CHAPTER 1
QUANTITATIVE HISTORY AND UNCHARTED PEOPLE

*Johan Fourie**

No archive, and therefore no history, is ever complete. Records of the powerful and privileged are preserved, while documents reflecting the lives of ordinary men and women, particularly the illiterate, the poor and the subjugated, are often missing, lost through dereliction, disaster or decree, or never saved at all. Such omissions in our recorded histories have consequences. Records inform our current identities and ideologies and shape our future. By distorting the past – and our place in it – an incomplete history can be a misleading guide for the future.

Historians have found creative ways to overcome the biases of the conventional archive. Sometimes they use oral testimonies to view history 'from below', the stories of how ordinary men and women shaped and were shaped by history. This can reveal the hidden histories of minority or marginalised groups. But such sources have their pitfalls.

This chapter, and indeed the whole book, makes the case for a different way to remedy the omissions of the archive. The method we use is quantitative history, the application of statistical data to the study of history. New quantitative tools and techniques now enable scholars to analyse information from administrative sources at a fraction of the cost and effort of a few decades ago. While administrative records are often of little interest individually, when combined with hundreds or sometimes even hundreds of thousands of similar records, they become a powerful means to test existing hypotheses and find new avenues to explore.

Quantitative history, I argue, can be especially powerful in South Africa, where large quantities of historical records are preserved but questions of bias and prejudice persist. With statistical tools we can extract information from historical documents and make it tell us far more than could have been imagined by those who collected and preserved it. This allows us to mitigate or condition for the bias. In some cases, it might even allow us to measure the extent of bias.

In this book's eleven research chapters, written by scholars associated with the *Biography of an Uncharted People* project at Stellenbosch University, we hope to

*Fourie: Department of Economics, Stellenbosch University, South Africa. I am grateful to Keith Breckenridge, Michiel de Haas, Jan Kok, Robert Ross and Charles van Onselen for their extensive comments on earlier versions of this chapter.

illuminate the ways in which the new tools and techniques can help us question what are taken to be historical truths and reveal new insights. The book is aimed at broadly two audiences: first, a younger generation of historians – from senior undergraduate students to PhDs and postdocs – interested in the opportunities (and pitfalls) that quantitative sources and tools offer. It is for that reason that we have created slides and tutorials for each of the chapters in the book.[1] A second audience is established academic historians (of South Africa) who are interested in finding new digital tools that can help to uncover sources perhaps previously ignored, new research questions to ask from their existing source material, or new techniques for writing histories from below.

We hope not to suffer from the overconfidence of an earlier generation of quantitative historians: we do not say that quantification should displace other methods of historical inquiry. It is, at best, a complementary approach to writing good history. There is much that quantitative sources and tools cannot do. Yet as long as archives have gaps, as long as some people are uncharted, the quantitative approach will be valuable in the quest for completeness of the historical record. And for shaping our identities, ideologies and future.

The rise of quantitative history

We are, of course, not the first to make use of quantitative techniques to study the past, nor the first to make the case for it. With its roots in the *Annales* school, a twentieth-century movement of predominantly French historians, quantitative history flourished from around 1950 to 1980. Peter Burke identifies three phases in the evolution of its interests: first, economic history (investigating prices in France, with the economist Francois Simiand and the historian Ernest Labrousse being the two main proponents), then social history, notably population studies and demography, and finally cultural history.[2]

It was its adoption by economists in the United States that propelled quantitative history forward. They made three telling contributions: the use of quantitative tools, and an enlargement of scale and scope. Although economic historians had long used statistics as supporting evidence, what was new, as explained by Douglas North, who in 1993 shared the Nobel Prize in Economics with Robert Fogel, was the use of quantitative tools to test hypotheses.[3] Rather than just calculating annual price fluctuations, as Simiand, Labrousse and many others had done, Harvard economists Conrad and Meyer used prices – slave prices, to be precise – applied to economic theory to make a causal claim, with counterfactual implications: that US slavery was more profitable than previously

[1]See the final chapter for more information.

[2]P. Burke, *The French Historical Revolution: The Annales School, 1929–2014* (Stanford: Stanford University Press, 2015): 60.

[3]D.C. North, 'The New Economic History after Twenty Years', *American Behavioural Scientist* 21, no. 2 (1977): 187–200.

thought, and that it was unlikely to have ended without the Civil War.[4] The objective was to use economic theory to find a general *causal* interpretation of historical events.[5]

Now known as 'cliometricians', the group increased the scale of their analysis.[6] They began to use information available at the individual (farm or household) level, rather than just regional or national averages. This is best exemplified by Fogel and Engerman's *Time on the Cross*, which relied on Parker and Gallman's 1860 Census of Agriculture sample of cotton-belt farms that was made accessible by being converted into machine-readable historical micro-data sets.[7]

Extending their scope, they made telling contributions to fields beyond economic history – even to the domain of policy. Reflecting on Fogel's *Railroads and American Economic Growth*,[8] another seminal contribution, economic historian Barry Eichengreen said that Fogel had argued that technical innovations that were thought indispensable, like the railroad, were not necessarily so, and generalised his findings to a rethinking of US economic policy and even the US space program.[9] The controversial findings of *Time on the Cross* caused a stir both within and far beyond the fields of economics and history.

Although it was the cliometricians who most enthusiastically adopted a more rigorous quantitative approach, quantification was not limited to the field of economic history. In fact, what defined quantitative history was its interdisciplinary nature. The need for history to move beyond the confines of mere narrative had already been apparent much earlier. In 1912 the American historian James Harvey Robinson extolled the virtues of interdisciplinary research, arguing that this 'New History', as he called it, 'will avail itself of all those discoveries that are being made about mankind by anthropologists, economists, psychologists and sociologists'.[10] But it was only in the 1950s, after Fernand Braudel's insistence that history and geography should march in lockstep over the *longue durée*, and in particular the 1960s and 1970s, that an enthusiasm for quantification with explanatory power took off.[11] One example of this was the launch in 1970 of the *Journal*

---

[4]A.H. Conrad, and J.R. Meyer, 'The Economics of Slavery in the Ante Bellum South', *Journal of Political Economy* 66, no. 2 (1958): 95–130.

[5]J.R. Meyer and A.H. Conrad, 'Economic Theory, Statistical Inference, and Economic History', *The Journal of Economic History* 17, no. 4 (1957): 524–44.

[6]'Cliometrics' combines 'Clio', the muse of history, and 'metrics', as in 'econometrics'. See D.N. McCloskey, 'The Achievements of the Cliometric School', *The Journal of Economic History* 38, no. 1 (1978): 13–28; C. Goldin, 'Cliometrics and the Nobel', *Journal of Economic Perspectives* 9, no. 2 (1995): 191–208. For a more recent overview, see C. Diebolt and M. Haupert, 'Cliometrics: Past, Present, and Future', in *Oxford Research Encyclopedia of Economics and Finance* (Oxford: Oxford University Press, 2021): 19.

[7]R. Fogel and S. Engerman, *Time on the Cross: The Economics of American Negro Slavery* (Boston: Little, Brown and Co., 1974).

[8]R.W. Fogel, *Railroads and American Economic Growth* (Baltimore: Johns Hopkins Press, 1964).

[9]B. Eichengreen, 'The Contributions of Robert W. Fogel to Economics and Economic History', *The Scandinavian Journal of Economics* 96, no. 2 (1994): 167–79.

[10]J.H. Robinson, *The New History: Essays Illustrating the Modern Historical Outlook* (New York: The Macmillan Company, 1912): 24.

[11]F. Braudel, 'Histoire et sciences sociales: la longue durée', *Annales. Histoire, Sciences Sociales* 13, no. 4 (1958): 725–53.

of Interdisciplinary History. In its first issue, the editors announced that: 'Whole new fields, such as historical demography, and entirely new techniques, such as computer data processing, have appeared and have made a broad impact on many areas of research'. In a review of five decades of the journal, Ruggles and Magnuson note that 'quantification was the central defining element of the "new" histories – the new social history, new economic history, and new political theory – that transformed the landscape of historical research in the 1960s and 1970s'.[12]

Social historians began to use probate inventories and marriage and death records and often employed new methods from the social sciences. The Cambridge Group for the History of Population and Social Structure, for example, reconstructed families from vital records to show that English households were more nuclear, more geographically mobile and more likely to marry later.[13] Others used soldiers' heights reported on military attestation forms to measure living standards across time – a field that became known as anthropometric history – or used the signatures on documents (or lack thereof) as a way to measure literacy in the past.[14] Political historians in the United States, unhappy with the lack of empirical rigour in political theory, began to analyse voting patterns, correlating voting outcomes with characteristics like class or religion.[15] By the 1970s, the need to organise these disparate disciplines under one roof resulted in the establishment of the Social Science History Association[16] whose explicit aim was to bring together social scientists who were working with historical materials. Its journal, *Social Science History*, was first published in 1976.

Interdisciplinary work also meant collaboration. Articles or books were generally co-authored, often by three or more authors. This was in contrast to the single-authored publications that were common in the field of history. Research projects also grew in size and scope. Teams of research assistants were often required to process the large quantities of information, initially by hand but later, as computers became more accessible, on punch cards. One welcome advantage of using the methods of the hard sciences was that historians could now apply for grants not usually accessible to the humanities.[17]

---

[12] S. Ruggles and D.L. Magnuson, 'The History of Quantification in History: The JIH as a Case Study', *Journal of Interdisciplinary History* 50, no. 3 (2019): 363–81.

[13] P. Laslett, 'Size and Structure of the Household in England Over Three Centuries', *Population studies* 23, no. 2 (1969): 199–223; P. Laslett, 'Characteristics of the Western Family Considered Over Time', *Journal of Family History* 2, no. 2 (1977): 89–115.

[14] R. A. Margo, and R.H. Steckel, 'Heights of Native-born Whites During the Antebellum Period', *The Journal of Economic History* 43, no. 1 (1983): 167–74; R.W. Beales, 'Studying Literacy at the Community Level: A Research Note', *The Journal of Interdisciplinary History* 9, no. 1 (1978): 93–102.

[15] L. Benson, *The Concept of Jacksonian Democracy: New York as a Test Case* (Princeton: Princeton University Press, 1961); A.G. Bogue, 'The Quest for Numeracy: Data and Methods in American Political History', *Journal of Interdisciplinary History* 21 (1990): 89–116.

[16] There seems to be some disagreement about the date: the American Historical Association reports the founding of the SSHA as 1972, the SSHA's own website reports 1974 and Wikipedia reports 1976.

[17] C. Lemercier and C. Zalc, *Quantitative Methods in the Humanities: An Introduction* (Charlottesville: University of Virginia Press, 2019).

Yet despite all this excitement, by the early 1980s, quantitative history was on the decline. There were good reasons for this. Historians, who had dabbled in quantitative history during these early years became emboldened by the attention from economists whose own profession was becoming increasingly mathematical, often used their sources uncritically to make bold assertions about the universality or generalizability of their findings. The particular and the idiosyncratic were lost in their sometimes anachronistic classifications. Quantification became a synonym for the positivist approach – the antithesis of the postmodern, relativist approach that was becoming popular just as the merits of quantitative history were being questioned.

And there were practical concerns too: quantification required large investments of time and resources, yet the returns were not immediately obvious. Many of the early proponents of quantification became increasingly sceptical and began to ask whether it was all worth it.[18] What questions could quantification solve that other approaches could not? For some, using sophisticated techniques took precedence over asking a relevant research question. As historian Lawrence Stone put it, they 'used the most sophisticated techniques either to prove the obvious or to claim to prove the implausible … The results sometimes combine the vices of unreadability and triviality'.[19] Historians Claire Lemercier and Claire Zalc put it plainly: 'Quantitative historians were often boring, and had themselves become bored.'[20]

For the next two decades at least, quantitative history took a back seat. It gave way to other approaches, from Marxism to postmodernism. E. P. Thompson's *The Making of the English Working Class*, published in 1963, stimulated new ways of viewing the past.[21] Gender history is one example of new subfields that flourished.[22] So, too, did earlier work by French philosophers like Michel Foucault and Jacques Derrida shift the historian's gaze. By the 1980s, objective truth was questionable and power relations explained all.

One thing these diverse schools had in common, however, was the desire to write 'history from below' or 'history from the bottom up'; that is, to focus explicitly on those who have not been accounted for, the marginalised, the inarticulate – whom we have called the 'uncharted' people.

The decline of quantitative history ended abruptly by the early 2000s. There are both demand- and supply-side reasons for this turnaround, or 'renaissance'.[23] On the demand side, the inability of postmodernism to explain (or even engage with) material progress – the rapid improvement in global living standards since the Industrial Revolution and

---

[18] L. Benson, 'The Mistransference Fallacy in Explanations of Human Behavior', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 17, no. 3 (1984): 118–31.

[19] L. Stone, 'The Revival of Narrative: Reflections on a New Old History', *Past & Present* 85 (1979): 11.

[20] Lemercier and Zalc, *Quantitative Methods,* 15.

[21] E.P. Thompson, *The Making of the English Working Class* (London: Gollancz, 1963).

[22] J.W. Scott, *Gender and the Politics of History* (New York: Columbia University Press, 1988).

[23] G. Austin and S. Broadberry, 'Introduction: The Renaissance of African Economic History', *The Economic History Review* 67, no. 4 (2014): 893–906.

particularly since the Second World War – necessitated a new approach. The new field of global history was interested in pinpointing the date when the gap between the West and the rest began to open up.[24] Economists and political scientists rediscovered institutions; history and its ability to explain persistent differences in living standards and political systems were suddenly top of the agenda.[25] And in economics a methodological revolution had begun: the credibility revolution, with empirical hypothesis-testing preferred to theoretical models.

On the supply side, the immense advances in computing power, the internet and the availability of statistical software packages lowered the costs of both statistical analysis and knowledge acquisition. As economic historian Steven Ruggles explains, a wealth of new data is now available with a few clicks: 'IPUMS has made freely available billions of cases of historical data from hundreds of censuses and surveys taken in more than 100 countries over the past 250 years.'[26] Instead of the large teams required to run quantitative history projects in the 1970s, individual historians, armed with a laptop and enough patience to watch a few online tutorials, can now construct their own artisanal datasets, create a pivot table and run a regression. As I explain in the final chapter, free software like R and its community of contributors further democratised (and continues to democratise) data discovery and coding. New user-generated packages give us new ways of documenting, describing and displaying data.

Despite this renewed interest in quantification, however, traditional historians remain largely sceptical. In fact, much of this revival has come not from historians but from other social scientists – economists, political scientists, demographers – just as the first wave of quantification also drew impetus from outside. What are the merits of quantification, then, and why is this time different? Will enthusiasm not inevitably be followed by disillusionment? In the next section, I explain why we can be optimistic that this is not simply history repeating itself.

A history from below

There have always been good reasons to count things. 'Many qualitative judgements or descriptions used by historians', said the economic historian Roderick Floud, 'have an implicit quantitative significance, which it is sometimes necessary to make explicit'. He added that 'many descriptions of the behaviour of individuals or groups have quantitative significance; such words as "usually," "normally," "often," "many," refer to quantitative

[24] K. Pomeranz, *The Great Divergence* (Princeton: Princeton University Press, 2021).

[25] D. Acemoglu, S. Johnson and J.A. Robinson, 'Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution', *The Quarterly Journal of Economics* 117, no. 4 (2002): 1231–94.

[26] S. Ruggles, 'The Revival of Quantification: Reflections on Old New Histories', *Social Science History* 45, no. 1 (2021): 21.

concepts, and although we may often not wish to test them exactly, in principle their significance or truth can only be established by quantitative measurement'.[27]

The demographic historian Peter Laslett was somewhat more blunt:

> Writers of history in [the] established, conventional mode are perpetually proclaiming that such and such an individual was typical, such and such an event or a trait was characteristic. But they do so without apparently recognising that to decide on typicality requires that the whole class of the phenomena under examination has to be surveyed and that all possible individual events or all possible traits must be taken into account.[28]

Quantification helps us identify whether an observation or sample of observations is representative of the total. Imagine a distribution of per capita incomes in the year 1500, or a distribution of marriage ages in the year 1920, or a distribution of experiences at a recent festival. The income per capita of Spain in 1500, or the age of a bride in 1920 or the complaints of a festival attendee can only be analysed once we know where on the distribution that income level, or occupation or experience fell. If we want to know whether Spain was rich or poor in 1500, we need to know whether other countries were richer or poorer. If we want to know whether it was common for a wife to be the same age as her husband in 1920, we need to know the ages of all women that married in that year – or at least a large proportion. If we want to know whether the experience of the festival attendee is the exception rather than the rule, we need to know the experiences of others.

Naturally, some things are easier to quantify than others. Incomes and ages are, by definition, numerical. Festival experiences less so. That is a concern that historians have had for some time, at least since Arthur Schlesinger Jr claimed that 'almost all important questions are important precisely because they are not susceptible to quantitative answer'.[29] There are two responses to this, one old and one new. It is true that not all things can be quantified, but it is equally true that quantitative evidence helps us understand the qualitative. As Floud put it: 'Quantitative evidence will almost certainly not provide a complete answer, but it may well provide some of the answer, and to throw it away unseen is both wasteful and irresponsible.'[30]

The new response is that, in the past two decades, new tools and techniques have allowed us to investigate a broader array of issues than was available to historians in the 1970s. The information and communication revolution has shifted not only how

---

[27]R. Floud, *An Introduction to Quantitative Methods for Historians* (London: Methuen, 1973): 2.

[28]P. Laslett, 'Signifying Nothing: Traditional History, Local History, Statistics and Computing', *History and Computing* 11, nos. 1 & 2 (1999): 130.

[29]A. Schlesinger Jr, 'The Humanist Looks at Empirical Social Research', *American Sociological Review* 27, no. 6 (1962): 770.

[30]Floud, *Introduction to Quantitative Methods for Historians*, 3.

sources are preserved but also the questions we ask and the ways we attempt to answer them. The historian of technology Adam Crymble notes that ‘in a field so quick to point out why it is important for us to have a deep knowledge of the past, historians have been slow to recognise how their work and their ideas have not been immune to the transformative rise of the computer, which has brought about the greatest social and cultural transformation the world has perhaps ever seen’.[31] Optical character recognition (OCR) tools allow us to rapidly transcribe large corpuses of text, increasingly even when handwritten. Geographic information systems (GIS) make accurate spatial analysis easy. Free statistical software packages are available to match observations and merge datasets, map and analyse networks of relationships, and even use machine learning tools to analyse historical photos, videos or soundbites. The computational toolkit has expanded to such an extent that even Schlesinger Jr would have to admit that *some* important questions are now within the purview of the quantitative.

Asking the important questions might indeed be what historians can bring to the table. Elsewhere, I have argued that historians, ‘because of their ability to operate at the intersections of the different social sciences – and perhaps the natural sciences in some ways too – … should be well-positioned to ask the big research questions about human and societal behaviour’.[32] That implies, however, being open to the use of numbers. Historian Caitlin Rosenthal, one of the main proponents of a ‘new history of capitalism’ literature, warns her history colleagues that ‘we ignore the data-driven work of economic historians at our peril’. She continues: ‘Even if we do not need their questions, their data and methods can help us to discipline our inquiries – to rule out interpretations that are incorrect and to suggest new ways forward.’[33]

Such data need not be Big Data. True, large datasets with thousands or even hundreds of thousands of observations are now easy to create and curate, owing to cheaper digitisation and transcription techniques, from OCR to outsourcing, and an abundance of (open-source) software packages, with R and Python being the most popular. But one strand of quantitative history involves what I call ‘artisanal’ quantification. Increasingly, our databases can include ‘messy’ entries, with notes. Because file size is of no concern anymore (no more need to fit everything onto a punch card), and because packages like Microsoft Excel are accessible and easy to use, and because we have a variety of techniques for dealing with both numerical values and text, we can preserve original entries in datasets before the data is ‘cleaned’ to allow for analysis. In fact, in some cases, it is exactly the ‘uncleaned’ versions of the data that allow the researcher to uncover new facts.

I will use an example from this book, Chapter 5, on race reclassification in apartheid-era Cape Town, to show what I mean. This study’s small sample of 100 cases, drawn from

[31]A. Crymble, *Technology and the Historian: Transformations in the Digital Age* (Urbana: University of Illinois Press, 2021): 2.

[32]J. Fourie, et al., ‘Making South African Historians Count’, *Historia* 66, no. 1 (2021): 4.

[33]C. Rosenthal, ‘Seeking a Quantitative Middle Ground: Reflections on Methods and Opportunities in Economic History’, *Journal of the Early Republic* 36, no. 4 (2016): 659–80.v

a population of more than 50,000 marriage records, was found by looking at notes in the marriage records indicating when a couple was reclassified. If a standardised, 'cleaned' dataset had been used, the reclassification details would have likely been ignored because they appeared on a very small percentage of records. This kind of study, say Lemercier and Zalc, 'can be part of an experimental practice of history that is playful, in the sense of not boring (yet attentive to ethics): numbers, when used as tools, not fetishes, allow familiarisation, comparison, and oblique readings of sources'.[34]

But there is another reason to use quantitative approaches, one that has not received the attention it deserves. Quantitative history can help to illuminate the experiences of people who have remained hidden in the conventional archive. Large administrative datasets include information about people who are absent from the archival sources generally used in narrative accounts, such as letters, diaries, newspapers and reports. Digitising, transcribing and analysing such administrative sources opens new opportunities to account for those people.

To date, ways of writing histories from below have had some successes and suffered from some shortcomings. To explain, it helps, again, to begin with the *Annales* school. Not only was the 'total history' promoted by *Annales* historians closely linked to quantification, often of the artisanal kind; their interest was explicitly less in the deeds of great men than in the experiences of ordinary people. While their focus was initially on compiling long-term series (of, say, prices and wages) and identifying trends and patterns of different duration, their emphasis had always been on incorporating the histories of those outside the domains of power and privilege. To do so, their national and regional focus increasingly shifted towards the local. One popular approach of later *Annalistes* was to write microstudies of villages, relying on detailed quantitative and qualitative information to bring to life the lives of women, children, peasants and other overlooked people.

These approaches can be seen as a prelude to the microhistory approach that would later develop in Italy; an approach, historian Giovanni Levi explains, that advocates for incorporating the research process into the narrative itself.[35] Microhistorians, says Lemercier and Zalc, 'allowed us to think of quantification in constructionist terms suited to the scales of interaction among individuals, and to approach it experimentally – open to non-standard procedures, including the idea that meaningful results can be obtained by studying a small, situated unit, such as a family or a village, systematically and even quantitatively'.[36] Not only did this kind of quantification help to identify the exceptional or outliers (rather than the mean or the median), it allowed data construction (and the source criticism associated with it) to become part of the text – 'not preliminary

---

[34]Lemercier and Zalc, *Quantitative Methods,* 27.

[35]G. Levi, 'On Microhistory', in *New Perspectives on Historical Writing*, ed. P. Burke (Cambridge: Polity Press, 1991): 93–113.

[36]C. Lemercier and C. Zalc, 'Back to the Sources: Practicing and Teaching Quantitative History in the 2020s', *Capitalism: A Journal of History and Economics* 2, no. 2 (2021): 480.

operations to be dispensed with or standardised'.[37] The focus on the exceptional or outliers, however, meant that these approaches had difficulty integrating the stories of those uncharted back into the larger narratives of history. Microhistories of this kind, apart from a few notable exceptions, had the unfortunate consequence that their impact was micro too.

The most popular 'histories from below', by contrast, did not rely on numbers. In fact, some historians feared that large datasets pushed them to write 'histories without people'. In Britain, an early initiative to write history from below, sometimes referred to as the 'Old History from Below', was by British Marxists. Starting in the 1940s, a group of historians that included E. P. Thompson, Eric Hobsbawm, Rodney Hilton, Christopher Hill, Raphael Samuel and George Rude combined their political activism in the Communist Party with a desire to write a new kind of social history, one of and for the working class.[38] In contrast to the more structural social histories of the time which, they believed, reduced the working class to a simple statistical artefact, theirs followed a humanist, narrative approach that explored the social and cultural dimensions of politics.[39] Thompson's widely popular *The Making of the English Working Class* is a classic example of this bottom-up view, emphasising the creation, by the late nineteenth century, of a working-class identity: 'The working class made itself as much as it was made'.[40] It was an argument that would prove immensely popular and powerful outside of England, notably in India, where Subaltern Studies would emerge as an energetic new approach to writing history from below.[41]

Subaltern Studies benefited from the popularity of postcolonial studies and critical theory, which had roots in postmodernism. By the 1970s, a group of French philosophers that included Michel Foucault, Jacques Derrida and Jean-François Lyotard began to offer a radical critique of objectivity, truth, reality and reason. Pointing to unequal power structures in the past and present, historians who adopted postmodernist approaches refused to prioritise one aspect of history over another, often preferring the trivial to the grand narrative and the symbolic to the real. In short, postmodern historians began to question not only the partialness of archival sources, but the archive itself.

In some cases, the belief that archival sources are more myth than material verged on the absurd: 'The idea of the historical past can … be considered as just one more example of the many imaginaries we have fabricated to help us make some sense of the apparent senselessness of existence', said the historian Keith Jenkins.[42] Not all historians,

---

[37]Ibid, 480.

[38]S. Scott-Brown, 'The Art of the Organiser: Raphael Samuel and the Origins of the History Workshop', *History of Education* 45, no. 3 (2016): 372–90.

[39]K. Gentry, 'Ruskin, Radicalism and Raphael Samuel: Politics, Pedagogy and the Origins of the History Workshop', *History Workshop Journal* 76, no. 1 (2013): 187–211.

[40]Thompson, *Making of the English Working Class*, 194.

[41]S. Bhattacharya, 'History from Below', *Social Scientist* 11, no. 4 (1983): 3–20; V. Chaturvedi, ed., *Mapping Subaltern Studies and the Postcolonial* (London: Verso, 2000).

[42]K. Jenkins, *Why History? Ethics and Postmodernity* (London: Routledge, 1999): 14.

fortunately, adopted such an extreme view. By the late 1990s, historian Richard J. Evans could declare that history 'is an empirical discipline, and it is concerned with the content of knowledge rather than its nature. Through the sources we use, and the methods with which we handle them, we can, if we are very careful and thorough, approach a reconstruction of past reality that may be partial and provisional, and certainly will not be objective, but is nevertheless true'.[43]

Another belief of postmodernist historians is that the authors' identity is inextricably bound up with their subject. Why and by whom history was written sometimes became more important than what was written. Although the idea that each group writes its own history is often unrealistic, it is true that people who write their own histories will bring their own concerns to bear on them. There is thus a justifiable concern that not only have some groups been systematically excluded from the archive, but those who should write their histories have been excluded from the academic profession. This is particularly true of quantitative history, a field that has not been very successful at attracting scholars from marginalised backgrounds. As I will argue in the conclusion, South Africa, with its history of racial discrimination, is a sad example of such inability and the reason why the Biography project, and the tools it will have created, is one small attempt at righting the imbalance.

Attracting a more diverse group of researchers is just one challenge in writing more complete histories from below. A more fundamental challenge, in the past and in the present, is the lack of sources. Some histories are missing because no evidence – of an event, a person, a thought – was ever recorded. One obvious example where this happens is an illiterate society. Oral history approaches have been of immense benefit. Just consider the excellent histories of African societies based on oral testimonies.[44]

But, as oral historians know only too well, such sources have shortcomings. Human memory is fickle. There are good reasons why humans remember poorly. Psychologists have shown that humans have a negativity bias: negative information and experiences are more likely to be remembered than equally extreme positive information and experiences.[45] People are more likely to search their memory for certain events if prompted to do so by other events. For example, patients who have been diagnosed with cancer are more likely to report previous minor illnesses than patients who have not been

[43]R.J. Evans, *In Defence of History* (London: Granta, 1997): 249.

[44]J.M. Vansina, *Oral Tradition as History* (Madison: University of Wisconsin Press, 1985). For South Africa, see B. Bozzoli and M. Nkotsoe, *Women of Phokeng: Consciousness Life Strategy and Migrancy in South Africa, 1900–1983* (Johannesburg: Ravan, 1991); I. Hofmeyr, *We Spend Our Years as a Tale that Is Told: Oral Historical Narrative in a South African Chiefdom* (Johannesburg: University of Witwatersrand Press, 1993). For more recent work in African history, see N. Achebe, *The Female King of Colonial Nigeria: Ahebi Ugbabe* (Bloomington: Indiana University Press, 2011); A. Wiemers, *Village Work: Development and Rural Statecraft in Twentieth-century Ghana* (Athens, US: Ohio University Press, 2021); E.L. Osborn, *Our New Husbands Are Here: Households, Gender, and Politics in a West African State from the Slave Trade to Colonial Rule* (Athens, US: Ohio University Press, 2011).

[45]C.J. Norris, 'The Negativity Bias, Revisited: Evidence from Neuroscience Measures and An Individual Differences Approach', *Social Neuroscience* 16, no. 1 (2021): 68–82.

diagnosed with cancer. Researchers need to avoid this 'recall bias'. Neuroscientists know that human evolution shaped story-telling into a useful tool for survival. But, argues the philosopher Alex Rosenberg, as narratives are almost always wrong, historians have turned this useful tool into a defective theory of human nature:

> If we humans are ever to move beyond our internecine histories, we will have to put historical 'understanding' behind us. We will have to recognise that even the best histories we can contrive are mostly wrong or, when right, are right by accident, that they fail to identify the real causal forces that drive events, that they obstruct efforts to really understand our past, and that they serve as harmful tools of the worse angels of our natures.[46]

The sociologist Zygmunt Bauman has argued that the need to reinvent the past also depends on our future outlook: 'Having lost all visions of an alternative – better – society of the future – associating the future, if not with "worse than the present" then with "more of the same"' – no wonder that, when seeking genuinely meaningful ideas, we turn, nostalgically, to the buried grand ideas of the past.'[47] In short, then, our recollections of the past are affected by our preference for the negative, the irrelevant and story-telling, and by our hopes for the future.

The problem is not only what gets remembered, but who is asked for their memories. Oral histories introduce concerns that are front-of-mind in survey design: How do we know we have a representative sample of views? Should we pay more attention to the average experience or the outlier? To what extent do we account for survivorship bias, the tendency to value the experiences of those who survived over the experiences of those who did not? What are the motives of the interviewer and interviewee? Not only are humans bad at remembering, but memory fades; testimonies become less reliable the further we go back in time, with serious implications for oral histories going back two or more generations.

Of course, oral testimonies are not the only method to have shortcomings. Where written records do remain, they invariably provide only a partial view of the past. The histories of the poor, the powerless and those on the periphery of society are often excluded because any evidence of their thoughts, feelings and experiences was lost or not preserved, either intentionally or unintentionally. British historian E. H. Carr observed in 1961 that historians are not impartial actors in the preservation of knowledge: prevailing preferences, politics and protocols shape what they preserve and protect, and the same can be said of archivists.[48]

---

[46] A. Rosenberg, *How History Gets Things Wrong: The Neuroscience of Our Addiction to Stories* (Cambridge, US: MIT Press, 2019).

[47] Z. Bauman, *Retrotopia* (Cambridge: Polity Press, 2017): 128.

[48] E.H. Carr, *What Is History?* (1961; repr. London: Penguin Random House, 2018).

Financial considerations also play a role in the creation and preservation of written records. For both supply and demand-side reasons, recordkeeping is expensive. In the late nineteenth-century Cape Colony, land surveys of African property, for example, were first simplified (in comparison to white landowners) to speed up the process and then, when the costs outweighed the benefits, abandoned completely.[49] In the past and present, the budgetary requirements of preserving large amounts of seemingly unimportant documentation in fiscally constrained environments have also meant the destruction of archival material. This is particularly true for poorer regions, those places where uncharted people most likely reside.

On the demand side, people may have had good reason to want to avoid being recorded, making it costly (or nigh impossible) for the state to record accurate information. Tax censuses are a classic example; in earlier work, I found that eighteenth-century Cape farmers substantially underreported their assets to avoid tax.[50] In addition, for economic, political or social reasons, the formerly oppressed may not want their histories to be known. Ally describes how the residents of the former Bantustan KaNgwane 'actively lay the past to waste', destroying the archive of the apartheid 'puppet' state that subjugated them.[51]

Where historical texts do remain, however, reading them against the grain is one way to remedy or, in the least, uncover biases. It can help to expose the unexamined beliefs and attitudes of the authors, and allow us to draw attention to the silences, contradictions and prejudices of the archive. In fact, as Stoler has shown, archives – and in particular colonial archives – themselves can and should be subjects of study, allowing the historian to not only work against but along the archival grain.[52] Quantitative history, I argue, can perform a similar function.

It can do this because of two recent technological advances: the digitisation and transcription of large quantities of archival material, and the availability of a statistical toolkit to analyse these at low cost. The digitisation of archival sources has opened up vast quantities of archival sources for analysis. The most common of these are the volumes of parliamentary papers, colonial reports, letters and other text-rich documents that historians have traditionally relied on. But another source – what I call administrative records – has the potential to illuminate parts of society that may have been hidden from view: birth records, death notices, probate inventories, voters' rolls and many more. These individual-level records are uninteresting on their own; other than, perhaps, for

---

[49] R.T. Ally, 'The Development of the System of Individual Tenure for Africans: With Special Reference to the Glen Grey Act, c. 1984–1922' (MA diss., Rhodes University, Makhanda, 1985).

[50] J. Fourie, 'The Remarkable Wealth of the Dutch Cape Colony: Measurements from Eighteenth-century Probate Inventories', *The Economic History Review* 66, no. 2 (2013): 419–48.

[51] S. Ally, 'Material Remains: Artifice Versus Artefact (s) in the Archive of Bantustan Rule', *Journal of Southern African Studies* 41, no. 5 (2015): 969–89.

[52] A.L. Stoler, *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense* (Princeton: Princeton University Press, 2009).

the individuals' immediate descendants, or for genealogists. But combining them into a large dataset and matching individuals across many such sources allows us to not only reconstruct trends in the previously uncharted groups – of their fertility, infant mortality, wealth, mobility and much more – but also to arrive at new causal insights that will provide a more nuanced and complete interpretation of historical change. This is especially applicable, as I argue elsewhere, in settings where conventional archival sources are limited.[53]

These methods are not without their detractors. As I mentioned earlier, the first wave of quantitative history, much of which was also based on large volumes of such administrative records, was criticised for its lack of definitive conclusions, inability to answer important questions, technical difficulty and high costs, among other things. The biggest concern now is undeniably ethical. Responding in a Broadstreet blog post to a January 2022 *New York Times* column by Jamelle Bouie on American slavery and quantification, political scientist Emily Sellars lists the ethical concerns when historians begin to quantify: 'the extent to which quantifying history can be "dehumanising", the practical and ethical consequences of relying on data generated by immoral historical processes like the slave trade, how our academic projects or datasets might be misused or misinterpreted, and, most importantly, what our ethical responsibilities are to those we study, to their descendants, and to society as a whole'.

These are all valid concerns, and are beginning to receive the attention they deserve.[54] There is no denying that working with numbers – rather than, say, an individual story – exacts less of an emotional toll. This has two consequences. First, it makes quantification less attractive for those hoping to effect change. As activists know all too well, statistics rarely make a good slogan: a picture is often worth a thousand numbers. Second, and perhaps more seriously, numbers that are abstracted from their context can indeed be misused and, in the extreme case, become dehumanising. At least since Foucault, historians have known that the data-generating process itself is often an exercise of power, usually by the state.

But as Sellars remarks in the same January 2022 post, the impersonality can also be an advantage: 'There are some ways in which quantitative research may be *less* exploitative than qualitative research. It may be easier to preserve the confidentiality of specific people, for example, which is an important consideration given that our research subjects can't usually consent to having their stories shared.'[55] Another advantage, already highlighted above, is that quantification helps to distinguish the average from the outlier. This is helpful when dealing with narratives, especially the narratives of those

---

[53] J. Fourie, 'The Data Revolution in African Economic History', *Journal of Interdisciplinary History* 47, no. 2 (2016): 193–212.

[54] A.M. Lerner, 'Quantifying the Archives: Leveraging the Norms and Tools of Data Science to Conduct Ethical Research on the Holocaust', *Holocaust Studies* 28, no. 3 (2021): 1–19.

[55] E. Sellars, 'What Are Our Ethical Responsibilities as HPE Scholars?', *Broadstreet*, 7 February 2022. Available online: https://broadstreet.blog/2022/02/07/what-are-our-ethical-responsibilities-as-hpe-scholars/.

who have been previously overlooked. If we find an example of a slave who did not want to be freed, does that mean that all slaves felt the same way? Says Sellars: 'One thing that quantitative research can contribute is to illustrate how distorted and misleading those narratives are'.[56] Quantification can also help us understand scale: 'quantifying the scope of processes like the slave trade makes it difficult to dismiss them as historical blips'.[57]

Finally, data generated by coercive systems can easily replicate the biases of those systems. This is, of course, true for both quantitative and qualitative data. But quantification has two properties that, I argue, have distinct advantages over qualitative approaches. Firstly, it can explicitly reveal the size of those biases. To give just one example: in work with Jonathan Jayes, I used a novel technique to calculate the extent of discrimination by white doctors when treating black patients during the Spanish flu.[58] (We found little evidence of such discrimination, which accords with the qualitative evidence of the time.) Secondly, quantification can help to answer questions orthogonal to the original purpose of data collection. Consider that anthropometric historians often make use of the heights of soldiers reported in attestation forms. Although the colonial recruits might have suffered from the prejudices of their superiors, it is unlikely that those prejudices would have manifested in biased measurements of a man's height.[59] But the standards of living that can be calculated using the height data have helped anthropometric historians to expose the ways in which colonial policies disadvantaged certain groups in society.

If we are to account for the people lost in the records or never recorded, we should do so in a way that is sensitive to the data-generating process, acknowledges the coercive nature of state bureaucracy, questions arbitrary classifications and categories and, in short, takes context seriously.

Quantitative history in South Africa

South Africa is a good place to show how quantification can help write 'histories from below'. That assertion might surprise many readers. According to historian Keith Breckenridge, civil registration in South Africa was a 'complete and prolonged failure'.[60] The proximate causes of such failure were obvious: a lack of state funding and the racist disparagement of African subjects by some white officials. A more basic problem, argues

---

[56]Ibid.

[57]Ibid.

[58]J. Fourie and J. Jayes, 'Health Inequality and the 1918 Influenza in South Africa', *World Development* 141 (2021): 105407.

[59]A. Moradi, 'Towards an Objective Account of Nutrition and Health in Colonial Kenya: A Study of Stature in African Army Recruits and Civilians, 1880–1980', *The Journal of Economic History* 69, no. 3 (2009): 719–54.

[60]K. Breckenridge, *Biometric State: The Global Politics of Identification and Surveillance in South Africa, 1850 to the Present* (Cambridge: Cambridge University Press, 2014): 136.

Breckenridge, was the absence of a political constituency demanding civil registration: 'Here the contrast with the many different forms of registration in other regions of the world is striking, because both elites and the poor – who, in other contexts, sought registration for many different ends – were silenced by the workings of customary law.'[61] The inertia of civil registration allowed the apartheid-era government to launch a project of universal fingerprint registration; a project that would, ironically and unintentionally, make possible the well-targeted social transfers programmes of democratic-era South Africa.[62]

Breckenridge's claim that civil registration in South Africa was a failure is only partially true. It is true for much of the region that became the Union of South Africa in 1910, but in the British Cape Colony (1806–1910) and for the Cape Province (1910–94), regions outside the scope of Breckenridge's study, civil registrations were widespread. This is true across population groups. It is true for those that would be classified under the later apartheid-era Population Registration Act as white as well as for those classified as coloured and black, although, as one would expect, there was likely to be some regional variation. As Chapter 12 on the Spanish flu shows, the criteria for registering a death were different in rural and urban areas.

The series on death registrations is just one of many sets of records that remain well preserved, in the Cape Archives and elsewhere. The British colonial authorities collected and preserved a copious amount of information on its subjects. Several censuses were undertaken, in 1865, 1875, 1891 and 1904, although, sadly, the original records, with a few exceptions, have been lost. The liberal franchise at the Cape meant that there was no legal reason to exclude certain groups. Except for the 1865 census, which includes only white and coloured inhabitants, the censuses made a full enumeration of the population in the territories annexed by the British. Chapter 3 uses these and the South African censuses to calculate and analyse under-five sex ratios by race.

The records go much further than census data. Huge numbers of probate inventories were preserved and these have now been made available online on FamilySearch.org.[63] More than 3 million civil records of the Cape Province are available on the same site, as well as 2.7 million civil marriage records for South Africa (which, presumably, would include the Cape Province). But it was not only the state that collected individual-level information. The two largest religious denominations, the Dutch Reformed Church and the Anglican Church, collected and preserved baptism and marriage records and membership lists. Chapter 2 matches Anglican baptism and marriage records to calculate bridal pregnancy, for example. Records of other denominations, like the Lutheran Church in Cape Town, remain well preserved too, although less accessible to the public. Churches collected not just vital statistics but financial statistics too: a series

[61]Ibid, 136.

[62]Ibid, 188–93.

[63]As of January 2022, the total number of probate records for the Cape Province (1834–1989) was 2,752,260.

of cashbook transactions across more than a century remains preserved in the Dutch Reformed Church Archive in Stellenbosch.

Why was recordkeeping so much better at the Cape than elsewhere in South Africa? Two reasons come to mind. Firstly, and most pertinently, Cape Town and colonial Cape society was founded by the first multinational company, the Dutch East India Company (*Vereenigde Oostindische Compagnie*), in the mid-seventeenth century. Because of the profit-maximising motive and for the purposes of reporting to shareholders, the VOC kept meticulous records of all its activities and subjects throughout the 143 years of its rule.[64] British rule (first from 1795 and then again from 1806) largely adopted the existing Dutch bureaucracy.

The VOC was remarkably meticulous in its recordkeeping. From the 1660s, an annual tax census was collected of all free households. The *opgaafrolle* include not only information about the settler family and the workforce (wage labourers, Khoe and slaves), but also a detailed breakdown of agricultural assets, inputs and outputs. It is a remarkable series, one that is currently being transcribed as part of the Cape of Good Hope Panel project.[65] The VOC-era records are not limited to farm production: the Master of the Orphan Chamber collected probate inventories, auction rolls and several other series. As Dutch economic historian Jan Luiten van Zanden remarked, the Cape Archive 'has a more complete record of eighteenth-century Cape Colony life than what is available in The Hague for economic historians of eighteenth-century Holland'.[66]

A second reason for the large difference in data quality between the Cape Province and the rest of South Africa can be found, perhaps, in the reasons for the Great Trek. One purported reason for frontier farmers in the 1830s to migrate into the interior was to escape British rule. The extent to which this explains the migration of more than 5,000 settler farmers is still debated, but it does explain why those who settled the interior were unwilling – quite literally – to be charted, or in other words, put on the map. This was especially true for those farmers who migrated into the territories that would become known as the Transvaal. Dimitruk, Du Plessis and Du Plessis have shown how the Transvaal's poor administrative capacity was reflected in less accurate maps, which hindered the provision of property rights.[67] It was only towards the second half of the twentieth century that more accurate maps became available. Poor state capacity for much of the late nineteenth and early twentieth century probably also explains the lack of good civil registration records.

---

[64] J. Fourie, 'The Quantitative Cape: A Review of the New Historiography of the Dutch Cape Colony', *South African Historical Journal* 66, no. 1 (2014): 142–68.

[65] J. Fourie and E. Green, 'Building the Cape of Good Hope Panel', *The History of the Family* 23, no. 3 (2018): 493–502; A. Rijpma, J. Cilliers and J. Fourie, 'Record Linkage in the Cape of Good Hope Panel', *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53, no. 2 (2020): 112–29.

[66] Quoted in J. Fourie and S. Schirmer, 'The Future of South African Economic History', *Economic History of Developing Regions* 27, no. 1 (2012): 114–24.

[67] K. Dimitruk, S. Du Plessis and S. Du Plessis, 'De jure Property Rights and State Capacity: Evidence from Land Specification in the Boer Republics', *Journal of Institutional Economics* 17, no. 5 (2021): 764–80.

Despite the larger quantity and better quality of administrative records for the Cape Colony and Cape Province, South African historians have largely neglected these sources. There have been exceptions. By the 1980s, just when historians elsewhere turned away from quantification, a new generation of South African historians (and historians interested in South Africa) turned to the study of Cape slavery. Inspired by the political changes that were afoot in the country, and an openness to counter-narratives of the colonial past, these scholars made extensive use of the abundance of VOC records to expose the centrality of slavery to the Cape economy.

Although historians had often used statistics to support their hypotheses, what was unique about this generation was their willingness to generate and analyse their own datasets. Nigel Worden calculated farm productivity.[68] Leonard Guelke reconstructed land grants to investigate inequality.[69] Robert Ross and Pieter van Duin used tax censuses to calculate population, production and export totals for the Dutch period.[70] Hans Heese used rich genealogical records meticulously reconstructed by amateur genealogists in the 1960s and 1970s to investigate the evolution of settler and slave families.[71] Many of these contributions appeared in *The Shaping of South African Society, 1652–1820*, an impressive collection edited by Richard Elphick and Hermann Giliomee.[72] Robert Shell's work deserves special mention. Studying for a master's degree at the University of Rochester in the late 1970s, Shell would certainly have been influenced by Stanley Engerman, co-author of *Time on the Cross*. He went on to do a PhD at Yale on 'Slavery at the Cape of Good Hope, 1680–1731' and continued to publish extensive series of statistics and visualisations on Cape slavery.[73] Before his untimely death in 2015, Shell was proud of being 'South Africa's only trained cliometrician'.[74]

Despite the work by Shell and others on slave statistics, the 1980s 'cultural turn' in history quickly stymied the possibility of further quantification.[75] I offer two further

[68]N. Worden, *Slavery in Dutch South Africa*, vol. 44 (Cambridge: Cambridge University Press, 1985).

[69]L. Guelke and R. Shell, 'An Early Colonial Landed Gentry: Land and Wealth in the Cape Colony 1682–1731', *Journal of Historical Geography* 9, no. 3 (1983): 265–86.

[70]J. Fourie, 'Subverting the Standard View of the Cape Economy: Robert Ross's Cliometric Contribution and the Work It Inspired', *Magnifying Perspectives: Contributions to History, a Festschrift for Robert Ross, ASC Occasional Publication* 26 (2014): 261–73.

[71]H.F. Heese, 'Slawegesinne in die Wes-Kaap, 1665–1795', *Kronos: Journal of Cape History* 4, no. 1 (1981): 38–48.

[72]R. Elphick and H. Giliomee, eds., *The Shaping of South African Society, 1652–1820* (Cape Town: Maskew Millar Longman, 1984).

[73]R. Shell, *Children of Bondage: A Social History of the Slave Society at the Cape of Good Hope, 1652–1838* (Hanover, US: Wesleyan and University Press of New England, 1995).

[74]In personal communication.

[75]Here, too, there were exceptions. A decade or more later, Susan Newton-King used probates to investigate the wealth of settler farmers and Wayne Dooling used land transactions to investigate social mobility. See S. Newton-King, *Masters and Servants on the Cape Eastern Frontier, 1760–1803*, vol. 97 (Cambridge: Cambridge University Press, 1999); W. Dooling, 'The Making of a Colonial Elite: Property, Family and Landed Stability in the Cape Colony, c. 1750–1834', *Journal of Southern African Studies* 31, no. 1 (2005): 147–62.

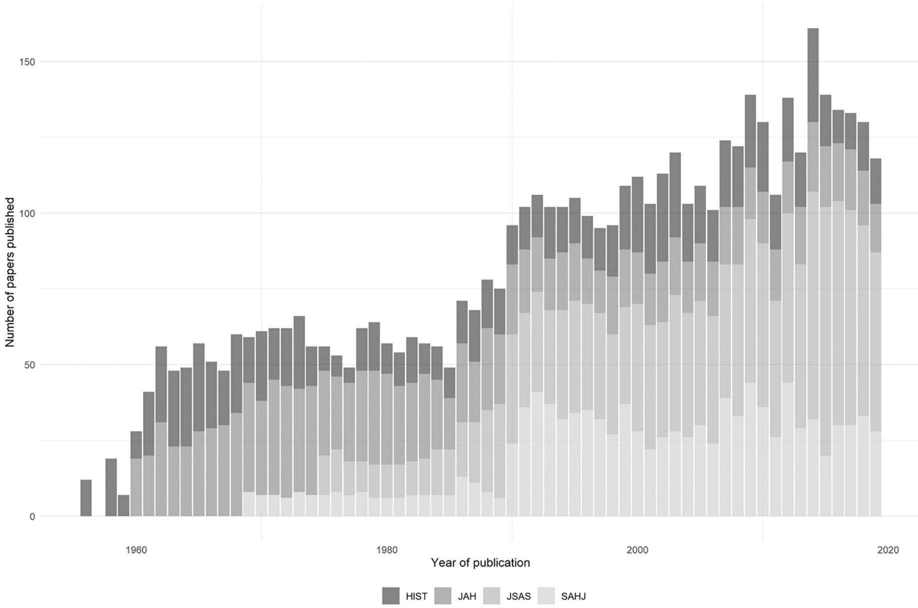

Figure 1.1 Total publications in four history journals, by year (1956–2019).

pieces of evidence, one qualitative and one quantitative. I recently asked Charles van Onselen, one of South Africa's leading historians, what role quantification played in South African historiography. He replied:

> Did quantification ever arise here in the sense that you are after? If it did then, as usual, I did not notice it. I think that there were some social scientists in the inter-war years who tried to quantify things on, say, poverty, and so on, but I am not sure that any historians attempted anything similar. I think that the historians tended to raid such findings to support their arguments, but I do not know they ever generated any primary data themselves.[76]

I found support for his opinion from a very different source. To quantify the extent of quantification in South African history, I analysed all the research papers published in four prominent history journals between 1956 and 2019: *Historia*, the *Journal of African History*, the *Journal of Southern African Studies* and the *South African Historical Journal*. A total of 5,217 papers were published over the period. Figure 1.1 shows the number of papers published by year between 1956 and 2019. The average is just over fifty for much

[76]Personal e-mail correspondence. 10 February 2020.

of the period until the 1980s, and then increases rapidly to a new equilibrium around 100. By the late 2000s, another shift occurs, although less obvious, rising to around 130 papers per year.

I wanted to know whether this increase in history output also resulted in more papers of a quantitative nature. To count such papers, I counted the number of tables, graphs, maps and images in each paper. I classified a paper that included a table as a quantitative history paper. This was of course a broad assumption – some papers that include tables might not be considered 'quantitative', while others that are quantitative might not include a table – but I believe it provides a useful proxy for research that uses numbers, to put it simply.

Besides recording these in-paper features, I also recorded whether a paper had one or more authors and the number of Google citations in 2020 for that paper. Table 1.1 provides an overview of the four journals, with the average for each of these features reported by journal.

The statistics reveal several noteworthy differences between the journals. The two more 'international' journals, the *Journal of Southern African Studies* and the *Journal of African History*, dedicated to publishing research beyond South Africa, have not only published a larger number of papers – a surprising finding, given that the *JSAS* was first published in 1975, almost two decades after *Historia* – but its papers also receive significantly more citations. Papers published in the *JAH* receive almost twenty times more Google citations than papers published in *Historia*. Another noteworthy difference is the number of single-authored papers. The *JAH* publishes very few co-authored papers, roughly one in sixteen. By contrast, one in seven papers in the *JSAS* has more than one author. This is probably due to the interdisciplinary nature of the *JSAS*. As we will see later, there is also a noticeable time trend associated with multi-authored papers.

A third noteworthy trend is the difference in the average number of tables, graphs, maps and images published per paper. Here, again, there are two groups. The *JSAS* and the *JAH* publish significantly more tables and graphs than do the two South Africa-specific journals: whereas the *JSAS* and *JAH* publish, on average, a table and a graph in every

Table 1.1 Descriptive statistics of journal papers.

Journal	*Number*	*Citations*	*Author*	*Tables*	*Graphs*	*Maps*	*Images*
SAHJ	1100	6.7	0.09	0.18	0.05	0.09	0.49
JSAS	1640	33.0	0.14	0.52	0.14	0.18	0.33
HIST	1091	2.0	0.07	0.24	0.03	0.14	0.61
JAH	1386	39.0	0.06	0.50	0.12	0.57	0.45

Notes: 'Number' refers to the number of observations (papers). 'Citations' refers to the average number of citations per paper. 'Author' reports the share of papers with more than one author. 'Tables', 'Graphs', 'Maps' and 'Images' count the number of tables, graphs, maps and images per paper, respectively.

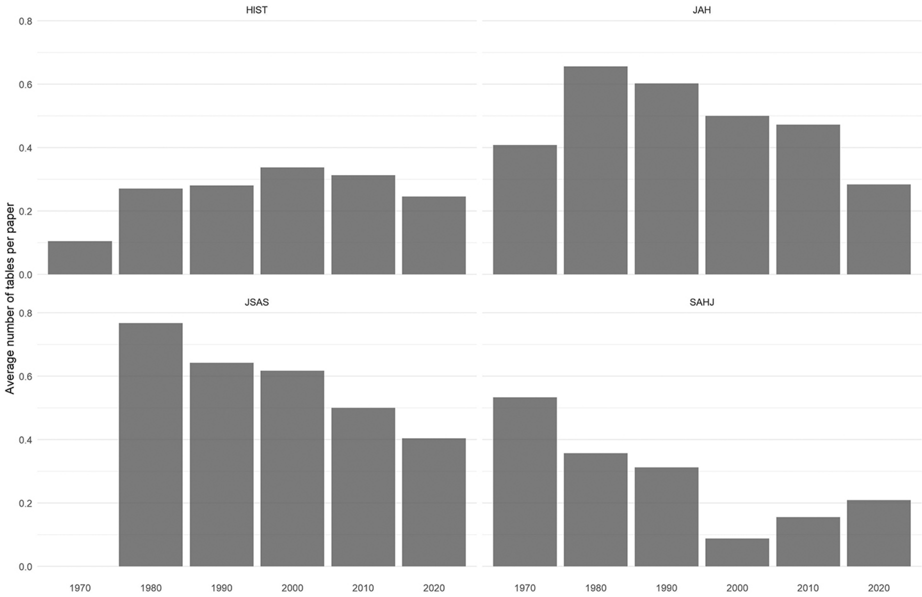

Figure 1.2 Average number of tables per paper, by journal and decade.

second paper, the *SAHJ* and *Historia* do so for one in every four and five, respectively.[77] I found a similar result for graphs. A different split occurs for maps. Here, the *JAH* is far more likely than the other three journals to publish maps, with one in every two papers having maps, compared to one in every five for the *JSAS* and one in every ten for the *SAHJ*. Of all four journals, *Historia* is the most likely to publish images.

I was, however, more interested in the trend over time. Has the number of tables, as a proxy for quantification, declined since the 1970s? Figure 1.2 shows the average number of tables per paper by decade for each of the four journals. Each decade reflects the previous ten years' papers; so, 1990 would include all papers published between 1981 and 1990.

The answer to the above question seems to be a resounding yes. The number of tables per paper declines sharply in the two journals that published the most tables, the *JAH* and *JSAS*. From a high of more than 0.7 tables per *JSAS* paper in the 1970s, the average falls to 0.4 by the 2010s. A steeper decline can be seen for the *JAH* (0.6 to less than 0.3) and *SAHJ* (above 0.5 to below 0.25). It is only *Historia* that, bar the pre-1970s period, maintains the same level, albeit at a very low rate of between 0.25 and 0.3.

[77]It is important to keep in mind that these are averages. It could be that one article includes several tables. Although I report it as one in every five to make it easily comparable, the average is pushed higher by one or two outlier articles. The reality is that the frequency of tables is, in fact, lower across the board.

My findings enabled me to test two hypotheses about the reasons for the decline. One hypothesis was that quantitative papers – defined as those with tables – received fewer citations, on average. Given that journals aim to attract citations, editors were thus less likely to accept those papers for publication. This hypothesis is easily refuted: I found that the 4,476 papers without a table had an average of twenty-one Google citations, while the 740 papers with at least one table had an average of thirty-five Google citations. The difference is statistically meaningful: it would suggest that papers with tables actually attract more citations. When I did an Ordinary Least Squares regression analysis with the number of Google citations as the dependent variable, and several factors, including the number of tables, the number of graphs, the number of maps and the number of images as independent variables, I found a statistically significant and positive coefficient for tables. On average, adding one additional table increases the number of Google citations by two.[78] Editors should prefer papers with tables because they attract more citations.

One interesting observation from the above regression is that multi-authored papers also attract more citations.[79] My second hypothesis was therefore that quantitative history papers are more likely to have more than one author. This would be because quantitative history, as explained above, is more likely to be interdisciplinary, requiring collaboration. The statistics support this conjecture. The 4,724 papers with a single author have, on average, 0.34 tables per paper. By contrast, the 493 papers with more than one author have, on average, 0.74 tables per paper.

Although historians still predominantly publish single-authored papers, there is evidence of more collaboration. Figure 1.3 shows the evolution of multi-authored papers over time. From almost universal single authorship in the early years, dual authorship increased somewhat during the 1990s and 2000s. It is only since the 2010s that papers with three or more authors have become common. Although the rise (albeit from a very low base) of multi-authored papers contrasts with the decline in the number of quantitative history papers, the fact that history journals are largely dominated by single-authored papers perhaps explains why interdisciplinary teams of scholars have looked elsewhere to publish their work.

This five-decade decline in quantification in South African history journals is mirrored by international trends, with one important exception. As Ruggles and Magnuson report, the number of papers per year published in the *Journal of Interdisciplinary History* by

[78] In the regression I also controlled for the number of pages, the issue number, and whether the paper has multiple authors. I also included journal and year dummies. Longer papers tend to attract more citations, the explanation for this probably being that the comments and editorials, that tend to be shorter, receive fewer citations. Articles with maps also tend to attract more citations, although the relationship is weak. The relationship between tables and citations is statistically significant with 99 per cent confidence. The regression output is available from the author on request.

[79] Moving from a single to two or more authors, conditional on all else staying constant, increases the number of citations by 4. This relationship is statistically significant with 90 per cent confidence.

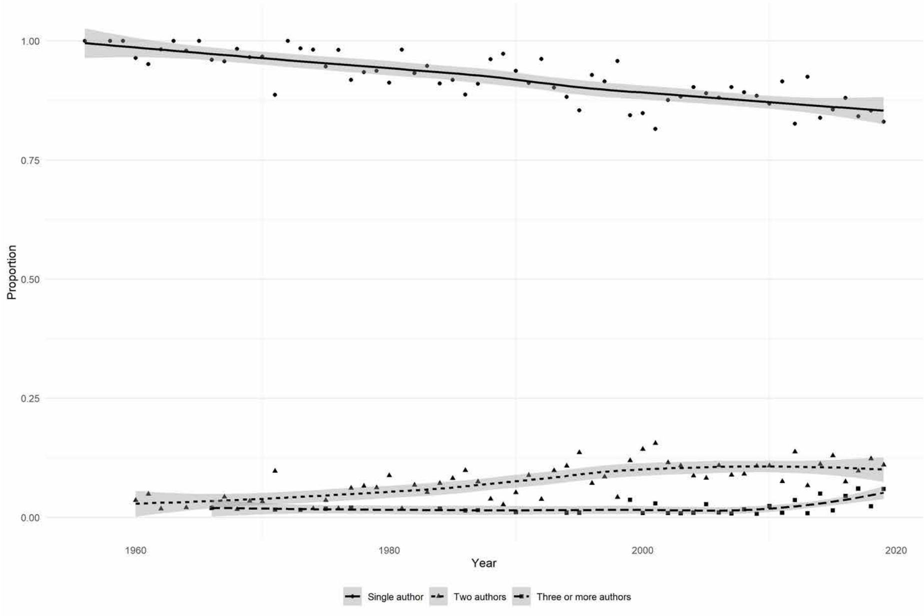

Figure 1.3 Proportion of papers with a single author, two authors or three or more authors, by year.

historians using quantitative methods fell by 72 per cent from the 1970s to the 2000s.[80] The parallel trend for papers in South African history journals is clear. Yet a major shift is currently underway in the leading history journals: the number of papers using quantitative techniques has increased in the past decade by 126 per cent. Although the magnitudes are smaller, this 'revival of quantification in historical research', as Ruggles and Magnuson have called it, is true also for the *American Historical Review*, the *Journal of American History*, the *Journal of Modern History* and *Past & Present*.[81] There is little evidence of this revival in South African history journals.

Look beyond history journals, however, and a more optimistic picture of quantitative history in South Africa emerges. As happened during the earlier wave of quantification, it was in economic history that the first green shoots began to appear. In the late 2000s, an economic history research cluster within the Department of Economics at Stellenbosch University, inspired by what later would be called the 'renaissance in African economic history', began to use historical datasets, notably of the Cape Colony, and analyse them using the economist's toolkit.[82] Several papers were published in some of the leading

[80]Ruggles and Magnuson, 'History of quantification in history'.

[81]Ibid, 379.

[82]This fledgling but dynamic group gained further momentum when Stellenbosch was announced as host of the 2012 World Economic History Congress.

international economic history journals.[83] In 2016, I published a paper in the *Journal of Interdisciplinary History*, reviewing the 'data revolution in African economic history'.[84] The team at Stellenbosch was very much part of that revolution.

By 2015, this research unit formalised into the Laboratory for the Economics of Africa's Past, with the aim of using quantitative methods to investigate the past. From the very beginning, LEAP encouraged interdisciplinary engagement, with historians, sociologists and even agriscientists with an interest in history often joining seminars. This meant that the research focus shifted to investigating questions beyond the confines of economic history – in social, gender, military, business and demographic history. A few examples convey the diversity of topic and audience. In 2018, Mpeta, Fourie and Inwood used the First World War and the Second World War attestation forms – and several other unconventional sources – to reconstruct the living standards of black South Africans over the twentieth century.[85] The paper was published in the *South African Journal of Science*. In 2019, Fourie and Inwood investigated interracial marriages in Cape Town using Anglican Marriage Records.[86] In 2020, Nyika and Fourie used nineteenth-century voters' rolls to investigate the effect of disenfranchisement on black voter registrations. In the same year, Fourie, Inwood and Mariotti published a paper in *Social Science History* which argued that changing military technology may account for the decline in the heights of recruits and, as a result, the apparent decline in living standards.[87] In 2021, Maphosa et al. used limited liability company records to describe the growth of the private capital market at the Cape at the turn of the twentieth century.[88] The paper was published in the *Economic History of Developing Regions*. In the same year, Cilliers and Mariotti used genealogical records to measure nineteenth-century birth spacing, building

[83]W.H. Boshoff and J. Fourie, 'The Significance of the Cape Trade Route to Economic Activity in the Cape Colony: A Medium-term Business Cycle Analysis', *European Review of Economic History* 14, no. 3 (2010): 469–503; J. Fourie and D. Von Fintel, 'The Dynamics of Inequality in a Newly Settled, Pre-industrial Society: The Case of the Cape Colony', *Cliometrica* 4, no. 3 (2010): 229–67; S. Du Plessis and S. Du Plessis, 'Happy in the Service of the Company: The Purchasing Power of VOC Salaries at the Cape in the 18th Century', *Economic History of Developing Regions* 27, no. 1 (2012): 125–49; J. Fourie and D. Von Fintel, 'A History with Evidence: Income Inequality in the Dutch Cape Colony', *Economic History of Developing Regions* 26, no. 1 (2011): 16–48; J. Fourie and D. von Fintel, 'Settler Skills and Colonial Development: The Huguenot Wine-makers in Eighteenth-century Dutch South Africa', *The Economic History Review* 67, no. 4 (2014): 932–63; D. Von Fintel, S. Du Plessis and A. Jansen, 'The Wealth of Cape Colony Widows: Inheritance Laws and Investment Responses Following Male Death in the 17th and 18th Centuries', *Economic History of Developing Regions* 28, no. 1 (2013): 87–108.

[84]Fourie, 'Data Revolution'.

[85]B. Mpeta, J. Fourie and K. Inwood, 'Black Living Standards in South Africa before Democracy: New Evidence from Height', *South African Journal of Science* 114, nos. 1–2 (2018): 1–8.

[86]Chapter 5 makes use of the same dataset to investigate race reclassification in Cape Town.

[87]J. Fourie, K. Inwood and M. Mariotti, 'Military Technology and Sample Selection Bias', *Social Science History* 44, no. 3 (2020): 485–500.

[88]L.M. Maphosa, et al., 'The Growth and Diversity of the Cape Private Capital Market, 1892–1902', *Economic History of Developing Regions* 36, no. 2 (2021): 149–74.

on earlier work that investigated South Africa's fertility transition.[89] Marco-Gracia and Fourie used census data to expose unbalanced sex ratios for infants in twentieth century South Africa.[90] And Ekama et al. published the first descriptive results of a large project that transcribed the full series of valuation rolls compiled during the period of emancipation in the 1830s.[91]

Two projects made possible by generous international funding deserve special attention. One is an ambitious attempt to fully transcribe a series of records of the Dutch and early-British colonial period, from the 1660s to the 1840s. In an attempt to account for the Khoesan in Cape economic history[92], the *Cape of Good Hope Panel* project intends to fully transcribe the Cape *opgaafrolle*, the annual tax censuses of all free households in the Colony.[93] The project, begun in 2015 and since 2021, funded by the Swedish Riksbankens Jubileumfonds, is ongoing.

The other is the *Biography of an Uncharted People* project. Generously funded by the Andrew W. Mellon Foundation, the project has two aims: to transcribe historical administrative records that have so far been largely ignored by historians and to encourage students to equip themselves with the quantitative tools to use these records in their research about marginalised, uncharted people. The choice of which source material to transcribe has been entirely up to the researcher. It has included baptism records, marriage records, voters' rolls, tax censuses, valuation rolls, death notices, company records, petitions to parliament, newspaper advertisements, civil service employee records, attestation forms, surveys, almanacs, probate inventories, auction rolls and mortgage records. Several of these sources form the basis for the chapters published in this volume. The transcribed sources have, where possible, been made available online to encourage wider usage.

This book

This edited collection brings together a selection of work on the *Biography of an Uncharted People* project. The purpose is to show the innovative ways in which quantitative tools and techniques can illuminate the past, shedding light on those that have often been

[89] J. Cilliers and M. Mariotti, 'Stop! Go! What Can We Learn about Family Planning from Birth Timing in Settler South Africa, 1835–1950?' *Demography* 58, no. 3 (2021): 901–25; J. Cilliers and M. Mariotti, 'The Shaping of a Settler Fertility Transition: Eighteenth-and Nineteenth-century South African Demographic History Reconsidered', *European Review of Economic History* 23, no. 4 (2019): 421–45.

[90] F. Marco-Gracia and J. Fourie, 'The Missing Boys: Understanding the Unbalanced Sex Ratio in South Africa, 1894–2011', *Economic History of Developing Regions* 37–2 (2022): 128–146.

[91] K. Ekama, et al., 'When Cape Slavery Ended: Introducing a New Slave Emancipation Dataset', *Explorations in Economic History* 81 (2021): 101390.

[92] J. Fourie and E. Green, 'The Missing People: Accounting for the Productivity of Indigenous Populations in Cape Colonial History', *The Journal of African History* 56, no. 2 (2015): 195–215.

[93] Fourie and Green, 'Building the Cape of Good Hope Panel'.

absent from the historical account. It is quantitative history in the service of those whom history has generally ignored, the uncharted.

Charting those who have been uncharted helps to inform several branches of history. One of the more surprising outcomes of the project has been the contribution of quantitative history to expanding women's history in South Africa.[94] Several of the chapters explicitly focus on women or gender. Richardson and Kok investigate bridal pregnancy in Chapter 2, studying popular attitudes towards premarital sexuality in Cape Town using matches of baptism and marriage records. Their results – of high variance between different Anglican parishes in Cape Town – offer a window into family economics of the time. They find, for example, that 'Cape Town working-class families were less concerned than other socio-economic groups with when and how their children married'.

Chapter 7 by Rommelspacher uses a representative survey to study domestic work in Cape Town. She shows that existing literature frequently makes assumptions that 'every white women was freed from domestic work by the employment of a domestic servant'. The survey reveals, however, that only 31 per cent of white households employed domestic workers, far less than previously thought.

In Chapter 8, Maphosa and Kerby use limited liability company records to expose the extent of participation of women in the Cape's private capital market. They find that female investors 'were diverse, and comprised of spinsters, widows, housewives and professionals'. Without the transcription of thousands of investor transactions, and the analysis of these transactions, we would not have been able to know the surprisingly large role women played as investors.

Just like during the 1960s, demographic history (or historical demography) profits from the renewed interest in quantitative history. Three chapters in this volume make explicit contributions to South African demographic history. Marco-Gracia and Fourie use census records in Chapter 3 to test the mechanisms that explain the surprisingly low infant sex ratios in South African (and African) history. In Chapter 4, Links uses tax censuses to study Khoe kinship networks, concluding that how households are defined has implications for the level of wealth inequality. Jayes and Fourie use death notices in Chapter 12 to calculate the extent of excess mortality during the Spanish flu in South Africa.

Our knowledge of South Africa's political history benefits, in particular, from Chapter 10 by Schoots on African nationalist organisations and from Chapter 11 by Dimitruk and Lemon on petitions to the Cape parliament. Schoots uses network analysis techniques to map the way proto-nationalist African organisations in the late nineteenth-century Eastern Cape were connected to each other through shared members. In doing so, he exposes how traditional leaders 'forged new alliances with mission-educated

[94] A. Rommelspacher, 'Restating the Case for Women's History in South Africa', *Economic History of Developing Regions* 36, no. 3 (2021): 445–50.

leaders, creating organisations which could both mobilise rural communities and make new demands on the colonial state'. It would be these innovations that would later guide the growing African Nationalist movement.

In Chapter 11, Dimitruk and Lemon transcribed petitions to the House of Assembly in Cape Town during the second half of the nineteenth century, documenting patterns in the types of petitions, types of petitioners and their geographic distribution. They then study the relationship between labour issues and temperance, noting the rich historiography of coercive labour laws and alcohol abuse. Dimitruk and Lemon find that both labour and temperance petitioners 'were concerned about the reliability of labour but suggested different legislative solutions to the problem'. They also find that petitioners increasingly formed part of Colony-wide campaigns involving organisations like the Afrikaner Bond and Women's Christian Temperance Union.

Migration remains a pressing issue in the developing world. Chapter 9 by Mbem and De Haas investigates the factors that affected black residents' move to Cape Town in the first half of the twentieth century. The chapter contributes not only to a broader understanding of South African migration history that has largely focused on migration to the mines, but also contributes to women's, demographic and political history.

Several chapters cross disciplinary boundaries too. Aside from its demographic history contribution, Links's study also adds to our understanding of the Cape Colony's economic history. So, too, does Chapter 8 by Maphosa and Kerby on female investors in the Cape. One chapter that makes a more direct contribution to economic history is Chapter 6 by Raaijmakers and Ekama on slave sale advertisements. They transcribed 1,180 newspaper ads published in *De Zuid-Afrikaan* between 1830 and 1834 to describe the reasons for sales being made in the period immediately preceding emancipation. They find, surprisingly, that advertisers increasingly referenced the wishes of the enslaved person to be sold. This, the authors argue, reveals 'not only a veiled agency on the part the enslaved (and responsiveness by slave owners) but also the role of the newspaper connecting town and countryside to facilitate that move'.

Several chapters broadly contribute to social history, although this is most evident in Chapter 5 where Chalmers, Fourie and Inwood investigate how social mobility was affected by race reclassification before and during apartheid. The one unifying theme of the book is, undeniably, the interdisciplinary contribution, in terms of both topics and methods, to writing a history from below.

The eleven chapters are not ordered chronologically. They are, instead, arranged as if following a person's life course: from conception to death. It is, indeed, a biography of an uncharted people. In the final chapter, I explore ways in which more quantitative history can be encouraged, in South Africa and elsewhere. The hope is that this book, and the teaching resources made available as a consequence of the *Biography of an Uncharted People* project, will help to equip a new generation of scholars to tell the stories of those who remain unaccounted for in South African history.

Bibliography

Acemoglu, D., S. Johnson and J.A. Robinson. 'Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution'. *The Quarterly Journal of Economics* 117, no. 4 (2002): 1231–94.

Achebe, N. *The Female King of Colonial Nigeria: Ahebi Ugbabe*. Bloomington: Indiana University Press, 2011.

Ally, R.T. 'The Development of the System of Individual Tenure for Africans: With Special Reference to the Glen Grey Act, c. 1984–1922'. MA diss., Rhodes University, Makhanda, 1985.

Ally, S. 'Material Remains: Artifice Versus Artefact (s) in the Archive of Bantustan Rule'. *Journal of Southern African Studies* 41, no. 5 (2015): 969–89.

Austin, G. and S. Broadberry. 'Introduction: The Renaissance of African Economic History'. *The Economic History Review* 67, no. 4 (2014): 893–906.

Bhattacharya, S. '"History from Below"'. *Social Scientist* 11, no. 4 (1983): 3–20.

Bauman, Z. *Retrotopia*. Cambridge: Polity Press, 2017.

Beales, R.W. 'Studying Literacy at the Community Level: A Research Note'. *The Journal of Interdisciplinary History* 9, no. 1 (1978): 93–102.

Benson, L. *The Concept of Jacksonian Democracy: New York as a Test Case*. Princeton: Princeton University Press, 1961.

Benson, L. 'The Mistransference Fallacy in Explanations of Human Behavior'. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 17, no. 3 (1984): 118–31.

Bogue, A.G. 'The Quest for Numeracy: Data and Methods in American Political History'. *Journal of Interdisciplinary History* 21 (1990): 89–116.

Boshoff, W.H. and J. Fourie. 'The Significance of the Cape Trade Route to Economic Activity in the Cape Colony: A Medium-term Business Cycle Analysis'. *European Review of Economic History* 14, no. 3 (2010): 469–503.

Bozzoli B. and M. Nkotsoe. *Women of Phokeng: Consciousness Life Strategy and Migrancy in South Africa, 1900–1983*. Johannesburg: Ravan, 1991.

Braudel, F. 'Histoire et sciences sociales: la longue durée'. *Annales. Histoire, Sciences Sociales* 13, no. 4 (1958): 725–53.

Breckenridge, K. *Biometric State: The Global Politics of Identification and Surveillance in South Africa, 1850 to the Present*. Cambridge: Cambridge University Press, 2014.

Burke, P. *The French Historical Revolution: The Annales School, 1929–2014*. Stanford: Stanford University Press, 2015.

Carr, E.H. *What Is History?* 1961; repr. London: Penguin Random House, 2018.

Chaturvedi, V., ed. *Mapping Subaltern Studies and the Postcolonial*. London: Verso, 2000.

Cilliers, J. and M. Mariotti. 'The Shaping of a Settler Fertility Transition: Eighteenth-and Nineteenth-century South African Demographic History Reconsidered'. *European Review of Economic History* 23, no. 4 (2019): 421–45.

Cilliers, J. and M. Mariotti. 'Stop! Go! What Can We Learn about Family Planning from Birth Timing in Settler South Africa, 1835–1950?' *Demography* 58, no. 3 (2021): 901–25.

Conrad, A.H. and J.R. Meyer. 'The Economics of Slavery in the Ante Bellum South'. *Journal of Political Economy* 66, no. 2 (1958): 95–130.

Crymble, A. *Technology and the Historian: Transformations in the Digital Age*. Urbana: University of Illinois Press, 2021.

Diebolt, C. and M. Haupert. 'Cliometrics: Past, Present, and Future'. In *Oxford Research Encyclopedia of Economics and Finance*, edited by Dixit, Avinash, Edwards, Sebastian, and Judd, Kenneth. Oxford: Oxford University Press, 2021. Available online: https://doi.org/10.1093/acrefore/9780190625979.013.552.

Dimitruk, K., S. Du Plessis and S. Du Plessis. 'De jure Property Rights and State Capacity: Evidence from Land Specification in the Boer Republics'. *Journal of Institutional Economics* 17, no. 5 (2021): 764–80.

Dooling, W. 'The Making of a Colonial Elite: Property, Family and Landed Stability in the Cape Colony, c. 1750–1834'. *Journal of Southern African Studies* 31, no. 1 (2005): 147–62.

Du Plessis, S. and S. Du Plessis. 'Happy in the Service of the Company: The Purchasing Power of VOC Salaries at the Cape in the 18th Century'. *Economic History of Developing Regions* 27, no. 1 (2012): 125–49.

Eichengreen, B. 'The Contributions of Robert W. Fogel to Economics and Economic History'. *The Scandinavian Journal of Economics* 96, no. 2 (1994): 167–79.

Ekama, K., J. Fourie, H. Heese and L. Martin. 'When Cape Slavery Ended: Introducing a New Slave Emancipation Dataset'. *Explorations in Economic History* 81 (2021): 101390.

Evans, R.J. *In Defence of History*. London: Granta, 1997.

Floud, R. *An Introduction to Quantitative Methods for Historians*. London: Methuen, 1973.

Fogel, R. *Railroads and American Economic Growth*. Baltimore: Johns Hopkins Press, 1965.

Fogel, R. and S. Engerman. *Time on the Cross: The Economics of American Negro Slavery*. Boston: Little, Brown and Co., 1974.

Fourie, J. 'The Remarkable Wealth of the Dutch Cape Colony: Measurements from Eighteenth-century Probate Inventories'. *The Economic History Review* 66, no. 2 (2013): 419–48.

Fourie, J. 'Subverting the Standard View of the Cape Economy: Robert Ross's Cliometric Contribution and the Work It Inspired'. *Magnifying Perspectives: Contributions to History, a Festschrift for Robert Ross, ASC Occasional Publication* 26 (2014): 261–73.

Fourie, J. 'The Quantitative Cape: A Review of the New Historiography of the Dutch Cape Colony'. *South African Historical Journal* 66, no. 1 (2014): 142–68.

Fourie, J. 'The Data Revolution in African Economic History'. *Journal of Interdisciplinary History* 47, no. 2 (2016): 193–212.

Fourie, J. and D. Von Fintel. 'The Dynamics of Inequality in a Newly Settled, Pre-industrial Society: The Case of the Cape Colony'. *Cliometrica* 4, no. 3 (2010): 229–67.

Fourie, J. and D. Von Fintel. 'A History with Evidence: Income Inequality in the Dutch Cape Colony'. *Economic History of Developing Regions* 26, no. 1 (2011): 16–48.

Fourie, J. and D. von Fintel. 'Settler Skills and Colonial Development: The Huguenot Wine-makers in Eighteenth-century Dutch South Africa'. *The Economic History Review* 67, no. 4 (2014): 932–63.

Fourie, J. and E. Green. 'The Missing People: Accounting for the Productivity of Indigenous Populations in Cape Colonial History'. *The Journal of African History* 56, no. 2 (2015): 195–215.

Fourie, J. and E. Green. 'Building the Cape of Good Hope Panel'. *The History of the Family* 23, no. 3 (2018): 493–502.

Fourie, J. and J. Jayes. 'Health Inequality and the 1918 influenza in South Africa'. *World Development* 141 (2021): 105407.

Fourie, J. and S. Schirmer. 'The Future of South African Economic History'. *Economic History of Developing Regions* 27, no. 1 (2012): 114–24.

Fourie, J., K. Inwood and M. Mariotti. 'Military Technology and Sample Selection Bias'. *Social Science History* 44, no. 3 (2020): 485–500.

Fourie, J., F. Ballim, G. Groenewald, J. Upton, T. Nyamunda and J. Parle. 'Making South African Historians Count'. *Historia* 66, no. 1 (2021): 2–38.

Gentry, K. 'Ruskin, Radicalism and Raphael Samuel: Politics, Pedagogy and the Origins of the History Workshop'. *History Workshop Journal* 76, no. 1 (2013): 187–211.

Goldin, C. 'Cliometrics and the Nobel'. *Journal of Economic Perspectives* 9, no. 2 (1995): 191–208.

Guelke, L. and R. Shell. 'An Early Colonial Landed Gentry: Land and Wealth in the Cape Colony 1682–1731'. *Journal of Historical Geography* 9, no. 3 (1983): 265–86.
Heese, H.F. 'Slawegesinne in die Wes-Kaap, 1665–1795'. *Kronos: Journal of Cape History* 4, no. 1 (1981): 38–48.
Hofmeyr, I. *We Spend Our Years as a Tale that Is Told: Oral Historical Narrative in a South African Chiefdom*. Johannesburg: University of Witwatersrand Press, 1993.
Jenkins, K. *Why History? Ethics and Postmodernity*. London: Routledge, 1999.
Laslett, P. 'Size and Structure of the Household in England over Three Centuries'. *Population Studies* 23, no. 2 (1969): 199–223.
Laslett, P. 'Characteristics of the Western Family Considered over Time'. *Journal of Family History* 2, no. 2 (1977): 89–115.
Laslett, P. 'Signifying Nothing: Traditional History, Local History, Statistics and Computing'. *History and Computing* 11, nos. 1 & 2 (1999): 129–33.
Lerner, A.M. 'Quantifying the Archives: Leveraging the Norms and Tools of Data Science to Conduct Ethical Research on the Holocaust'. *Holocaust Studies* 28, no. 3 (2021): 1–19.
Levi, G. 'On Microhistory'. In *New Perspectives on Historical Writing*, edited by P. Burke, 93–113. Cambridge: Polity Press, 1991.
Lemercier, C. and C. Zalc. *Quantitative Methods in the Humanities: An Introduction*. Charlottesville: University of Virginia Press, 2019.
Lemercier, C. and C. Zalc. 'Back to the Sources: Practicing and Teaching Quantitative History in the 2020s'. *Capitalism: A Journal of History and Economics* 2, no. 2 (2021): 473–508.
Marco-Gracia, F. and J. Fourie. 'The Missing Boys: Understanding the Unbalanced Sex Ratio in South Africa, 1894–2011'. *Economic History of Developing Regions* 37–2 (2022): 128–146.
Margo, R.A. and R.H. Steckel. 'Heights of Native-born Whites during the Antebellum Period'. *The Journal of Economic History* 43, no. 1 (1983): 167–74.
Maphosa, L.M., A. Ehlers, J. Fourie and E.M. Kerby. 'The Growth and Diversity of the Cape Private Capital Market, 1892–1902'. *Economic History of Developing Regions* 36, no. 2 (2021): 149–74.
McCloskey, D.N. 'The Achievements of the Cliometric School'. *The Journal of Economic History* 38, no. 1 (1978): 13–28.
Meyer, J.R. and A.H. Conrad. 'Economic Theory, Statistical Inference, and Economic History'. *The Journal of Economic History* 17, no. 4 (1957): 524–44.
Moradi, A. 'Towards an Objective Account of Nutrition and Health in Colonial Kenya: A Study of Stature in African Army Recruits and Civilians, 1880–1980'. *The Journal of Economic History* 69, no. 3 (2009): 719–54.
Mpeta, B., J. Fourie and K. Inwood. 'Black Living Standards in South Africa before Democracy: New Evidence from Height'. *South African Journal of Science* 114, nos. 1–2 (2018): 1–8.
Newton-King, S. *Masters and Servants on the Cape Eastern Frontier, 1760–1803*. Vol. 97. Cambridge: Cambridge University Press, 1999.
Norris, C.J. 'The Negativity Bias, Revisited: Evidence from Neuroscience Measures and an Individual Differences Approach'. *Social Neuroscience* 16, no. 1 (2021): 68–82.
North, D.C. 'The New Economic History after Twenty Years'. *American Behavioural Scientist* 21, no. 2 (1977): 187–200.
Osborn, E.L. *Our New Husbands Are Here: Households, Gender, and Politics in a West African State from the Slave Trade to Colonial Rule*. Athens, US: Ohio University Press, 2011.
Pomeranz, K. *The Great Divergence*. Princeton: Princeton University Press, 2021.
Rijpma, A., J. Cilliers and J. Fourie. 'Record Linkage in the Cape of Good Hope Panel'. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 53, no. 2 (2020): 112–29.
Robinson, J.H. *The New History: Essays Illustrating the Modern Historical Outlook*. New York: The Macmillan Company, 1912.

Rommelspacher, A. 'Restating the Case for Women's History in South Africa'. *Economic History of Developing Regions* 36, no. 3 (2021): 445–50.
Rosenberg, A. *How History Gets Things Wrong: The Neuroscience of Our Addiction to Stories*. Cambridge, US: MIT Press, 2019.
Rosenthal, C. 'Seeking a Quantitative Middle Ground: Reflections on Methods and Opportunities in Economic History'. *Journal of the Early Republic* 36, no. 4 (2016): 659–80.
Ruggles, S. 'The Revival of Quantification: Reflections on Old New Histories'. *Social Science History* 45, no. 1 (2021): 1–25.
Ruggles, S. and D.L. Magnuson. 'The History of Quantification in History: The JIH as a Case Study'. *Journal of Interdisciplinary History* 50, no. 3 (2019): 363–81.
Schlesinger Jr, A. 'The Humanist Looks at Empirical Social Research'. *American Sociological Review* 27, no. 6 (1962): 768–71.
Scott, J.W. *Gender and the Politics of History*. New York: Columbia University Press, 1988.
Scott-Brown, S. 'The Art of the Organiser: Raphael Samuel and the Origins of the History Workshop'. *History of Education* 45, no. 3 (2016): 372–90.
Sellars, E. 'What Are Our Ethical Responsibilities as HPE Scholars?'. *Broadstreet*, 7 February 2022. Available online: https://broadstreet.blog/2022/02/07/what-are-our-ethical-responsibilities-as-hpe-scholars/.
Shell, R. *Children of Bondage: A Social History of the Slave Society at the Cape of Good Hope, 1652–1838*. Hanover, US: Wesleyan/University Press of New England, 1995.
Stoler, A.L. *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*. Princeton: Princeton University Press, 2009.
Stone, L. 'The Revival of Narrative: Reflections on a New Old History'. *Past & Present* 85 (1979): 3–24.
Thompson, E.P. *The Making of the English Working Class*. London: Gollancz, 1963.
Vansina, J.M. *Oral Tradition as History*. Madison: University of Wisconsin Press, 1985.
Von Fintel, D., S. Du Plessis and A. Jansen. 'The Wealth of Cape Colony Widows: Inheritance Laws and Investment Responses Following Male Death in the 17th and 18th Centuries'. *Economic History of Developing Regions* 28, no. 1 (2013): 87–108.
Worden, N. *Slavery in Dutch South Africa*. Vol. 44. Cambridge: Cambridge University Press, 1985.