>59</sup> Nowadays, indices are typically based on actual transaction prices, and are published regularly on a monthly or quarterly basis, often, but not always, by the national statistical agency. Sometimes indices are produced by a private company in the real estate or financial sector. Given the heterogeneous nature of the housing market, a vast number of HPIs have appeared for different geographical units like cities, sub-regions and states, as well as for different types of houses, like single homes, small semi-detached houses and apartments.

In some countries we have also observed recent efforts to improve the existing house price indices and/or the construction of new ones.<sup>60</sup> There have also been numerous contributions to the academic literature discussing pros and cons of the different methodologies involved in constructing quality-adjusted HPIs.<sup>61</sup>

### The representative house

Wang and Zorn (1997) criticise the academic debate on the properties of HPIs for a lack of focus because what the different HPIs aim to achieve is left unspecified. An HPI may be “good” for one goal and “bad” for another. Across the board, a widely accepted and reasonable objective for an HPI is to capture the general developments in house prices for a “representative house” located in a given area. In this case, one

<sup>58</sup> The recommendations underlined the importance of clearly stating the goal and purpose of making the HPI (why?) and that the methods adopted by statistical agencies to construct constant-quality HPIs (how?) would vary among countries and would be dictated in large part by the availability of data generated by the processes involved in buying and selling a property (Eurostat (2013), p 156–7).

<sup>59</sup> See National Statistician (2010), see Eurostat (2013) for some case studies, Mack et al (2011) documents an international house price database, maintained at the Federal Reserve Bank of Dallas, primarily collected from a wide range of national public sources, which contains quarterly data for 22 countries going back to the first quarter of 1975. Knoll et al (2017) have recently documented an international house price database which contains annual historical HPIs for 14 countries going all the way back to 1870. This historical database is updated and maintained as part of the international macro-history database project documented in Jordà et al (2016, 2017).

<sup>60</sup> See Office for National Statistics (2013) for a recent example for the United Kingdom.

<sup>61</sup> A series of comparative studies have been conducted over the past two decades, see eg Gatzlaff and Ling (1994), Meese and Wallace (1997), Wang and Zorn (1997), de Vries et al (2009) and Silver (2011, 2014).

must decide whether the HPI should be representative of the entire housing stock (stock-weighted) or of the subset of existing houses which are traded in a given period (sales-weighted). If the area is a country this also requires solving the aggregation problem, which arises from substantial regional variations in house price developments across individual geographical sub-areas.

The measurement of HPIs faces two main methodological challenges largely due to two main characteristics of the housing market. Firstly, houses are inherently heterogeneous in nature because of obvious quality differences between houses with different dwelling types, sizes, amenities and locations. Secondly, houses are put on the market infrequently. Only a small fraction of the housing stock is subject to a sale/purchase transaction in any given year. House price indices based only on broad summary statistics, such as the annual mean or median sales price, may be of some interest as they are more readily available, but they are crude indicators of house price developments. However, it is important to recognise that such crude indicators might be heavily affected by compositional effects. The quality of a *given* house may also change over time as a result of renovation and depreciation. As a result, changes in simple price index measures like the mean or the median can be due to changes in the composition of houses transacted, themselves reflecting, say, investment in real assets, rather than in the underlying general level of housing prices. Quality adjustment is required but this is data-intensive.

Ideally, an HPI would allow us to identify some factor at time  $t$ ,  $P(t)$ , which represents the general house price development for a representative house, and separate it from other factors which characterise differences in quality, location and amenities,  $Q(i, t)$ , such that the price of an individual house,  $P(i, t)$ , can be decomposed and written as:

$$(4.1) P(i, t) = P(t) + Q(i, t) + \varepsilon(i, t)$$

where a mean zero stochastic variable  $\varepsilon(i, t)$  represents a random error term. In short, we can think of a “first-best” HPI as a quality-adjusted HPI.

In the following paragraphs we provide a brief overview of the methods to control for quality that are used in practice, their informational requirements and their main advantages and disadvantages.<sup>62</sup> We also include some paragraphs on “second-best” methods, when information on individual real estate transactions and/or information on quality is insufficient or missing.

### Hedonic model

Hedonic price indices are based on a regression of individual house prices on data for available housing characteristics.<sup>63</sup> They use information on  $Q$  to back out the general price component  $P$ . Hedonic price indices have a number of advantages. They provide for an efficient use of available information, as all data contribute to estimating property prices. The method controls for changes in the property mix in the sample and, if sufficient property characteristics are included, can equally control for differences in quality. Unlike in methods using stratification, ie the creation of subsamples as explained below, price estimates can also be produced in situations

<sup>62</sup> This overview borrows from a similar overview for the United Kingdom (Office for National Statistics (2013, Section 2)), for a more detailed technical discussion, see Eurostat (2013).

<sup>63</sup> See Rosen (1974) for theoretical foundation and Hill (2013) for a recent summary of the literature on hedonic house price models.

when certain house types or locations have only a few, or even zero, recorded transactions in a given period.

Hedonic models however, also come with a number of disadvantages. Data requirements are high as information on all key housing characteristics, which capture quality differences, is needed. In addition, hedonic regression models can be implemented in many different ways, based on the available housing characteristics, choice of functional form etc. As a result, the method is subject to bias if key house quality characteristics are left out from the model, or if the parameters, which represent the true effects of these quality characteristics, are subject to non-modelled time variations. In particular – and this is an issue for very long historical series – hedonic price indices ought, in principle, to deal with *changing valuations of quality characteristics*. The consequence is that there are very few, if any, studies, which have data on quality characteristics over a very long period.<sup>64</sup>

HPIs which are based on model estimation, like hedonic indices (this is also true for repeat sale and hybrid models), are typically subject to revisions as time goes by and the sample is updated. Repeat sale models suffer from revisions shown to be of a larger order of magnitude when compared, say, with hedonic models. HPIs may also be designed such that historical values are never revised, eg using chain-weighting.<sup>65</sup>

### Repeat sales model

Repeat sales models offer an alternative if only a limited amount of information on housing characteristics is available. Repeat sales HPIs are based on a regression of the changes in individual house prices between two consecutive sale transactions of the same house, on a time-varying constant term. In some applications this method makes use of data on available housing characteristics, in which case we denote this as a “hybrid model” (combining the repeat sales and hedonic models).<sup>66</sup>

The key advantage of repeat sales models is that they only need a reliable way to identify and match the traded properties. The method does not rely on housing characteristics between matching pairs, and automatically controls for the traded property mix, location and qualities that can be hard to represent in a hedonic model, or can be unobservable. The method does not require a correct functional form to be found, as is crucial for the hedonic model. At the same time, however, the method is wasteful of data as it ignores all available data on single-sale transactions, as of the houses traded in a particular period, only some have been traded previously. The efficiency of the method thus depends on the length of the period over which price data are available as well as the frequency with which houses are traded, which varies across markets and time.<sup>67</sup> In markets with little turnover, the method is less easy to apply. While repeat sale models control for the main characteristics of the property, changes in those characteristics (renovation, depreciation, extensions and conversions) are not automatically taken into account. In particular, properties that

<sup>64</sup> Depending on the implementation, the method may also give rise to HPI revisions when the sample is updated and coefficients change retrospectively. In the construction of historical price indices, this issue is of minor concern, however.

<sup>65</sup> See eg Shiller (1993), Clapham et al (2006) and Silverstein (2014) for further discussion.

<sup>66</sup> See Englund et al (1998), Clapp and Giaccotto (1998), Englund et al (1999a,b) and Hwang and Quigley (2004).

<sup>67</sup> When a new repeat sales pair is included in the sample this entails revisions of the estimated index. Again, this is of no concern for historical indices, which operate on a fixed set of data.



are traded after a long period may be subject to significant quality changes.<sup>68</sup> Also, because the most frequently traded houses tend to have properties which make them less representative, such as “lemons”, starter homes or flip sales, this can lead to a sample selection bias.<sup>69</sup> Finally, the practical possibility of creating a repeat sales index is subject to the organisation of the primary material. If land registries classify the documents containing the prices chronologically and not by property, sampling can easily become prohibitively expensive.

## SPAR

Like the repeat sales model, the SPAR model builds on the comparison of matched pairs of prices collected at two different points in time. Instead of using two transaction prices, however, SPAR matches the individual house sale price to a corresponding value appraisal of the same house, which may have been made for taxation or some other administrative purpose.<sup>70</sup> The main advantage of SPAR is its efficiency as, in principle, all sales data can be used, apart from those sales that cannot be matched with an appraisal or those of properties built after the appraisal period. It is easy to implement from administrative sources when these are available. If property valuations are done frequently and these take into account substantial improvements to the properties, this method is better able to deal with quality changes than a repeat sales model. No information on the characteristics of the traded houses is needed. At the same time, however, appropriate sources with good coverage are required. The method hinges, critically, on the quality of the appraisals.<sup>71</sup> If appraisals are subject to structural shifts, say from taxation values to refinance appraisals, or cross-sectional variations in quality, this will potentially bias SPAR HPIs. In many cases not all regions or all appraisers use the same methods. If appraisals are infrequent the method cannot deal with quality changes.

## Mix-adjustment with a stratification matrix

A final alternative is provided by mix-adjustment methods. These create HPIs based on statistical summary measures, such as means or medians, from a sample of traded houses. These are subject to a weighting procedure which meets the requirements for a “representative house”, such that different house types – like single family homes, semi-detached houses and apartments – and different geographical locations

<sup>68</sup> Some authors suggest giving such transactions a lower weight in the index (Case and Shiller (1987)).

<sup>69</sup> Repeat sale indices have been popular in the United States since the late 1980s, largely due to the impact of the work on HPIs by Case and Shiller (1987, 1989) who improved on the original work by Bailey et al (1963). Their potential weaknesses have been studied in a series of academic articles, see eg Goodman and Thibodeau (1998) (heteroscedasticity); Goetzmann and Peng (2002) (unbiasedness); Pennington-Cross (2005) (aggregation bias); Clapham et al (2006) (revisions); Harding et al (2007) (depreciation); Francke (2010) (thin markets); and Silverstein (2014) (revisions).

<sup>70</sup> HPIs based on SPAR have been popular in the Netherlands, New Zealand and Sweden, see Chinloy et al (1997), Bourassa et al (2006) and de Vries et al (2009). For Sweden the recent HMFS project undertaken by Sveriges Riksbank has documented the construction of historical HPIs for two metropolitan areas, Gothenburg (Bohlin (2014)) and Stockholm (Söderberg et al (2014)), using two alternative methods, the repeat sale method and the SPAR-method, from 1875 to 1957.

<sup>71</sup> See Shimizu and Nishimura (2006, 2007) who have compared HPIs for Japan based on appraisal values and transaction prices over the 25-year period 1975–1999, and reported large and systematic discrepancies between actual transaction prices and the published land prices, which indicated substantial biases (valuation errors) in appraisal land price information for Japan. The appraisers used different approaches to value land, the comparison approach, the income capitalisation approach and the cost approach, and it is unclear how the valuations are weighted.

are all represented to a sufficient degree in the sample of traded houses. The main advantages of mix-adjustment are the lower information requirements compared with hedonic models while at the same time no matching is needed as in the case of repeat sales and SPAR. The stratification variables typically used are the property type and location, which are both easily available. The procedure is also straightforward to understand and assess. The main disadvantage of mix-adjustment is the effective limit on the number of stratification variables that can be used. Increasing the number of stratification variables  $x$  and categories  $c$  rapidly leads to a high number of subsamples  $c^x$  with potentially too few observations or even some combination of, say, house type and location, for which there are no trades in a given period. Good judgment is therefore required when choosing the stratification variables.

#### Indirect methods based on rental value and construction costs<sup>72</sup>

If no direct evidence on transaction prices is available, there are a number of indirect methods that have been used to obtain a valuation of property. The first of these second-best methods is to capitalise the value of rental payments on dwellings. To capitalise the rental value into a selling price, a rate of return or discount factor can be used with a standard present value formulation. However, often a direct measure of capitalisation – known as the “years purchase” price of rented property – may also be available. For example, in the United Kingdom the sales prices of a rented property and land were often quoted in terms of the rental value and number of “years purchase” – the number of years of rent required to buy the property. Often properties with different rentals might be sold at the same or similar “years purchase” that reflected the overall expected return on land/property as an investment. That implies rental values derived from a broad set of properties might be combined with a more limited sample of years purchase data to create an aggregate property price index.

$$\text{Property Price} = \text{Rental Value} \times \text{Years Purchase}$$

Holmans (2005) was able to construct such an index for the United Kingdom for the period before the First World War. This was until recently the only viable method of constructing a house price index starting in the late 19th century. He used a series for rents derived from valuations for house duty which he combined with “years purchase” information from a separate tax – estate duty data – to convert these into a house price index for the period 1895–1913.

A second indirect method is the cost function approach. This assumes that the price of a residential property can be adequately captured by the cost of building the dwelling plus the cost of the land on which it is built. With adequate indices for building costs and land prices, a proxy for residential house prices can be constructed with the use of a cost function  $c$ .

$$\text{Property Price} = c(\text{Price of Land}, \text{Cost of Construction})$$

The issue here is specifying the cost function, which will embody the implicit substitutability between land and the structure. Higher land prices will imply building more units on a given plot and the degree of substitutability between land and structure will imply different cost functions and/or parameters. Often, simple assumptions like fixed proportions or Cobb-Douglas specifications are used but

<sup>72</sup> Thanks to Ryland Thomas for contributing to this text by explaining how we may apply indirect methods to construct historical HPIs.

indices may need to be chain-linked if substitutability is imperfect and the share of land in the cost of the property is changing over time. Also, the index created will strictly apply only to newly built homes and will be less adequate as a proxy for the overall stock of housing. Braae (1960), for example, uses a cost function method for the United Kingdom for the period 1920–39 in which house prices are derived from available construction costs and land prices.<sup>73</sup>

Neither indirect method is likely to meet many of the criteria for a good HPI. Although some stratification of data on rents and land might be possible, the sub-components used in both methods are likely to apply to different samples of the housing stock and so only a rough approximation to the price of an average house is likely to be obtained. In the case of the cost function approach, good data on land prices are required. Building costs for properties may be producer price indices or measures of input costs (based on construction wages and the cost of materials). Typically, these are more relevant for the price of new homes. Residential land price data may be difficult to acquire but proxies may be available. Again the price of agricultural land might have to be used as a proxy for residential land, at least for rural areas.

Perhaps one advantage of these measures is that they may be able to cover a broader set of properties and regions than some of the first-best methods if, for example, the data on rents cover a wider geographical region than transactions or appraisal prices. They may therefore provide a cross-check on house price estimates that use a better methodology but only have a narrow regional or city-based coverage. Indirect methods may also be useful to decompose and evaluate house price movements to derive the respective contribution of the cost of land and the dwelling itself, as done in many of the sources relied upon in Knoll et al (2017). Finally, they can be used to construct estimates of “fundamental” house prices which can be set against actual price indices from transactions to evaluate bubbles and price overhangs. For example, a long-term risk-free rate can be used to capitalise the value of rentals to create such a fundamental measure.

## Splicing

Constructing long-run HPIs often requires splicing together HPIs of different types between different subperiods. Exactly as would occur in fact for any price index, alternative HPIs based on the different methods described above coexist for a given period in the history of a country, reflecting different primary material, methodologies, types of property or geographic areas. Since different HPIs were designed to meet different targets, one has to exercise care and judgment when splicing different HPIs to form a long-run time series. For lack of an alternative, crude statistical measures that do not control for changes in quality must sometimes be combined with more high-end HPIs that attempt to control for quality changes. In other cases, splicing involves HPIs generated by heterogeneous methods.<sup>74</sup>

<sup>73</sup> From 1928 to 1939 Braae’s source was estimated average construction costs for private dwellings for which plans were approved by local authorities. For years before 1928 values were estimated from contract prices for new houses for local authorities.

<sup>74</sup> Knoll et al (2017, Table 1, p 336) provides an overview for 14 countries, see also Ryland Thomas’s documentation of historical HPIs for the United Kingdom in the appendix to this publication.

## Quality-adjusted historical HPIs for Norway 1819–2020

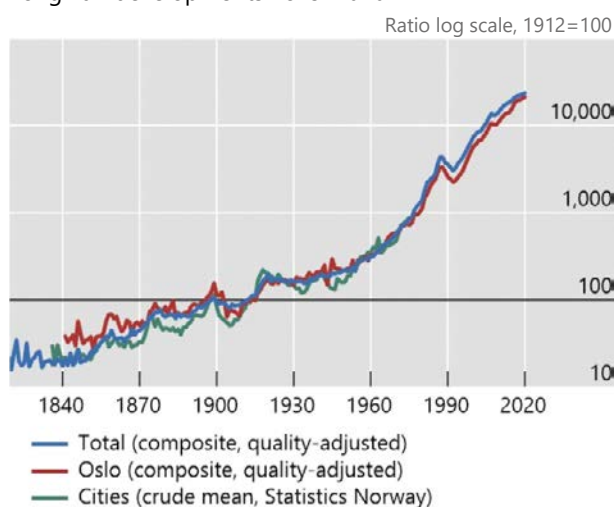
Graph IV.1 (left-hand panel) shows two composite constant-quality HPIs for Norway that provide estimates for the capital Oslo (Christiania) since the 1840s (red line) and estimates of a total country average since 1819 (blue line). For comparison, we have also included a data series reported by Statistics Norway up to 1975, which reports crude averages of traded properties in Norwegian cities (black line). Constant-quality HPIs control for changes in the decomposition of the properties that happen to be traded in a period. We observe notable differences between the observed HPIs for Norway, which underscores the benefits of controlling for quality in a systematic way, say, by controlling for the composition of the type of dwellings which are traded, and important quality attributes such as their size and location.

One example is developments during the turbulent 1920s and 1930s, starting with high inflation during and immediately after the First World War, followed by a severe post-war recession, deflation and a systemic banking crisis. This was followed by a period characterised by rapid resumption of the exchange rate back to its pre-war gold parity, which was completed in 1928. Then came the international downturn following the 1929 stock market crash, leading to the Great Depression. During these crisis years, in our case exemplified by 1920s Norway, it is of particular importance to control for changes in the composition of traded houses, since this composition may be quite different from what we observe in normal years.

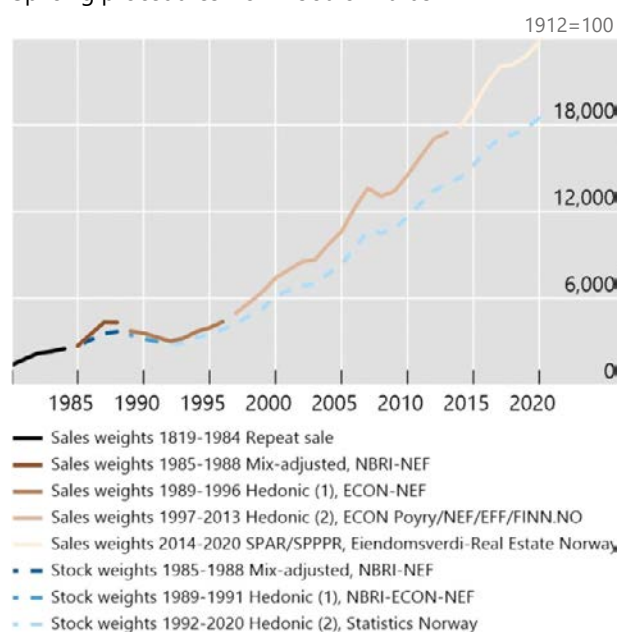
### HPIs for Norway

Graph IV.1

#### Long-run developments 1819–2020



#### Splicing procedures from 1980 onwards



Sources: Eitrheim (2022) and sources cited therein; Statistics Norway

The right-hand panel in Graph IV.1 shows more details about the splicing of available composite constant-quality national HPIs for Norway since 1980. We have set all HPIs equal to 100 in 1912 and report the splicing of two main types of HPIs, which, over this long period of time, have been made available by different researchers and public and private producers of house price statistics.

The splicing of HPIs based on sales weights involves the combination of five types of HPI starting with a repeat sale HPI followed by one mix-adjustment, two hedonic (see Econ (2004) for details) and one SPAR-type HPI. These HPIs have been developed by researchers (Eitrheim and Erlandsen (2004)) or in different private partnerships like the current one mentioned above between Real Estate Norway and Eiendomsverdi AS.

The splicing of HPIs for Norway based on stock weights from 1985 onwards involves the combination of a mix-adjustment HPI and two types of hedonic HPIs, one produced by the original partnership between the Real Estate Agents Association (NEF) and the Norwegian Building Research Institute – which was established in the late 1980s – from 1989 onwards, and one from Statistics Norway from 1992 onwards.

The main distinction is between HPIs based on sales weights – indicated with solid lines in Graph IV.1 (right-hand panel) – which change with the sample of traded houses in a period; and HPIs based on stock weights, indicated with dotted lines. The latter translate price changes observed in the sample of traded houses to estimates of changes in the value of the total housing stock in the country by applying a different weighting procedure. We note that from 1985 onwards, the sales-weighted HPIs show stronger growth and larger fluctuations than the stock-weighted HPIs. A more complete discussion about the real-time properties of these two types of HPIs is offered in Eitheim (2022).

The composite constant-quality HPIs splice together, for each of the two types of HPI, indices which are based on some of the methods listed in Section 4.2. The variation in the use of these methods across this long period reflect the fact that constant-quality HPIs may have been constructed for different purposes by different producers, but also that the amount of information available may have been limited, which may have restricted the choice of method.

The data for Norway illustrate an example of the latter consideration since before 1985 the only available historical HPIs for Norway are repeat sale HPIs, which are based on sales weights, eg the estimates for the capital Oslo (Christiania) and the total country average shown in Graph IV.1 (left-hand panel) above.

### *Summary*

At present there are two sets of HPIs which are published on a regular basis. The first is a sales-weighted monthly HPI, a SPAR-type HPI (sales price to predicted price ratio (SPPPR)), which is produced in a private partnership between Real Estate Norway and Eiendomsverdi AS, a private bank-owned firm. The second is an official stock-weighted quarterly HPI, a hedonic HPI, which is published by Statistics Norway.

Whereas the HPI from Statistics Norway is publicly available and well documented in a series of publications, which can be downloaded from Statistics Norway's website, the monthly HPIs produced by Real Estate Norway/Eiendomsverdi AS are proprietary and subject to licensing, and there is only limited documentation available, which offer little detail about the construction procedures and model properties. Eiendomsverdi AS provides Automated Valuation Model (AVM) services to a variety of clients, including real estate agents and financial institutions.

Against this backdrop, splicing calls for solid documentation of the series' characteristics and a discussion about why the behaviour of the old and the new series is expected to be similar (or different). Generally, it is useful to have a period when both series overlap. In practice, spliced series often increase geographical scope over time, with earlier indices covering a single or several important cities, and later extend to the entire national territory. In these cases, it might be of interest to construct and report a sub-index that refers to the initial smaller geographical unit alongside the new national index, at least for a time of transition. The composite HPIs may also need some break adjustments accompanied by words of caution in their use. Growth rates are typically found to be more reliable than levels. Composite long-run housing price series for Norway can serve as an example (Box IV.1).

### Selecting among the methods

Current handbooks on the construction of HPIs discuss fine-grained hierarchies among alternative methods.<sup>75</sup> Hedonic HPIs are typically ranked above other approaches owing to their ability – given sufficient information – to capture more explicitly individual quality changes and isolate “pure” price movements. Furthermore,

<sup>75</sup> See eg the OOH-manual by Eurostat (2017).

within the class of hedonic HPIs, distinctions are made between alternatives.<sup>76</sup> We will argue that for our purpose, with a strong focus on the documentation of historical data, it suffices to distinguish more broadly between HPIs that are quality-adjusted and HPIs that are not. The choice of a method is then driven by the available primary material as well as the specifics of the market under study, which can make one method better suited than another.

As stated, methodological choices also depend on the research objectives. To give an example, different priorities regarding the use of stock-weighted or sales-weighted HPIs are called for. Depending on whether the value of the entire housing stock or the value of the houses sold at a given point in time is the variable of interest, different priorities regarding the use of stock-weighted or sales-weighted HPIs are called for.<sup>77</sup> From a pragmatic point of view, there may be no well defined hierarchy of methodological choices, but rather a need to clearly explain the reasons for the choices taken. In fact, methodologies can be (and are) combined in hybrid HPI models. Such combinations might be attractive in a historical context if particular types of information are not available.<sup>78</sup> Due to the relative novelty of the collection of long-run HPIs, an open-minded approach is strongly recommended. We have not yet reached the stage where anything like a best historical practice can be defined.

<sup>76</sup> For instance, there are distinctions between time dummy, rolling time dummy, imputation, average quality and repricing methods. A discussion of pros and cons on this level of detail falls outside the scope of the HMFS project and the interested reader is pointed towards Eurostat (2017).

<sup>77</sup> HPIs are also used today for many different purposes and by many different groups of users. They are, of course, of great value for the real estate sector and credit institutions involved in the house sale process such as mortgage lending institutions. But HPIs are also of value for other groups of users like policymakers, business cycle analysts and insurance companies, to mention a few. Preferences concerning the construction of HPIs differs accordingly.

<sup>78</sup> Examples include combinations of sales prices and appraisal values in repeat sale HPIs; Clapp and Giaccotto (1992) use appraisal values and sales prices to construct a repeat sale HPI, while Gatzlaff and Ling (1994) use the two measures in reverse order. SPAR-type HPIs use a combination of sales prices and predicted prices from hedonic models. SPAR-type hybrid HPIs for Norway, produced in a private sector partnership, cover monthly indices for 23 regions, three house types and a nationwide index for the country as a whole. The method combines transaction prices and estimated appraisal values generated from house price predictions from hedonic pricing model using matching housing characteristics. See the chapter on Norway in the accompanying BIS Paper for more details.

## The house sale timeline – which price to select?

Any sale transaction is typically associated with several prices (which refer to different concepts), are collected by different agents and stored in different sources. This multitude of available prices has been discussed in the context of the production of current HPIs, regarding which the availability of data in real time is a concern. Time lags to publication vary and some HPIs are constructed from the price information available early in the house selling process, which may be subject to later revisions when final sale prices are registered. Thus, there will be a trade-off between user demands for timeliness and the desire for HPIs to have minimal revisions. For historians it is important to keep in mind that prices generated at different stages of the sale timeline have different meanings and characteristics.

Prices can be ordered along the timeline of a real estate transaction. Today, initial ask prices are readily available from specialised webpages, while in earlier years these were announced in printed news media. However, since not all houses on the market are actually sold, and, for those houses which are sold the final transaction prices may deviate considerably from the initial ask prices, most available HPIs are based on final sale prices. For earlier years, final sale prices eventually became available after they were registered in the land registry, typically at regional notaries' offices. Today, sale prices are available from modern land registry databases, some of which are operated and maintained by government-owned companies, which may demand some fees for their services.

However, since most people in the process of buying a house today also need a mortgage loan to finalise the transaction, there are often more timely data available on ask prices, appraisal values and final sales prices from real estate agents and their associations, and from mortgage companies who are involved in the house purchase process. Some of these private sector agencies also maintain databases, whose coverage of transaction prices is steadily increasing, and which include information about characteristics of the housing stock, which are relevant for hedonic models.

In the light of these developments, available HPIs may vary according to whether the index is based on the initial ask prices, on appraisal values used for securing a loan, on the final transaction price registered by a private sector agency or the land registry, or on a combination of the latter two.

In the context of historical price indices, the timeliness of publication and the revision of the index following the inclusion of additional data are of less concern. There is however the question of whether there are systemic biases if offer prices or assessed values are used instead of transaction prices, and how these biases can be obviated.<sup>①</sup>

① Shimizu et al (2011, 2012) found substantial differences between the statistical distributions of housing prices stemming from different sources in the sale/purchase process. However, when they controlled for quality differences only small differences remain.

## 4.3 Historical context and primary sources

Societies differ widely in the ways in which the ownership and use of land and real property are organised. Agricultural land can be held by small farmers, rented out by large landowners or cultivated directly by them. The same is true for urban housing. In some cities most people live in single-family dwellings, which they own, while in others, most people rent apartments in multi-family residential buildings, which are owned by investors or public institutions.<sup>79</sup>

Such structural features determine the extent of the market, in terms of both *who* participates as buyer and seller as well as *what* is bought and sold, ie the share of the total housing stock that is regularly transacted and its characteristics. Causality runs in both directions of course, and market performance will in turn influence the evolution of the city structure. In the same vein, the long existence of active markets in land and housing should not lead one to underestimate the role of regulation. This is obviously true for feudal Europe with its significant restrictions on the rights of

<sup>79</sup> For an overview of the various forms of organising the spatial structure, housing, transport and segregation observed in modern cities, see eg Paddison (2001).

landlords, tenants and peasants to sell land and use it to secure creditors. Transactions outside the market through the family, village communities or the manor can coexist with markets and, in specific historical periods, these transactions predominated. But modern land and housing markets have been subject to government intervention as well, concerning rents and zoning which limited the types of use to which land can be put and, crucially, they depended on public investments in infrastructure.<sup>80</sup> Lescure (1980, 1982) provides a classic example of the interactions that develop between government, the capital market and the housing market in his discussion of the long-run evolution of the French real estate market (1820–1940).

The main point is that there is no clear historical break between an early modern world based on customary transfers and a modern world based on market exchange (Polanyi (1944)). Instead, markets and price building always need to be understood in their institutional context.

## Context

Investigating and exploring the specific environment of historical housing markets brings several benefits. First, structural features influence the representativeness of the recorded market transactions included in the calculation of an index. Many long-term studies such as for Amsterdam (Eichholtz et al (1997)) rely on a small number of properties located in the city centre, since they are the only ones for which repeated sales are available over a long stretch of time. Often, indices used to track housing prices at the country level in fact refer to the capital city or a handful of large, economically important, conurbations. Alternatively, they may document agricultural land. If the distribution of prices in space follows a well defined gradient as posited in the monocentric models dating back to Ricardo and Von Thünen, evidence from city centres might well be representative of developments in a larger geographical area. However, socio-economic and technological factors, such as faster transportation, can change patterns of spatial market integration over time, and markets that, for a time, were leading the general trend might become subject to idiosyncratic developments (Meen (2016)).<sup>81</sup> Another interesting example for the importance of structural features in the choice of the relevant market is provided by Japan, where the separation of the ownership of houses and land in the 1870s meant that the price of land rather than houses became the main indicator for the real estate market.<sup>82</sup>

Second, the impact of property prices on, and interaction with, the financial system also depends on the structure of the housing market. A price boom in a market dominated by large investors will have different implications than in a market with many small owner-occupiers. Access to credit and therefore the degree of leverage can vary, and so too can the financial instruments that are linked to property. The structure of the financial system and notably of mortgage finance, in turn, impacts the types of houses built, their owners and the markets in which they are transacted

<sup>80</sup> For a long-run view on rent control, see Willis (1950).

<sup>81</sup> An example at the city level would be the decline in the price of inner-city residential property in the United States in the 1960s and the migration of affluent urban dwellers to the suburbs (Collins and Margo (2007)). At the national level, Eitrheim and Erlandsen (2005) document the decoupling of house prices in Kristiania (Oslo) from trends in other cities during the 1890s boom. The case of Kristiania is also interesting with respect to the decline in market activity that followed the end of the boom (See Box IV.2).

<sup>82</sup> See the chapter on Japan in the accompanying BIS Paper.



(Blackwell and Kohl (2019)). Ideally, therefore, house price indices should come with a description of the market where they were produced.

Finally, institutional features influence the availability and organisation of the primary documentation. In systems characterised by a high number of owner-occupied single homes there will be more transactions and more publicly available information, eg in newspaper advertisements. If, on the other hand, most real estate is in the hands of long-term institutional investors, the number of transactions might be limited, some types of property might never make it to the market, and offers and transactions might only be covered in special media or kept as private information. In the two settings different types of historical records will be produced.

## Sources

Quality-adjusted indices require both information on prices and on the characteristics of the properties traded. These are often found in sales contracts, which naturally provide information on price and key characteristics of the property traded such as location, size etc. Historically, contracts were archived because buyers and sellers needed proof of ownership and some of that material has survived.

For the early modern and modern periods, an important role is played by notarial records, although their features vary widely depending on the particular organisation of the notary as an institution. This includes the existence (or not) of a legal obligation to have the transaction certified by a notary. In France, Italy, Spain and many Latin American countries, notarial archives provide rich repositories for real estate transactions (Poisson (1985), Hoffman et al (2000)). In other regions and countries, information on contracts was preserved in public registries, serving as proof in case of disputes and allowing buyers to claim good faith when a transaction was not recorded in the register. In Britain, a number of local deed registries were established in the early modern era until the Land Registry Act 1862 rendered the process uniform and resulted in the creation of a centralised depository. In German speaking countries, the medieval *Stadtbücher* and *Urbare* were local forerunners of modern land registries set up, as in Britain, over the course of the 19th century. In Qing China the government provided standardised contracts for house sales, subject to a proportional stamp tax, and kept the record in public archives.<sup>83</sup>

Another source for price information are government surveys. Governments started to set up cadastres of land ownership for tax purposes early on and these included information on plot size, use and yields, although accuracy and detail varied.<sup>84</sup> An early example is the Florence catasto of 1427, which has been widely exploited in the historical literature (Herlihy and Klapisch-Zuber (1978)). Unlike contracts, cadastres estimate values or yields. The question becomes how property values were assessed. In Piedmont, *estimi* (estimates) available for some towns from as early as the mid-14th century, were regularly updated and mechanisms were present that ensured that assessments closely followed changes in market values (Alfani (2015)). In principle, therefore, such sources could be used to track secular trends in property prices.

<sup>83</sup> Raff et al (2013) use these archival records to derive a hedonic price index for Beijing during the period 1644–1840.

<sup>84</sup> The history of the national cadastre starts early, in England with William the Conqueror who initiated the *Domesday book* of the 11th century, in Norway the history of the cadastre is traced back to King Magnus the Law-mender in the 12th century.

Other fiscal sources do contain actual transaction prices. This is the case, for example, for the Paris *sommier foncier*, a document established by the city's administration responsible for the collection of taxes on legal acts, which compiled various information on owners, inheritances and rent contracts to check whether values declared for tax purposes were plausible.<sup>85</sup> The *sommier* starts around 1800 and served as the basis for (one of) the earliest repeat-sales price indices in the literature (Duon (1943)). In those cases, where fiscal sources contain assessed values only, these can still serve as input for SPAR-type models. As tax values were an important information for potential buyers, they were sometimes published. Bohlin (2014) and Söderberg et al (2014) use the official assessments reproduced in contemporary address books as base values in their SPAR-indices for Gothenburg and Stockholm.

In the second half of the 19th century, the newly created national statistical offices started to put together statistics by counting the total number and value of transactions recorded by the land registries. Aggregate statistics have the considerable advantage of covering the full set of property transactions in a given jurisdiction, but they allow the calculation of simple averages only. In some cases, when statistical publications provide separate totals by regions or by type of real estate, eg agricultural versus urban land or single- versus multi-family dwellings, it might at least be possible to use mix-adjustment methods and control for changes in composition when calculating a national aggregate index.<sup>86</sup>

But when researchers attempt to create a quality-adjusted index based on information on individual properties and transactions, there is no alternative to mining the mass of material stored in land registers and notarial archives. In addition to sales contracts, these archives include many other legal documents including inheritances, inter vivos transfers and mortgages; and the possibility of identifying the relevant information and transforming it into useable data hinges on the archive's organisation. Information about individual properties is easier to extract when the register or cadastre is spatially organised such that all relevant documents pertaining to an individual property are kept in separate ledgers, thereby facilitating the calculation of a repeat-sales index.<sup>87</sup> When only some of the relevant documents can be located in spatially organised sources such as notaries' registries or other available publications, important information on where to find the remaining information necessary to calculate HPIs may be present.<sup>88</sup> If locating specific contracts is not

<sup>85</sup> See also Friggit (2008). For some institutional background see Eichholtz et al (2019), while Eichholtz et al (2021) use the information on rents available in the *sommier* to calculate a repeat-rent index matched to the repeat-sales index to arrive at an estimate of the total return to Parisian residential real estate for the period 1809–1942.

<sup>86</sup> Carmona et al (2014) go a step further and calculate a hedonic index for interwar Spain based on aggregate data. They regress mean prices at the provincial level on average quality characteristics of the provincial housing stock to back out the general price trend  $P(t)$ .

<sup>87</sup> The primary data used in Eitrheim and Erlandsen (2004) are from the national city registers organised spatially after individual property unit numbers. The ledgers also contained necessary information about sales prices and housing characteristics to estimate repeat-sale indices. The individual property unit numbers used in the Norwegian cadastre were already in place from 1838. From 1935 data were available from scanned documents which could be downloaded from the webpages of Norsk Eiendomsinformasjon AS (today Ambita AS).

<sup>88</sup> In Sweden, local governments kept ledgers, ordered by individual property that contained the relevant documents on ownership titles and mortgages, but not transaction prices and taxation values, which were needed to calculate HPIs by the SPAR method. In Bohlin (2014) the author collected data on transaction prices and taxation values for the selected properties from

feasible because of the sheer volume of documents, the repeat-sales method will not be an option. Instead, sampling for a hedonic index may be the only possibility.

Turning from public to private sources, advertisements in newspapers and other media are important sources of price information today. However, the use of newspaper advertisements has its own specific challenges. Unlike land registers, newspaper listings do not refer to the full universe of transactions and there is a question about the extent to which the sample is representative. In addition, listings often do not provide prices at all, and if they do then these are offer prices which may deviate from what was finally paid. The spread is probably smaller in the case of rents than sales, and in historical work advertisements have so far mostly been used in the calculation of rent indices (Margo (1996), Kholodilin et al (2021)). However, offer prices may be a sufficient source if care is taken, rates of change rather than levels are focused on and hedonic models are used, as argued by Keely and Lyons (2020) in their study on post-1945 house prices in Dublin.

Media do not only publish advertisements. Due to the publicity of land registers and notary records in many countries, media have access to, and indeed report, information on actual transactions as well. Such publications are typically addressed to wealthy investors and professionals in the real estate business, with a corresponding focus on large residential or commercial structures. If used systematically, such secondary sources can significantly reduce the work otherwise involved in collecting individual transaction data. As readers of these media were interested in additional information – eg the size of the lot, year of build, the number of stories, building materials or building designation – reports might contain many of the variables necessary for hedonic regressions. An example of such a source is the French *l'Economiste français*, founded in 1873, which published a weekly list of buildings sold in Paris including the prices paid. Similar listings in the *Real estate record and builder's guide* have been used by Nicholas and Scherbina (2013) in their study on the New York housing market during the 1920s and the Great Depression. Another example of a private collection of publicly available information, although in this case for private purposes and not publication in a newspaper, is the archive of the Amsterdam-based firm Jan Brouwer & Zn used by Eichholtz et al (2021) for the calculation of a repeat-sale index for Amsterdam. According to Eichholtz et al (2021), the existence and structure of the source is due to Amsterdam's history of selling property in public auctions, not only for foreclosure but also in regular sales for investment purposes. Before auctions, information on rents and taxes were published, and the hammer price could be recorded by interested investors. In the case of Brouwer, the information was stored in a card system that allows researchers to easily match and track sales and rent transactions for individual properties over time.

Records of real estate firms and property lenders may also be useful sources of price data and appraisals. As Holmans (2005) discusses in the case of Britain, the house price data for much of the period since the Second World War are derived from information collected by building societies and other house purchase lenders about the dwellings that their mortgages financed. But it is necessary to take care if the information comes from the business of a single lender (or a narrow set of lenders). That means shifts in the market share and the mix of business of those lenders will affect how representative its data are of property prices more generally. In the 1960s

chronologically ordered protocols using pointers from the individual property ledger. Söderberg et al (2014) collected taxation values from one protocol, which was ordered spatially (by geographical address), and combined this with data on sales prices which was collected from a different protocol, which was ordered chronologically.

and 1970s, for example, the prices of houses purchased with loans from building societies could represent, with a high degree of certainty, the prices of all dwellings bought with loans, but from the 1980s onwards they no longer did so because banks increasingly entered the mortgage market. Information about the prices of houses from lenders obviously only reflects the price of dwellings purchased with a loan. This will not cover houses of poorer quality which tend to be at the bottom end of the market, and were not acceptable as security for loans.

Records of institutional endowments provide another potential source for information on the property market, especially if those institutions are big and long-lived. Chambers et al (2021) look at the real estate portfolio of colleges in Oxford and Cambridge for the period 1901–83. By definition, long-term investors rent out rather than sell and buy, but their sample still includes a significant number of sales and purchases. Hospitals, the church and monasteries are classical sources for medieval and early modern prices and wages, but also provide information on rents, as used by Drehlichman and González Aguado (2014) for 15th-century Toledo. Clark (2002) uses a large sample of dwellings owned by charities in England and Wales from the 16th century onwards, in which information comes from 19th century investigations into the activities of charities.

A final example of a source on house prices are surveys on household wealth, which are rather a special case. Wealth surveys ask for the current (estimated) market value of the property owned by the household but sometimes also for the original price and year of purchase. In this case a (pseudo) repeat-sales index can be calculated. Such a survey is the basis for the index for US house prices by Grebler et al (1956), which served as input to the Shiller (2015) house price index for the US.

While records of individual transactions provide the preferred primary material for analysing past developments in housing prices, additional sources exist which can supplement transaction prices or provide a substitute if no index based on transaction prices can be compiled or has yet been compiled. For Canada, for instance, house prices for 1921–49 are proxied by an average replacement value, composed of estimated construction costs and land value. As argued above, land prices and estimates of total construction costs are of interest in themselves, as they can serve as an indicator of the fundamental value of real estate. With the focus on quality adjusted indices, these sources are not further discussed here.

#### 4.4 Conclusions and lessons going forward

As this survey has made evident, recent years have seen a boom in the construction of historical housing price indices. The boom is a result of demand for academic and policymaking purposes, and of methodological advances which have led to the broad adoption of SPAR or hedonic models as reference methods for constructing indices. These are particularly useful when devising long- and very long-run price series, as the quality and composition of the housing stock will change dramatically from a secular perspective. The boom is also unlikely to be over soon. If anything, the breadth of sources used in recent publications has shown that there is much primary material out there and that new research is worthwhile. Besides land registers and notarial archives, promising sources seem to include annuaries and address books that list (assessed) property values along with key characteristics of the property, and which are available for many European cities starting in the 19th century. Another avenue is the specialised press, as well as other contemporary compilations undertaken by

mortgage institutions, banks or investors. The archive of the Amsterdam-based firm Jan Brouwer & Zn is an example.<sup>89</sup>

At the same time, the availability of many new indices enables an empirical comparison of different methods for quality adjustment, and assessment of their relative merits in historical contexts. In many cases, the structure of sources and lack of sufficient explanatory variables for hedonic regressions has made repeat-sales methods the obvious choice. In order to be tractable, repeat-sales indices often rely on small numbers of observations, in the order of 20 to 40 per year, and in the case of the Herengracht on just 5.2 on average for the 17th century. If repeat-sales indices are constructed for long time spans, they typically also cover small areas, such as the 10 streets in Dublin examined by Deeter et al (2017). Hedonic methods allow the potential sample to expand as observations need not be matched, but they require more and different data. For researchers facing these difficult trade-offs it would be helpful to get a better understanding how different types of index fare in practice. In the same vein, new quality-adjusted indices can be used to benchmark older existing crude average or median prices to see whether under specific circumstances these can provide a reasonable alternative to the time- and resource-intensive calculation of quality-adjusted series.<sup>90</sup> Another methodological question that could potentially unlock significant data sources is whether and, if so, how appraisal values can be used to track developments in housing prices. The address books used by Bohlin (2014) and Söderberg (2014) give information on all houses in Stockholm and Gothenburg in tabular form. Using OCR and machine learning, these tables could be converted into a panel data set not of a sample but of the entire housing stock of the two cities. We are thus very far from the small sample problem faced by most repeat-sales indices. If appraisal values were adjusted regularly to follow market prices, then these sources, which are available for many cities, could in fact provide significant information, if not on annual fluctuations in market prices, then at least on long-term trends.<sup>91</sup>

Finally, this chapter has discussed commercial property and land prices only in passing. Rents have not been discussed either, and Chambers et al (2021) and Eichholtz et al (2021) have recently argued that the matching of transaction prices and rents for individual properties yields significantly lower total returns compared with studies that combine aggregate price and rent indices (Jordà et al (2019)). This might imply that to reap economies of scope, data collection should include both prices and rents, if primary sources permit. Similarly, information on individual mortgages, which are also available from land registries and notaries, could be

<sup>89</sup> In the best of cases, some historians have already undertaken the collection and editing of the relevant information, without, however, constructing an index. The Herengracht index, for instance, is based on a social, cultural and art history of all buildings in the Herengracht, published at the occasion of the 750th anniversary of the city of Amsterdam in 1975 (Eeghen et al (1976)).

<sup>90</sup> Edvinsson et al (2021) calculate both a SPAR and a hedonic index for Stockholm. Both methods rely, critically, on the assumption of constant-quality since they are tracing the same individual house over time. But this assumption may be difficult to verify, except that the geographical location remains the same. For instance, it is only in cases in which the value depreciation from age and the net additions to value from renovation and modernisation exactly cancel each other out that the constant-quality assumption may actually hold in practice.

<sup>91</sup> de Vries et al (2009) discuss the use of appraisal data sets in house price measurement in their study of HPIs for the Netherlands for the period from January 1995 to March 2009, comparing the SPAR method with the repeat sale method. Further details and statistical analysis of SPAR HPIs is provided in de Haan et al (2008).

incorporated to arrive at a better micro-based understanding of housing credit and its dynamics.

To conclude, as we look ahead, the wealth of primary data waiting to be exploited – combined with thoughtful and historically informed choices of method – promise significant advances in our understanding of historical property markets and their interaction with the financial system.

## References

Alfani, G (2015): “Economic inequality in northwestern Italy: a long-term view”, *Journal of Economic History*, vol 75, no 4, pp 1058–96.

Bailey, M, R Muth and H Nourse (1963): “A regression method for real estate price index construction”, *Journal of the American Statistical Association*, vol 58, no 304, pp 933–42.

Bank for International Settlements (BIS) (1990): *60th Annual Report*, June.

——— (2005): “Real estate indices and financial stability”, *BIS Papers*, no 21, proceedings from a conference jointly organised by the BIS and the IMF in Washington DC, 27–28 October 2003.

Bavel, B van (2016): *The invisible hand? How market economies have emerged and declined since AD 500*, Oxford University Press.

Bellhouse, D (2017): *Leases for lives: life contingent contracts and the emergence of actuarial science in 18th century Britain*, Cambridge University Press.

Blackwell, T and S Kohl (2019): “Historicizing housing typologies: beyond welfare state regimes and varieties of residential capitalism”, *Housing Studies*, vol 34, no 2, pp 298–318.

Bochove, C, H Deneweth and J Zijderduijn (2015): “Real estate and mortgage finance in England and the Low Countries, 1300–1800”, *Continuity and Change*, vol 30, no 1, pp 9–38.

Bohlin, J (2014): “A price index for residential property in Göteborg, 1875–2010”, in R Edvinsson, T Jacobson and D Waldenström (eds), *House prices, stock returns, national accounts, and the Riksbank balance sheet, 1620–2012*, Stockholm, Sveriges Riksbank and Ekerlids Förlag, pp 26–62.

Bourassa, S, M Hoesli and J Sun (2006): “A simple alternative house price index method”, *Journal of Housing Economics*, vol 15, no 1, pp 80–97.

Braae, G (1960): *Fluctuations in investment in Britain and America in the inter-war years 1919–39*, Oxford, Bodleian Library, unpublished thesis.

Carmona, J, M Lampe and J Rosés (2014): “Spanish housing markets, 1904–1934: new evidence”, *Revista de Historia Económica/Journal of Iberian and Latin American Economic History*, vol 32, no 1, pp 119–50.

Case, K and R Shiller (1987): “Prices of single-family homes since 1970: new indexes for four cities”, *New England Economic Review*, September, pp 45–56.

——— (1989): “The efficiency of the market for single-family homes”, *American Economic Review*, vol 79, no 1, March, pp 125–37.

- Cerman, M (2008): "Social structure and land markets in late medieval central and east-central Europe", *Continuity and Change*, vol 23, no 1, pp 55–100.
- Chambers, D, C Spaenjers and E Steiner (2021): "The rate of return on real estate: long-run micro-level evidence", *The Review of Financial Studies*, vol 34, no 8, pp 3572–607.
- Chinloy, P, M Cho and I Megbolugbe (1997): "Appraisals, transaction incentives, and smoothing", *Journal of Real Estate Finance and Economics*, vol 14, pp 89–111.
- Clapham, E, P Englund, J Quigley and C Redfearn (2006): "Revisiting the past and settling the score: index revision for house price derivatives", *Real Estate Economics*, vol 34, no 2, pp 275–302.
- Clapp, J and C Giaccotto (1998): "Price indices based on the hedonic repeat-sales method: application to the housing market", *Journal of Real Estate Finance and Economics*, vol 16, no 1, pp 5–26.
- Clark, G (2002): "Shelter from the storm: housing and the industrial revolution, 1550–1909", *The Journal of Economic History*, vol 62, no 2, pp 489–511.
- Clark, G and A Clark (2001): "Common rights to land in England, 1475–1839", *Journal of Economic History*, vol 61, no 4, pp 1009–36.
- Collins, W and R Margo (2007): "The economic aftermath of the 1960s riots in American cities: evidence from property values", *The Journal of Economic History*, vol 67, no 4, pp 849–83.
- Deeter, K, F Quinn, and D Duffy (2017): "Dublin house prices: a history of booms and busts from 1708–1949", *Journal of the Statistical and Social Inquiry Society of Ireland*, vol 46, pp 1–17.
- de Haan, J, E van der Wal and P de Vries (2008): *The measurement of house prices: a review of the sale price appraisal ratio method*, [www.cbs.nl](http://www.cbs.nl).
- de Vries, P, J de Haan, E van der Wal and G Mariën (2009): "A house price index based on the SPAR method", *Journal of Housing Economics*, vol 18, no 3, pp 214–23.
- Drelichman, M and D González Agudo (2014): "Housing and the cost of living in early modern Toledo", *Explorations in Economic History*, vol 54, October, pp 27–47.
- Duon, G (1943): "Évolution de la valeur vénale des immeubles parisiens", *Journal de la société statistique de Paris*, vol 84, pp 169–92.
- Econ (2004): "Justeringer i Eiendomsmeglerbransjens boligprisstatistikk" [Adjustments in house price statistics from the Real Estate Agents' Association] (in Norwegian), *Econ Analyse*, July.
- Edvinsson, R, K Eriksson and G Ingman (2021): "A real estate price index for Stockholm, Sweden 1818–2018: putting the last decades housing price boom in a historical perspective", *Scandinavian Economic History Review*, vol 69, no 1, pp 83–101.
- Eeghen, I, G Roosegaarde Bisshop and H Wijnen (1976): *Vier eeuwen Herengracht*. Amsterdam, Stadsdurkkerij Amsterdam.
- Eichholtz, P (1997): "A long run house price index: the Herengracht index, 1628–1973", *Real Estate Economics*, vol 25, no 2, pp 175–92.
- Eichholtz, P, M Korevaar, T Lindenthal and R Tallec (2021): "The total return and risk to residential real estate", *The Review of Financial Studies*, vol 34, no 8, pp 3608–46.

- Eitrheim, Ø (2022): "Revisions and break-adjustments in composite house price indices", in Ø Eitrheim, J Klovland and J Qvigstad (eds), *Historical Monetary and Financial Statistics for Norway*, Occasional Papers, Norges Bank, forthcoming.
- Eitrheim, Ø and S Erlandsen (2004): "House price indices for Norway 1819–2003", in Ø Eitrheim, J Klovland and J Qvigstad (eds), *Historical Monetary Statistics for Norway, 1819–2003*, Norges Bank Occasional Papers, no 35, Chapter 9 .
- (2005): "House price indices for Norway 1819–1989", *Scandinavian Economic History Review*, vol 53, no 3.
- Englund, P, T Gordon and J Quigley (1999a): "The valuation of real capital: a random walk down Kungsgatan", *Journal of Housing Economics*, vol 8, no 3, pp 205–16.
- Englund, P, J Quigley and C Redfearn (1998): "Improved price indexes for real estate: measuring the course of Swedish housing prices", *Journal of Urban Economics*, vol 44, no 2, pp 171–96.
- (1999b): "The choice of methodology for computing housing price indexes: comparisons of temporal aggregation and sample definition", *Journal of Real Estate Finance and Economics*, vol 19, no 2, pp 91–112.
- Eurostat (2013): *Handbook on residential property price indices (RPPIs)*, technical report, Eurostat.
- (2017): *Technical manual on owner-occupied housing and house price indices*, Eurostat.
- Faroqhi, S (1987): *Men of modest substance: house owners and house property in seventeenth-century Ankara and Kayseri*, Cambridge University Press.
- Francke, M (2010): "Repeat sales index for thin markets: a structural time series approach", *Journal of Real Estate Finance and Economics*, vol 41, no 1, pp 24–52.
- Friggit, J (2008): "Comparing four secular home price indices", *Working Paper Version 9*, Conseil Général des Ponts et Chaussées, Ministry of Ecology, Sustainable Planning and Development.
- Gatzlaff, D and D Ling (1994): "Measuring changes in local house prices: an empirical investigation of alternative methodologies", *Journal of Urban Economics*, vol 35, no 2, pp 221–44.
- Goetzmann, W (2016): *Money changes everything: how finance made civilization possible*, Princeton University Press.
- Goetzmann, W and L Peng (2002): "The bias of the RSR estimator and the accuracy of some alternatives", *Real Estate Economics*, vol 30, no 1, pp 13–39.
- Goodman, A and T Thibodeau (1998): "Dwelling age heteroscedasticity in repeat sales house price equations", *Real Estate Economics*, vol 26, no 1, pp 151–71.
- Grebler, L, D Blank and L Winnick (1956): *Capital formation in residential real estate: trends and prospects*, NBER, Princeton University Press.
- Harding, J, S Rosenthal and C Sirmans (2007): "Depreciation of housing capital, maintenance, and house price inflation: estimates from a repeat sales model", *Journal of Urban Economics*, vol 61, no 2, pp 193–217.
- Herlihy, D and C Klapisch-Zuber (1978): *Les Toscans et leurs familles: une étude du catasto florentin de 1427*, Paris, Presses de la Fondation Nationale des Sciences.



- Hill, R (2013): "Hedonic price indexes for residential housing: a survey, evaluation and taxonomy", *Journal of Economic Surveys*, vol 27, no 5, pp 879–914.
- Hoffman, P, G Postel-Vinay and J-L Rosenthal (2000): *Priceless markets, the political economy of credit in Paris*, University of Chicago Press.
- Holmans, A (2005): *Abstract of historical statistics of British housing*, Cambridge Centre for Housing and Planning Research, [www.cchpr.landecon.cam.ac.uk/system/files/documents/Report\\_45.pdf](http://www.cchpr.landecon.cam.ac.uk/system/files/documents/Report_45.pdf).
- Hwang, M and J Quigley (2004): "Selectivity, quality adjustment and mean reversion in the measurement of house values", *Journal of Real Estate Finance and Economics*, vol 28, no 2/3, pp 161–78.
- Ito, S (2013): "Registration and credit in seventeenth-century England", *Financial History Review*, vol 20, no 2, pp 137–62.
- Jordà, Ò, K Knoll, D Kuvshinov, M Schularick and A Taylor (2019): "The rate of return on everything, 1870–2015", *The Quarterly Journal of Economics*, vol 134, no 3, pp 1225–98.
- Jordà, Ò, M Schularick and A Taylor (2016): Jordà-Schularick-Taylor Macrohistory Database, technical report, [www.macrohistory.net/data](http://www.macrohistory.net/data). JSTDatasetR2, second edition, November.
- (2017): "Macrofinancial history and the new business cycle facts", in M Eichenbaum and J Parker (eds), *NBER Macroeconomics Annual 2016*, vol 31, no 1, University of Chicago Press, pp 213–63.
- Jursa, M (2014): "Factor markets in Babylonia from the late seventh to the third century BCE", *Journal of the Economic and Social History of the Orient*, vol 57, no 2, pp 173–202.
- Keely, R and R Lyons (2020): "Housing prices, yields and credit conditions in Dublin since 1945", *Journal of Real Estate Finance and Economics*, pp 1–36.
- Kholodilin, K, L Limonov and S Watzl (2021): "Housing rent dynamics and rent regulation in St Petersburg (1880–1917)", *Explorations in Economic History*, vol 81, July.
- Knoll, K, M Schularick and T Steger (2017): "No price like home: global house prices, 1870–2012", *American Economic Review*, vol 107, no 2, pp 331–53.
- Lescure, M (1980): *Les sociétés immobilières en France au XIXe siècle: contribution à l'histoire de la mise en valeur du sol urbain en économie capitaliste*, Paris, Publications de la Sorbonne.
- (1982): *Les banques, l'état et le marché immobilier en France à l'époque contemporaine 1820–1940*, Paris, Editions de l'Ecole des Hautes Etudes en Sciences Sociales.
- Mack, A, E Martínez-García and V Grossman (2011): "A cross-country quarterly database of real house prices: a methodological note", *Working Paper*, no 99, Federal Reserve Bank of Dallas, Globalization and Monetary Policy Institute, revised January 2019.
- Margo, R (1996): "The rental price of housing in New York City, 1830–1860", *Journal of Economic History*, vol 56, no 3, pp 605–25.

- Matsubara, K (2017): "The Meiji land reform and the formation of modern land rights in Japan", in S-W Cheung (ed), *Colonial Administration and Land Reform in East Asia*, London, Routledge, pp 37–48.
- Meen, G (2016): "Spatial housing economics: a survey", *Urban Studies*, vol 53, no 10, pp 1987–2003.
- Meese, R and N Wallace (1997): "The construction of residential housing price indices: a comparison of repeat-sales, hedonic-regression, and hybrid approaches", *Journal of Real Estate Finance and Economics*, vol 14, no 1, pp 51–73.
- National Statistician (2010): *National Statistician's review of house price statistics*, Kew: Crown Copyright.
- Nicholas, T, and A Scherbina (2013): "Real estate prices during the roaring twenties and the Great Depression", *Real Estate Economics*, vol 41, no 2, pp 278–309.
- Office for National Statistics (2013): *Official house price statistics explained*, technical report, 4 April 2013.
- Paddison, R (2001): *Handbook of urban studies*, London, Sage Publishing.
- Pennington-Cross, A (2005): "Aggregation bias and the repeat sales price index", *BIS Papers*, no 21, proceedings of a joint conference organised by the BIS and the IMF in Washington DC, 27–28 October 2003.
- Poisson, J-P (1985): *Notaires et société; travaux d'histoire et de sociologie notariale*, Paris, Economica.
- Pomeranz, K (2000): *The great divergence: China, Europe, and the making of the modern world economy*, Princeton University Press.
- Raff, D, S Wachter and S Yan (2013): "Real estate prices in Beijing, 1644 to 1840", *Explorations in Economic History*, vol 50, no 3, pp 368–86.
- Rosen, S (1974): "Hedonic prices and implicit markets", *Journal of Political Economy*, vol 82, no 1, pp 34–55.
- Saito, O (2009): "Land, labour and market forces in Tokugawa Japan", *Continuity and Change*, vol 24, no 1, pp 169–96.
- Samy, L (2015): "Indices of house prices and rent prices of residential property in London, 1895–1939", *Discussion Paper in Economic and Social History 134*, Winton Institute for Monetary History, University of Oxford.
- Scatigna, M and R Szemere (2014): BIS collection and publication of residential property prices, [www.bis.org/ifc/events/7ifcconf\\_scatigna\\_szemere.pdf](http://www.bis.org/ifc/events/7ifcconf_scatigna_szemere.pdf).
- Shiller, R J (1993): *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks*, Clarendon Lectures in Economics, Oxford University Press.
- (2015): *Irrational exuberance*, Princeton University Press, third edition.
- Shimizu, C and K Nishimura (2006): "Biases in appraisal land price information: the case of Japan", *Journal of Property Investment & Finance*, vol 24, no 2, pp 150–75.
- (2007): "Pricing structure in Tokyo metropolitan land markets and its structural changes: pre-bubble, bubble, and post-bubble periods", *The Journal of Real Estate Finance and Economics*, vol 35, no 4, pp 475–96.

- Shimizu, C, K Nishimura and T Watanabe (2011): "House prices at different stages of the buying/selling process", *Working Paper Series*, no 69, Research Center for Price Dynamics, Institute of Economic Research, Hitotsubashi University.
- (2012): "House prices from magazines, realtors, and the Land Registry", *BIS Papers*, no 64.
- Silver, M (2011): "House price indices: does measurement matter?", *World Economics*, vol 12, no 3, pp 69–85.
- (2014): "The degree and impact of differences in house price index measurement", *Journal of Economic and Social Measurement*, vol 39, no 4, pp 305–28.
- Silverstein, J (2014): "House price indexes: methodology and revisions", *Research Special Report*, Federal Reserve Bank of Philadelphia, Research Department, June.
- Söderberg, J, S Blöndal and R Edvinsson (2014): "A price index for residential property in Stockholm, 1875–2011", in R Edvinsson, T Jacobson and D Waldenström (eds), *House prices, stock returns, national accounts, and the Riksbank balance sheet, 1620-2012*, Sveriges Riksbank and Ekerlids Förlag, pp 63–100.
- Temin, P (2001): "A market economy in the early Roman Empire", *The Journal of Roman Studies*, vol 91, pp 169–81.
- Van Zanden, J (2009): *The long road to the industrial revolution: the European economy in a global perspective, 1000–1800*, vol 1, Leiden, Brill.
- Wang, F and P Zorn (1997): "Estimating house price growth with repeat sales data: what's the aim of the game?", *Journal of Housing Economics*, vol 6, no 2, pp 93–118.
- Willis, J (1950): "Short history of rent control laws", *Cornell Law Review*, vol 36, no 1, pp 54–94.

## Previous volumes in this series

<b>No</b>	<b>Title</b>	<b>Issue date</b>
BIS Papers No 126	Corporate digital identity: no silver bullet, but a silver lining	June 2022
BIS Papers No 125	Gaining momentum – Results of the 2021 BIS survey on central bank digital currencies	May 2022
BIS Papers No 124	The design of a data governance system	May 2022
BIS Papers No 123	CBDCs in emerging market economies	April 2022
BIS Papers No 122	The monetary-fiscal policy nexus in the wake of the pandemic	March 2022
BIS Papers No 121	Covid-19 and the monetary-fiscal policy nexus in Africa	February 2022
BIS Papers No 120	Virtual banking and beyond	January 2022
BIS Papers No 119	Non-bank financial institutions and the functioning of government bond markets	November 2021
BIS Papers No 118	A taxonomy of sustainable finance taxonomies	October 2021
BIS Papers No 117	Fintech and the digital transformation of financial services: implications for market structure and public policy	July 2021
BIS Papers No 116	CBDCs beyond borders: results from a survey of central banks	June 2021
BIS Papers No 115	Multi-CBDC arrangements and the future of cross-border payments	March 2021
BIS Papers No 114	Ready, steady, go? – Results of the third BIS survey on central bank digital currency	January 2021
BIS Papers No 113	Financial market development, monetary policy and financial stability in emerging market economies	December 2020
BIS Papers No 112	The dawn of fintech in Latin America: landscape, prospects and challenges	November 2020
BIS Papers No 111	Inflation dynamics in Asia and the Pacific	March 2020
BIS Papers No 110	Measuring the effectiveness of macroprudential policies using supervisory bank-level data	February 2020
BIS Papers No 109	The digital economy and financial innovation	February 2020

All volumes are available on the BIS website ([www.bis.org](http://www.bis.org)).

*Promoting global monetary  
and financial stability*

Bank for International Settlements (BIS)

[www.bis.org](http://www.bis.org)  
[email@bis.org](mailto:email@bis.org)