



ARTICLE

A Microlevel Analysis of Danish Dairy Cooperatives: Opportunities for Large Data in Business History

Paul Sharp 

Sofia Henriques 

Eoin McLaughlin 

Xanthi Tsoukli 

Christian Vedel 

We contribute to the argument for a “new” business history employing a quantitative approach. We illustrate opportunities for new perspectives from this approach using a novel microlevel longitudinal database comprising 131 variables for 1,419 cooperative creameries in Denmark for the period 1898–1945, which we also document and make available to the scholarly community. We present a number of applications of the data, including investigating regional productivity differences, expenditure on fire insurance, and survivorship and reporting biases.

Keywords: Dairying, Denmark, New business history, longitudinal data

Introduction

As with any dynamic field of research, the goals and methods of business history are frequently reflected upon.¹ For example, the focus on big American manufacturers, following the

This project was made possible through generous funding from Independent Research Fund Denmark (DFF – 6109- 00123). Moreover, the digitization required many research assistant hours. Two of the authors of this paper started as research assistants, but we would also like to thank Peter Baunsgaard, Peter Gerner Ejersbo, Nick Ford, Jakob Baltzer Henriksen, Maja Uhre Pedersen, Andreas Slot Ravnholt, and Emil Nørmark Sørensen for extremely conscientious research assistance. Various iterations of this project have been presented at the Cork FRESH Meeting, the 4th HEDG Mini-Workshop for Young Scholars, the 2nd SDU Workshop on Economic and Business History, the 3rd Annual Meeting of the Danish Society for Economic and Social History, two DGPE workshops, and several times at the PhD seminar series of the Department of Business and Economics at the University of Southern Denmark. We would like to thank participants at these events for insightful questions and comments. Finally, we would like to express our appreciation for feedback from the editor and two anonymous reviewers who provided the best of constructive criticism and helped us to position our work more successfully within business history.

This article has been updated since its original publication. See <http://doi.org/10.1017/eso.2023.13>.

1. See e.g., Redlich, “Approaches to Business History”; Hidy, “Business History”; Jones and Friedman, “Business History.”

© The Author(s), 2023. Published by Cambridge University Press on behalf of the Business History Conference. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

pioneering work of Alfred D. Chandler and John Kenneth Galbraith, was questioned by Youssef Cassis.² Business history today has moved beyond the narrow confines of the United States, and big business no longer dominates research, although there are still arguments in favor of a wider focus.³ Building on this, the aim of the present work is twofold. First, we present a new database of Danish dairy cooperatives (“creameries”) 1898–1945, which will allow for a variety of directions for future research due to three qualities: a long time span, a large sample, and in-depth coverage through many variables. Second, the insights for the field of business history, which can be gained from quantitative analyses, provide a motivation for a “new” business history.

Several business historians, most recently Abe de Jong, have argued that case research based on rich primary sources, which dominates the field of business history, could usefully contribute to a theorization of the field, which in turn would allow it to connect more easily to related disciplines, such as economics and economic history.⁴ We are economic historians who have contributed to business history in the past. There is an increasing focus on crossing interdisciplinary boundaries today, and as the theoretical assumptions economics has been based on are increasingly called into question, the core insights from business history have an opportunity to teach important lessons. Indeed, there is a general consensus that economics and economic history today are dominated less by *homo economicus* and more by advanced statistical methods.⁵ Nevertheless, as the Nobel Laureate George Akerlof has so pertinently pointed out, this has meant that economics stands guilty of the “sins of omission,” whereby economic research ignores important problems when they are hard to tackle using its methodologies taken from the “hard” sciences.⁶ Although this is beyond the scope of the present work, one can easily imagine how the qualitative work of business historians might help fill this gap through knowledge of entrepreneurs and entrepreneurship, and how firms act within their political and social context. In short, economic and business history have much to teach each other, and we hope that the data we present here will whet the appetite of business historians as a complement to work more qualitative in focus.

The type of data we present here allows for the consideration of industries made up of many small firms, which typically do not leave behind the sort of comprehensive archives needed

2. Chandler, *Strategy and Structure*; Chandler, *Visible Hand*; Chandler, *Scale and Scope*; Galbraith, *New Industrial State*; Cassis, *Big Business*.

3. Jones and Zeitlin, Introduction to *The Oxford Handbook*; Scranton, “Foundations and Futures.”

4. de Jong, “Research in Business History.”

5. See e.g., Noah Smith, “Data Geeks Are Taking Over Economics,” Bloomberg, August 25, 2016, <https://www.bloomberg.com/opinion/articles/2016-08-25/data-geeks-are-taking-over-economics>; David Sloan Wilson, “Economic Theory Is Dead. Here’s What Will Replace It,” Evonomics, January 12, 2016, <https://evonomics.com/economic-theory-is-dead-heres-what-will-replace-it/>;

Jérémie Cohen-Setton, “The Empirical Shift in Economics,” *Bruegel* (blog), Bruegel, June 15, 2015, <https://www.bruegel.org/blog-post/empirical-shift-economics>. For a dissenting view, see Beatrice Cherrier, “Is There Really an Empirical Turn in Economics?” *Perspectives* (blog), Institute for New Economic Thinking, September 29, 2016, <https://www.ineteconomics.org/perspectives/blog/is-there-really-an-empirical-turn-in-economics>. (All accessed November 29, 2022.) On economic history and econometrics, see Margo, “Integration of Economic History.”

6. Akerlof, “Sins of Omission.”

for case studies. Our database covers 1,419 Danish cooperative creameries over the period 1898–1945 and 131 variables. The rapid spread of industrialized dairying from the 1880s until World War I, based on the use of a new technology—the steam-powered automatic cream separator, a centrifuge—was a defining moment in the development of Denmark, leading to its rapid convergence with the leading economies of the time, and an unusually balanced growth path between town and country.⁷ This case helps to shed light on the business history of cooperation, in which there has been a tendency to ignore cooperatives and other noncapitalist enterprises with a few notable exceptions, such as Patrizia Battilani and Vera Zamagni on Italian cooperative enterprises; John F. Wilson, Anthony Webster, and Rachael Vorberg-Rugh on the Co-operative Group in the UK; Eva Fernández on the diffusion of cooperatives in thirteen countries 1880–1930; and Jordi Planas and Samuel Garrido on cooperative wine making in Spain.⁸

De Jong terms the use of empirical techniques in business history “bizhismetrics.”⁹ We demonstrate how our use of big data allows us, for example, to consider firm-level productivity and the determinants of this, and to examine the issue of survival bias, connecting to the work of Laura Panza, Simon Ville, and David Merritt, who use a survival analysis on firm-level data from the Australian Stock Exchange 1901–1930 to investigate the determinants of longevity of firms, finding that firm size is a poor predictor but that age and profitability are statistically significant.¹⁰ In our case, none of the individual creameries we consider have survived. Waves of mergers over the course of the twentieth century led to a near monopoly for the large Danish-Swedish dairy cooperative Arla. Mads Mordhorst employed a cultural-historical framework, narrative theory, and Pierre Nora’s notion of memory, and argued that the process of globalization has led to Arla being seen as monopolistic and undemocratic, in contrast to the Danish national story of the democratic development of the country, serving to illustrate the importance of our case beyond the confines of economic and business history.¹¹ It should be noted that our database also contains information on, for example, technology, the adoption of which reflects the sort of management decisions of interest to business historians, and we present other potential avenues for future research in the conclusion. Quantitative methods are prevalent in teaching and research in business schools today, as they are seen as important core competencies in decision making. This is also the origin of our database, as data was seen as the foundation of cooperative management decisions—the reason, indeed, that the data was collected and published in the first place by contemporaries.

7. Bjørn, *Dansk mejeribrug 1882–1914*; Henriksen, “Transformation of Danish Agriculture”; Khaustova and Sharp, “Note on Danish Living Standards.”

8. Battilani and Zamagni, “Managerial Transformation”; Wilson, Webster, and Vorberg-Rugh, *Building Co-operation*; Fernández, “Selling Agricultural Products”; Planas, “Emergence of Winemaking”; Garrido, “Cooperatives, Opportunism and Quality Product.”

9. de Jong, “Research in Business History.”

10. Panza, Ville, and Merrett, “Drivers of Firm Longevity.” See also Colli and Vasta’s network analysis of large business groups in Italy, and Van Lieshout, Bennett, and Smith, who use data on workforce numbers from the British censuses of 1851–1881. Colli and Vasta, “Large and Entangled”; Van Lieshout, Bennett, and Smith, “British Business Census.”

11. Mordhorst, “Arla and Danish National Identity.”

The remainder of this paper proceeds as follows. In the following section, we provide our argument for quantitative business history. In section 3, we explain the background to the collection of the creamery statistics and the challenges that were faced, and in section 4 we provide an overview of the database. Section 5 presents some applications of the data, and section 6 concludes.

Toward a Quantitative Business History

During the past twenty-five years or so, there has been an expansion in business history research, which has seen a shift away from the study of individual firms toward a “richer understanding of business systems” and a methodological expansion from qualitative to quantitative research methods. Some have therefore argued for a plurality of methodologies, among them quantitative hypothesis testing along the lines of that used in the “new” economic history, although business history as a discipline has traditionally shied away from this.¹² This mirrors earlier calls for a broader business history encompassing more than single company descriptive case studies toward a more generalizable and analytical discipline, although one that is not isolated from other social sciences.¹³ Thus, there have been calls for business history to move beyond the traditional Chandlerian and post-Chandlerian descriptive case study of core industries and large firms.¹⁴

A special issue of *Enterprise & Society* from 2013 on “How to Do Business History” took up this very debate. Daniel M. G. Raff noted the survival bias evident in much of business history, argued that we should try to understand all participants in a market, and pushed for a focus on the choices and processes behind the various outcomes reached.¹⁵ He noted that “it is easiest to carry out this forward-looking approach to history writing when there is information about a genuine cross-section of the population” but that “this is rarely obtained in business history.” Therefore, a major contribution of the present work is that we present such a cross section. Similar issues were taken up in another special issue of *Enterprise & Society* in 2020, with Raff again arguing for the importance of understanding how decisions are made and Philip Scranton backing this up, noting that, although business history has moved beyond the narrow confines of big business, “outcomes are contingent, unstable, and

12. Eloranta, Ojala, and Valtonen, “Quantitative Methods”; Decker, Kipping, and Wadhvani, “New Business Histories!”; Jones and Friedman, “Business History”; de Jong, Higgins, and van Driel, “Towards a New Business History?”; Jones, “Debating Methodology”; Wilson et al., Introduction to *The Routledge Companion*; de Jong, “Research in Business History.”

13. Redlich, “Approaches to Business History”; Glover, “Comment Approaches to Business History”; Hidy, “Business History.”

14. Kipping, Kurosawa, and Wadhvani, “Revisionist Historiography”; Toms and Wilson, “Business History.”

15. Raff, “How to Do Things with Time.” Responding to this, Popp noted that it is important not to focus on outcomes (e.g., failure/success) but to understand why decisions are made in the first place, and Rosen added that we should remember that externalities such as pollution are not internalized by the firm, so we need to go beyond the individual firm-level decision process. Popp, “Making Choices in Time”; Rosen, “What Is Business History?”

temporary, and ... they are also plainly the wrong place to start an inquiry.¹⁶ It thus becomes crucial to ask questions that locate actors at the outset (or in the middle) of emergent processes whose trajectory is only partly understood and whose success or failure is hardly within their span of control.” Our database is rooted in the microlevel but allows for quantitative work with implications for the macrolevel, covering a period when the dairy sector expanded before World War I, faced economic difficulties during the war and immediately afterward, and then began a gradual process of consolidation. The industry was highly competitive, both on the domestic market and against rivals abroad, and the period marked considerable technological and institutional changes. As Geoffrey Jones and Walter A. Friedman have argued, “Rigorous analysis of large datasets has been shown to transform our understanding of generalizations based on qualitative research.”¹⁷ In fact, the dairy industry in Denmark teaches us that, if business history only focuses on qualitative accounts, then the histories of many smaller firms will be lost; much of the information we have on them survives in the form of the data that we describe below.

The dearth of quantitative research in business history is surprising, given the prevalence of quantitative methods in teaching and research in business schools. Broadly speaking, quantitative methods are widely taught within business schools, as they are seen as key competencies for managers to make informed decisions.¹⁸ Within wider business research, quantitative methodologies have been a dominant strategy in research.¹⁹ This is evident from a recent survey of all empirical articles published in the *Journal of International Business Studies* between 1970 and 2019 that found 87 percent of articles were quantitative in nature.²⁰ Similarly, David Strang and Kyle Siler highlight the methodological diversity within publication trends in *Administrative Science Quarterly* from 1958 to 1970, showing the increased engagement of social scientists with organization studies and how this has led to a wider array of approaches with a strong quantitative element.²¹ As it is widely acknowledged that qualitative and quantitative methods are appropriate for different research questions and contexts, we are not arguing for a superior methodological approach.²² Rather, the relative absence of one is surprising, given the evolution within social sciences and their influence on business research, thus indicating that there could be scope for greater engagement of business historians outside traditional academic silos if a broader methodological range is used.²³

The creation of new ventures is a core theme in management, organizational theory, and strategic management.²⁴ Explaining variation in new ventures, survival and termination and exit, requires longitudinal (repeated measurement of the same subject over time) data, such as the U.S. Longitudinal Business Database.²⁵ Using data from the U.S. LBD, Vincent Sterk, Petr

16. Raff, “Business History”; Scranton, “Foundations and Futures.”

17. Jones and Friedman, “Business History.”

18. Waters, *Quantitative Methods for Business*.

19. Bryman and Bell, *Business Research Methods*.

20. Nielson et al., “Fifty Years of Methodological Trends.”

21. Strang and Siler, “From ‘Just the Facts.’”

22. Lo, Rey-Martí, and Botella-Carrubi, “Research Methods in Business.”

23. Abbott, *Chaos of Disciplines*.

24. For a review, see Shepherd, Souitaris, and Gruber, “Creating New Ventures.”

25. Chow et al., “Redesigning the Longitudinal Business Database.”

Sedláček, and Benjamin Pugsley find that *ex ante* variation in firm characteristics (such as business models) are more important than persistent *ex post* shocks (such as demand shocks) in explaining the longevity of U.S. start-ups.²⁶ Other studies use different longitudinal surveys to assess the influence of entrepreneurial optimism for the establishment of new ventures and the changing risk preferences as ventures age.²⁷ Elsewhere, Panayiotis Georgallis, Glen Dowell, and Rudolphe Durand look at the factors that lead to the state support for new industries, finding this to be greater where the sectors were uncontested.²⁸ Recent work has seen attempts to finesse the research method by separating the likelihood of survival from the magnitude of growth.²⁹

Another avenue of interest is the wide disparity of productivity among businesses. Productivity differences across firms (businesses) has been widely documented, even businesses within the same industry, and this has led to continued use of firm (business) level data to explore this issue in more detail.³⁰ The work of Erik Brynjolfsson and Lorin Hitt is an example of an early study using firm-level data to document heterogeneity in productivity of firms adopting information technology.³¹ Another aspect of this, which has been focused on recently, concerns how firms respond to uncertainty in business decision making. Hitesh Doshi, Praveen Kumar, and Vijay Yerramilli find uncertainty affects firm investment decisions, but this varies by firm size, with smaller firms more affected.³² Alon Kalay, Suresh Nallareddy, and Gil Sadka find a close relation between firm-level performance shocks and economy wide uncertainty and how these two can interact.³³ Finally, Nicholas Bloom et al. show how manager expectations are influenced by historical experience and how these expectations influence investment decision making.³⁴

The Case of the Danish Creameries and the “Operational Statistics”

The case of the emergence of the Danish butter industry is well-known, and a sizeable literature—Danish and international—has been devoted to exploring how a small and relatively backward country, reeling from the loss of most of its territory (Norway in 1814 and Schleswig and Holstein in 1864), managed to create a world-beating dairy industry—one based on well over one thousand small, democratically organized cooperatives of peasant farmers. Denmark became the largest exporter of butter in the world prior to World War I, successfully outcompeting traditional suppliers of the important British market, such as Ireland, and would later become an important provider of know-how to other countries, such

26. Sterk, Sedláček, and Pugsley, “Nature of Firm Growth.”

27. Anglin, McKenny, and Short, “Impact of Collective Optimism”; Wennberg, Delmar, and McKelvie, “Variable Risk Preferences.”

28. Georgallis, Dowell, and Durand, “Shine on Me.”

29. Delmar, Wallin, and Maged Nofal, “Modeling New-Firm Growth.”

30. Syverson, “Product Substitutability”; Syverson, “What Determines Productivity?”

31. Brynjolfsson and Hitt, “Paradox Lost?”

32. Doshi, Kumar, and Yerramilli, “Uncertainty, Capital Investment, and Risk Management.”

33. Kalay, Nallareddy, and Sadka, “Uncertainty and Sectoral Shifts.”

34. Bloom et al., “Business-Level Expectations and Uncertainty.”

as the United States and Russia, the latter of which would go on to become the second largest exporter before the 1917 October Revolution.³⁵ Common to these explanations is an impression that the cooperatives rolled over Denmark in something of a uniform wave, which is often compared to relative failure in other countries. In the present work, as a first application of the database, we use a simple fixed effects setup to demonstrate that there was also heterogeneity between regions of Denmark, something that contests a homogenous linear business history of creameries in Denmark.³⁶

Many reasons have been given for the success of the Danish cooperative creameries. Ingrid Henriksen demonstrates that cooperation was particularly well suited to dairying and that, given the technology of the time, they were ideally suited to overcoming the problems of potential lock-in and asymmetric information, although they were not necessarily technologically “savvy” and often dragged their feet in the implementation of new technologies.³⁷ Denmark was, however, quick to embrace “winter dairying,” which allowed producers to enjoy higher prices at times of the year when more traditional operators were unable to produce.³⁸ The homogeneous population of Denmark, in contrast to the divisions of Ireland, have also been suggested as a reason the cooperation found special favor, although Henriksen et al. 2012 and Henriksen et al. 2015 demonstrate that social cohesion was not enough, and explain that the legal system, in particular the ability to enforce contracts, was also important. Eoin McLaughlin and Paul Sharp demonstrate that a lack of sizeable proprietary competitors also explained Denmark’s relative success compared to Ireland.³⁹ Henriksen et al. 2011 demonstrate that the productivity of the cooperatives in terms of butter production was owed mostly to their rapid adoption of the centrifuge rather than their organizational form, and Markus Lampe and Sharp show that farmers were able to increase milk yields through the introduction of multiple innovations in, for example, breeding and feeding.⁴⁰ Sofia Teives Henriques and Paul Sharp explain that Denmark was also fortunate in having a particular geography with a long coastline, which allowed coal, a vital input, easy and cheap access to the entire country, although this followed centuries of searching for coal in a country where it was practically nonexistent.⁴¹ Finally, Lampe and Sharp, and Nina Boberg-Fazlic et al., have demonstrated the role of traditional landed elites who introduced new practices to Denmark from the eighteenth century, and through continuous innovation laid the basis for the rapid spread of industrial dairying in the final decades of the nineteenth century.⁴²

Understanding the decision-making that allowed for this process means understanding the decisions of hundreds of managers and boards of individual cooperatives, as well as the “knowledge elites” in government, educational institutions, and more. The background of

35. Boberg-Fazlic and Sharp, “Immigrant Communities”; Korchmina and Sharp, “Denmark and Russia.”

36. Controlling for variables that are constant across firms.

37. Henriksen, “Avoiding Lock-In”; Henriksen and Hviid, “Diffusion of New Technology.”

38. Henriksen and O’Rourke, “Incentives, Technology and the Shift.”

39. O’Rourke, “Late 19th Century Denmark”; O’Rourke, “Culture, Conflict and Cooperation”; Henriksen, Lampe, and Sharp, “Strange Birth”; Henriksen, McLaughlin, and Sharp, “Contracts and Cooperation”; McLaughlin and Sharp, “Competition Between Organisational Forms.”

40. Henriksen, Lampe, and Sharp, “Role of Technology”; Lampe and Sharp, “Just Add Milk.”

41. Henriques and Sharp, “Danish Agricultural Revolution”; Ranestad and Sharp, “Success Through Failure?”

42. Lampe and Sharp, *Land of Milk and Butter*; Boberg-Fazlic et al., “Getting to Denmark.”

this decision-making is, however, to a large extent the sort of statistical information we present here, and the editors of the statistics offered advice based on their reading of the information they had collected. In fact, a major contribution of the aforementioned elites was the early adoption of sophisticated agricultural accounting, probably unique in a world perspective at the time, which allowed farmers to make rational and informed decisions regarding both how and what to produce.⁴³ Part of this was the dissemination of knowledge through educational establishments, scientific journals, extension services, and other publications, which allowed for the rapid diffusion of best practice. Although the collection and standardization of accounting material for landed estates had a longer history, the first attempt to do this for the cooperatives was under M. C. Pedersen in 1884, just two years after the first was established in 1882, at the request of the then chair of the Association of Agricultural Societies of Jutland's dairy committee, Frederik Friis, who hoped to be able to demonstrate their inferiority compared to privately owned concerns. In fact, Pedersen's report, which covered just seven cooperative creameries in western Jutland, ended up demonstrating the opposite and thus providing an important boost to the cooperative movement.

The dairy consultant Bernhard Bøggild carried on this work in a number of publications in the 1880s. When creamery associations, representing the cooperatives at a regional level, were established, they took responsibility for this, and various reports were published surveying increasing numbers of creameries for different parts of the country between 1891 and 1897.⁴⁴ In 1897, it was decided to request that all cooperatives submit accounts for publication by a central organization, and it is these "Operational Statistics of Creameries" (*Mejeridriftsstatistik*, MDS) that we have hand collected from the original published volumes and that allow us to construct detailed microlevel longitudinal data over many decades. Our database comprises 1,419 creameries over the period 1898–1945 and up to 131 variables. We are not aware of any other such detailed microlevel database for the major industry of any country covering such a long period, and in such detail. MDS continued to be published until the early 1970s, but increased rationalization of the industry and a lack of comparability to earlier volumes means that we chose to end in 1945. This database provides us with the scale and scope to view the dynamics of the dairying industry over time, revealing key elements of the decision-making process and how factors such as regional geography influenced decisions, even for a small country such as Denmark.

Some background on MDS is necessary in order to understand their representativeness, reliability, and significance. Thus, the first volume of MDS (1897/1899) provides a helpful summary of the battle to establish the Operational Statistics. In the winter of 1897, the creamery associations of Jutland, Funen, and Zealand, which represented the cooperatives in their respective regions, applied for government funding of up to 4,000 kroner from the Ministry of Agriculture to collect statistics from their members.⁴⁵ They received a reply on September 24, 1897, addressed to the chairman of the United Creamery Associations of

43. Lampe and Sharp, "Quest for Useful Knowledge"; Lampe and Sharp, "Accounting for the Wealth."

44. To avoid confusion, we refer to the published volumes of MDS throughout with MDS followed by two years. The first is the year of the data covered in that volume, and the second is the year of publication. MDS, 1897/1899.

45. To provide context, a male urban laborer could expect to earn around 2.6 kroner per day in 1897. Khaustova and Sharp, "Note on Danish Living Standards."

Jutland (samvirkende jydsk Mejeriforening). The ministry explained that they had exchanged letters with the Danish Royal Agricultural Society, an organization founded in the eighteenth century and largely representing traditional landed elites, and had ultimately concluded that they could not proceed without first receiving more details on exactly how the information was to be collected and processed, how much it would cost, and whether it was clear that a large proportion of creameries would be willing to share their accounts. Moreover, they requested that further planning should be in collaboration with the Royal Agricultural Society, which had declared itself willing. The Creamery Associations wrote to the Royal Agricultural Society on October 6, 1897, and held their first meeting with them on October 26 of the same year. There was general agreement about the importance of the project, but disagreement regarding how exactly it should be organized. Ultimately, however, they agreed on the following points: (1) that it should be led by a committee of three members, one from the Royal Agricultural Society, one from the Creamery Associations, and one from the Ministry of Agriculture's dairy consultant; (2) that the annual work plan should be determined at a meeting of delegates, at which each Creamery Association should participate with one representative, and other provincial agricultural organizations and the Danish Association of Dairymen (Dansk Mejeristforening) should also be represented (the idea being that this would facilitate fruitful collaboration); and (3) that the costs should be calculated as at least 6,000 kroner the first year, and the requested support was increased to this amount.

One member subsequently decided to take issue with the first point, and this led to continued negotiations, which were additionally delayed by the poor health of the president of the Royal Agricultural Society, Jørgen Carl la Cour. Under the new president, Count Gustav Wedell Wedellsborg, point 1 was changed so that the committee should consist of four members: the president of the Royal Agricultural Society, its dairy consultant, and two representatives of the Creamery Associations, while the other points were left substantially unchanged. The Ministry agreed to support this initiative with 4,000 kroner for the budget year 1898–1899, and parliament accepted this. This was communicated in a letter from the Ministry dated April 15, 1898, to the United Creamery Associations of Jutland, stating that they would leave the responsibility to the Royal Agricultural Society and the Creamery Associations to perform the task to the extent that the amount offered could cover.

On May 16, 1898, the committee as constituted met for the first time, and it was decided that responsibility for organizing the statistics should mostly rest with the Creamery Associations, so that three committee members would now come from them, and just one from the Royal Agricultural Society, and that point 2 should be removed. A new proposal was sent to the ministry proposing that the 4,000 kroner would be administered by those four members. Half was to be distributed in amounts of 5 kroner per creamery that was a member of the associations, under the condition that each association should contribute the same amount, that they submit accounts that could be used for the Operational Statistics, that a report should only be printed for an association if at least fifteen usable accounts were submitted, and that only two copies of the statistics would be sent to each creamery in each association. In order for the accounts to be usable, creameries were required to use a certain standardized form, covering 365 days and with an accounting year ending sometime between October 1, 1897, and January 1, 1898. The other half of the money would be at the disposal of the committee to cover expert

advice and printing costs, with 1,000 kroner to each activity. The ministry accepted this revised plan on September 27, 1898.⁴⁶

Although it was originally planned that the statistics would be compiled separately for each association, the committee eventually agreed that it made more sense to collect the reports in one volume, which was subsequently published each year with a short introduction. Ten copies were distributed to each Creamery Association, and one was sent to creameries outside the associations, to parliament members, to chairs of agricultural societies, etc.⁴⁷ By 1901, the organization was running an increasing deficit, and it asked the government for an extra 1,500 kroner, but ultimately the MDS were combined with the publication of butter price statistics, which were previously published elsewhere, with government support of 9,000 kroner.⁴⁸ Production was turned over to a Committee for Creamery Statistics (Udvalg for Mejeri-Statistik), consisting of one representative from the Royal Agricultural Society, two from the United Danish Agricultural Associations (De samvirkende danske Landboforeninger), three from the United Danish Dairy Associations (De samvirkende danske Mejeriforeninger), and one from the Association of Danish Dairymen (Dansk Mejeristforening).⁴⁹

Besides the obvious value of the massive amount of data collected, the introduction to each volume provides a wealth of important qualitative information and analysis of the published statistics, touching on issues such as fuel prices, tuberculosis, and war, as well as the more mundane aspects of how to run a creamery based on the results of their statistical analysis. This material provides the qualitative material for understanding decisions made for the industry on the macrolevel, with considerable influence for actors in the individual creameries, and based of course on the statistical information. Of course, these decisions could only be as good as the data collected, and the editors frequently emphasized the importance of submitting accounts, and complained about low compliance rates in the early years, although this improved from around a third in the first years of the twentieth century, to about half by 1910, and two-thirds by the late 1920s, at which point it is stated that “it can be considered completely responsible to use the calculated averages as an expression of the general situation of dairying.”⁵⁰ Although the first submissions of accounts from the individual creameries were patchy and many could not be used, and were thus not published, already in the volume for 1899 improvement is noted, which was greatly facilitated by the publication of standardized accounting books.

From the beginning, the editors encouraged more creameries to submit their accounts, explaining that the comparison would facilitate improvements for the individual creamery, and for the situation of dairying in the country as a whole. One editor provided a particularly pertinent motivation:

It can easily be understood that replying to the given questions gives a very reliable picture of the individual creamery before and now, not to mention the historical and general statistical

46. MDS, 1897/1899.

47. *Ibid.*

48. MDS, 1901/1902.

49. MDS, 1902/1903.

50. MDS, 1929/1930.

interest that the collection of this information has at a time when it is still possible to collect authentic data from the beginning of the cooperatives. The present generation will, quite naturally, not attach great value to this information, since the development is occurring before their eyes; but future descendants will quite surely value it highly.⁵¹

A later editor explained that “accounts are the compass of the creamery. Without them one manages blindly, and if one is to have one’s accounts accounting-wise in order, it is important that they are not misleading, but point in the direction of the greatest possible return under the given conditions.”⁵² Occasionally, editors speculated about the reasons why some creameries did not submit their accounts, for example, because they were embarrassed by bad results, because their accounting system was incompatible with that of the MDS, or because they were simply happy to freeride on the efforts of others, something which was even once described as morally wrong.⁵³ There is no doubt that the entire sector was put under a great deal of pressure to provide accurate accounts, and despite the weaknesses in the data discussed below, we believe that our database can be considered to provide a good—if imperfect—snapshot of the condition of dairying in Denmark over the period covered. If nothing else, our case study illustrates some of the challenges involved for our proposed “big data” approach to business history.

Overview of the Database

The data we decided to collect from MDS can be roughly divided into four categories:

1. Basic characteristics (e.g., name, year of establishment, number of shareholders, etc.)
2. Financial statistics (e.g., amount spent on fire insurance, debt, dividends, etc.)
3. Input/Output (number of cows, amount of milk processed, butter produced, milk/butter ratio, etc.)
4. Technology and energy sources (use of refrigeration, types of fuel used, number of centrifuges, etc.)

Appendix A contains tables with a description of all the variables presented by category, as well as the range of years for which the variable is available (Tables A.1-A.4). Furthermore, appendix B contains summary statistics for all of these variables (Tables B.1-B.4). It should be noted that, partly because of time constraints, we did not collect every piece of published information, and chose to exclude some for reasons of lack of comparability between volumes, or our own subjective lack of interest.

As noted above, the data collected is not without its flaws, in particular during the early years, with the most obvious issue being that not every creamery submitted accounts. We can only speculate, as did the editors of MDS, on the reasons for this, and whether there might be certain characteristics common to those who did not send their accounts. For example,

51. MDS, 1900/1901.

52. MDS, 1910/1911.

53. MDS, 1902/1903; MDS, 1928/1929.

although it might be that “worse” or less efficient creameries were less likely to have kept good accounts, which would have made it possible for them to participate, it might also be the case that the most modern creameries simply had less interest in the project, perhaps because they received useful information from elsewhere, or they were content to free ride, which would certainly have been rational at the level of the creamery. Whatever the case, our analysis below demonstrates quite some heterogeneity in productivity between the individual creameries, which might only have been greater if we had access to the full sample.

In their introductions to each volume of MDS, the editors describe the most important issues of the accounts submitted. Some were, of course, simply not compatible with the need to provide meaningful data for comparison between creameries. For example, the first accounts, which only cover Funen and Jutland, totaled 156 of which only 101 could actually be used, either because the remainder, which were thus not published, did not cover the relevant period, or because of other reasons. Another issue that was raised almost every year was that creameries failed to harmonize their accounting years, although the actual dates covered are recorded in our database. The accounts for cheese production are mentioned as being particularly incomplete for the early years, reflecting the lack of importance attached to this minor part of Danish dairying.⁵⁴ Making the accounts comparable was also a considerable task, although it became easier as standardized accounting forms became more widely used. In the first volume, some associations chose to anonymize their data, and we thus did not include them.⁵⁵ In the MDS for 1899 (1900), very few accounts were received from Maribo, and these were thus merged with Sorø and Præstø, which also included, by mistake, those for Odden (Holbæk). As the accounting forms became more standardized, certain creameries stopped participating for some years, as it took some work to make their accounts compatible, although this issue was only temporary.⁵⁶ Pricing the biproducts of butter production (i.e., buttermilk and skim milk, as well as whey from cheese production) also proved difficult, and the system for this was changed with the MDS for 1916 (1917), meaning that the prices before and after this year are not directly comparable. Finally, the small association in Bornholm did not send accounts for 1926 and 1927 because of changes in the accounting year.

The variables not digitized are within the following approximate categories: classical accounts (incomes, expenses, etc.); cheese statistics; and milk production and milk fat distributed over the months of the year.

A key issue with the construction of large-scale hand-entered databases such as this one is the validation of the data. We have, however, used several techniques to ensure the greatest possible accuracy. First, we used a judgement-based method. From previous work, we have knowledge of realistic values of, for example, the milk/butter ratio, which was the main productivity measure recorded (see also below). Thus, we were able to verify or correct unlikely values. Furthermore, we have visually inspected all observations in plots and checked all unlikely entries. Second, we employed statistical methods, in particular a method centered on outliers in individual variables and a regression-based technique.

54. Although, see Henriksen, Hviid, and Sharp, “Law and Peace.”

55. MDS, 1897/1899.

56. MDS, 1909/1910.

All of these approaches are presented in detail in [appendix C](#), and might be of value to others considering similar exercises.

Some Applications of the Database

We now illustrate the power of our database with three applications: first, we consider the extent of heterogeneity in terms of productivity between the individual creameries; second, we investigate the use of fire insurance as just one example of how one might look at decision-making (risk aversion) at the level of the creamery; and finally, we consider the extent of survivor bias.

Heterogeneity of Productivity

As a first application of the database, we aim here to determine whether the productivity of the individual creameries differed to a significant extent between regions, after controlling for trivial explanations for this—after all, the microlevel adds little detail to our prior knowledge without such heterogeneity. As noted above, a common efficiency measure of the time was the milk/butter ratio (*MB_ratio*), which is the weight of milk necessary to produce one weight-unit of butter.⁵⁷ Thus, in order to document heterogeneity in productivity, we model this ratio nonparametrically, that is, using regional dummies and annual fixed effects:

$$MB_ratio_{it} = \beta_0 + \sum_j \beta_j 1[Region_consistent_{it} = j] + \sum_t \gamma_t 1[Year_{it} = t] + \varepsilon_{it} \quad (1)$$

The set that j iterates over contains all regions (except the excluded region), and the set that t iterates over contains all years. $1[x]$ is an indicator function returning 1 if x is true and 0 otherwise. The reference year is chosen as 1898 (the first year), and Silkeborg is chosen as the reference region, as it is the one with an average *MB_ratio* closest to the overall average. If any element of β is different from 0, then this is evidence that there are regional differences in productivity. Thus, we test the hypothesis $\beta = [0, \dots, 0]$ with a classical F-test to see whether this is the case. If the hypothesis is rejected, then this is direct evidence that levels of productivity differed across creameries in different regions.

Principally, we can think of efficiency of butter production in terms of milk input being affected by two factors: first, the technology used in processing the milk into butter (i.e., the ability to extract as much butterfat from the milk as possible), which depended on a minimum efficient scale of production; and second, the quality of the milk (i.e., the fat or cream content of the milk)—the higher the fat content, the more can be extracted per unit of milk.

Our main specification does not, however, capture possible heterogeneity induced from the differing quality of milk, which might principally be determined by variation in the breed of cows, with varying geographic accessibility.⁵⁸ In order to argue that any heterogeneity observed can be found in the production itself, we need to rule out this source of variation.

57. See e.g., Henriksen, Lampe, and Sharp, “Role of Technology.”

58. See Lampe and Sharp, “Just Add Milk,” for an analysis of cow productivity.

Another determinant of productivity is the scale of production. Appendix D documents how some regions reported creameries of much larger size than others. Taken together, this motivates the following alternative specification:

$$MB_ratio_{it} = \beta_0 + \sum_j \beta_j 1[Region_consistent_{it} = j] + \sum_t \gamma_t 1[Year_{it} = t] + \delta_1 \log(Cows_now_{it}) + \delta_2 Milk_fat_{it} + \varepsilon_{it} \quad (2)$$

Here *Cows_now* is the number of cows supplying the creamery, and *Milk_fat* is the fat content of the milk. We only have information on milk fat content from 1929. For this reason, we show results both with and without *Milk_fat*. We refer to the specification with *Milk_fat* as 2A and the specification without as 2B.

Finally, we might suspect that the regional heterogeneity changes over time, and this motivates the following flexible specification:

$$MB_ratio_{it} = \beta_0 + \sum_{jt} \beta_{jt} 1[Region_consistent_{it} = j \wedge Year_{it} = t] + \delta_1 \log(Cows_now_{it}) + \varepsilon_{it} \quad (3)$$

A problem in using *MB_ratio* as an outcome variable is that some of the heterogeneity will be caused by randomness in the production process, making the ratio an unfit measure of efficiency of the individual creamery. This could be addressed by decomposing the random and inefficiency elements of the heterogeneity in a Stochastic Frontier Analysis (SFA) approach.⁵⁹ However, we are not interested in obtaining a measure of individual productivity of each creamery, and for this reason the *MB_ratio* suffices for our purpose and simplifies the analysis somewhat. Nevertheless, an analysis based on SFA was also performed and yielded similar results. We used the specification of George E. Battese and Timothy J. Coelli, in which explanations of the inefficiency are modeled directly in a maximum-likelihood framework.⁶⁰ However, we encountered the so-called wrong skew problem, which causes the Hessian and thereby all classical standard errors to be undefined, thus motivating our focus on the nonparametric specification outlined above.⁶¹ An alternative method would be to supplement the current analysis with a variance decomposition—for instance, estimating Shapley values, which is an estimate of how much a specific variable (or groups of variables) contribute to explaining the variation in the outcome (here *MB_ratio*). Doing this, we find, depending on the specification, that the region contributes between 3 and 44 percent of the variation in the *MB_ratio*. Such a spread highlights the importance of further work with more complete models in order to understand this variation.

Table 1 presents our regression results from the first two specifications.⁶² The reference region is chosen to match the mean *MB_ratio* for that period, meaning that deviations can be roughly compared to the average. Note that there is considerable regional variation in productivity, as revealed by the very strong F-statistic on the regional dummy parameters. This does not qualitatively change when taking milk quality and production scale into account. It should be noted that, although there is regional variation in creamery size, the parameters on

59. See e.g., Henriksen, Lampe, and Sharp, “Role of Technology”; Lampe and Sharp, “Just Add Milk.”

60. Battese and Coelli, “Model for Technical Inefficiency.”

61. Kumbhakar, Parmeter, and Zelenyuk, “Stochastic Frontier Analysis.”

62. Regression tables are generated with Leifeld, “texreg.”

Table 1. Regression table for specification 1 and 2

	Specification 1	Specification 2A	Specification 2B
(Intercept)	26.463 (0.075)***	41.198 (1.120)***	27.539 (0.267)***
Region Bornholm	-0.164 (0.016)***	-0.073 (0.039)	-0.108 (0.009)***
Region Fyn	-0.304 (0.005)***	-0.054 (0.055)	-0.316 (0.014)***
Region Hobro	0.263 (0.009)***	0.060 (0.020)*	0.255 (0.004)***
Region Holbæk	-0.325 (0.016)***	-0.103 (0.038)*	-0.362 (0.025)***
Region København & Frederiksborg	-0.247 (0.013)***	0.032 (0.020)	-0.206 (0.012)***
Region Maribo	-0.286 (0.004)***	-0.004 (0.037)	-0.230 (0.012)***
Region Midtjydsk	-0.287 (0.008)***	-0.013 (0.023)	-0.284 (0.014)***
Region Randers	0.065 (0.010)***	-0.052 (0.005)***	0.065 (0.009)***
Region Ribe	0.323 (0.002)***	0.111 (0.021)***	0.297 (0.003)***
Region Ringkøbing	0.699 (0.012)***	0.214 (0.035)***	0.685 (0.012)***
Region Sorø, Praestø & Møn	-0.340 (0.004)***	-0.079 (0.035)	-0.345 (0.006)***
Region Sønderjydsk	0.681 (0.023)***	0.188 (0.033)***	0.677 (0.025)***
Region Thisted	0.335 (0.006)***	-0.080 (0.045)	0.445 (0.012)***
Region Vejle	-0.219 (0.002)***	-0.102 (0.017)***	-0.266 (0.003)***
Region Vendsyssel	0.288 (0.023)***	-0.016 (0.007)	0.283 (0.026)***
Region Viborg	0.325 (0.003)***	0.083 (0.031)*	0.334 (0.006)***
Region Aalborg	0.757 (0.010)***	0.070 (0.043)	0.746 (0.013)***
Region Aarhus	-0.351 (0.002)***	-0.027 (0.025)	-0.351 (0.004)***
log(Cows now)		-0.076 (0.017)***	-0.163 (0.036)***
Milk fat		-4.659 (0.310)***	
Regional F	5609979717***	78701263***	1651790***
Year fixed effects	Yes	Yes	Yes
Years	1898-1945	1929-1945	1898-1945
R2	0.766	0.830	0.763
Num. obs.	32951	8525	27517

*** $p < 0.001$;

** $p < 0.01$;

* $p < 0.05$.

'Regional F' is the F-statistic of the null-hypothesis that all regional parameters equal zero. Clustered standard error. P-values based on Satterthwaite-corrected t-statistics

the regional dummies do not change much between specifications 1 and 2B. This is evidence that the regional variation is not only determined by regional heterogeneity in creamery size, and that some other factor must be of importance. The parameters do, however, change a lot when controlling for milk quality.

Because of the relatively large amount of data, it is not hard to provide precise estimates, but these also show that our results are economically significant and consistent with at least one feasible mechanism. Thus, MB ratios are, in general, larger in regions that are in the geographic periphery of Denmark. In Ringkøbing (the worst performer), around 1.05 kg of additional milk is needed for every kg of butter when compared to Aarhus (the best performer). Moreover, 0.24 kg of this difference can be attributed to factors other than milk quality, at least for the years 1929–1945, for which we have data on this.⁶³ Aarhus is the largest city in Jutland, whereas Ringkøbing is in the rural western part of Jutland.

63. One might be concerned that the changing range of years generates this changed difference. Appendix E contains a robustness check with no qualitative difference in the conclusion, with a difference between Aarhus and Ringkøbing of 1.21 kg, when not controlling for Milk fat in 1929–1945.

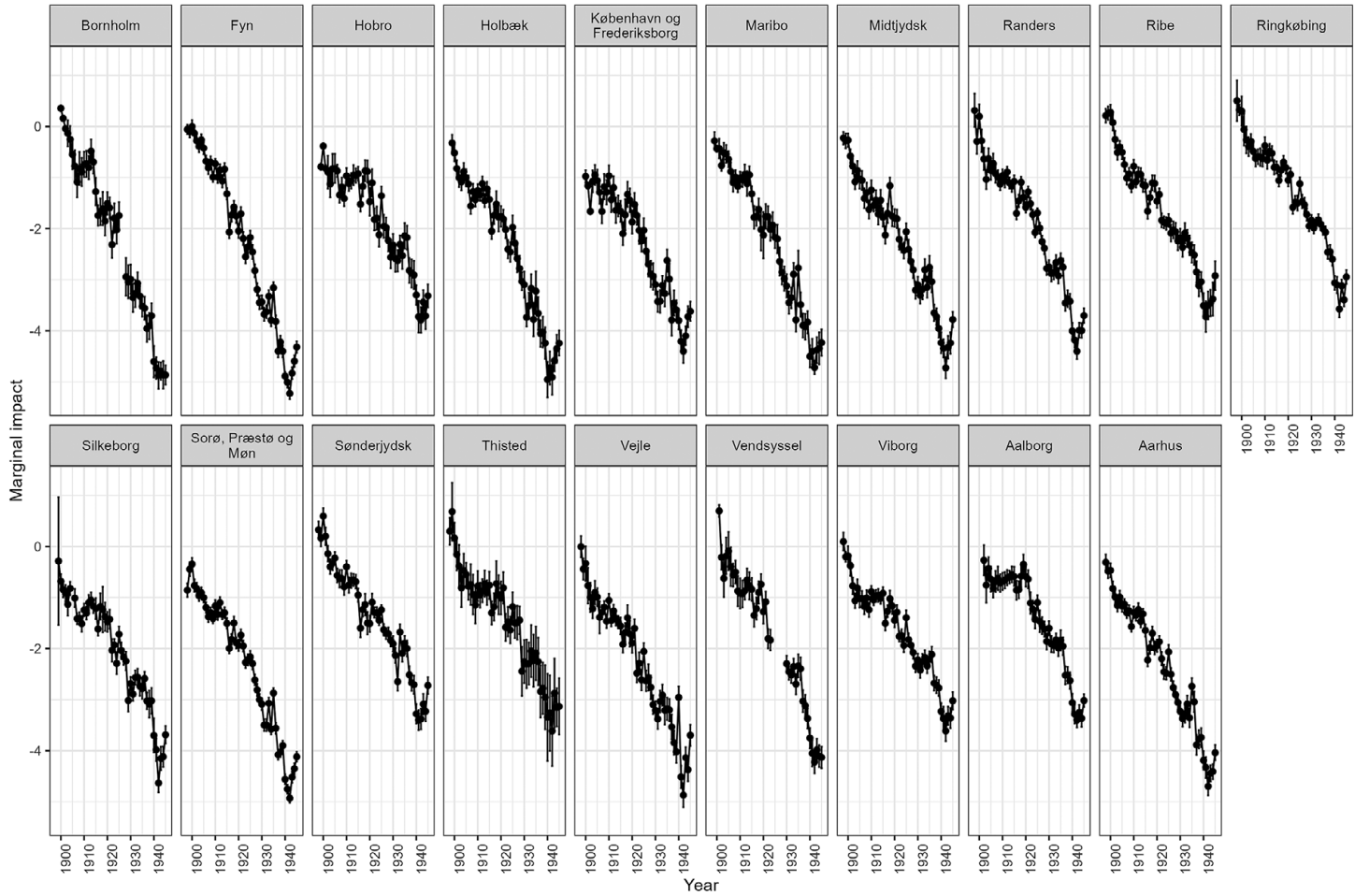


Figure 1. Productivity over time.

Source: MDS. Each of these plots should be interpreted as differences from Silkeborg in 1898 at fixed creamery size measured by $\log(\text{cows_now})$.

Specification 3 contains 890 individual parameters. For this reason, it is simply not feasible to show a regression table for every one of them. However, a figure for each estimate can communicate the information in a far more efficient manner, and [Figure 1](#) contains all the parameter estimates. This shows changes in MB ratio, which are not attributable to the scale of production each year, compared to Silkeborg in 1898, and each point represents a parameter estimate. We have added error bars representing a 95 percent confidence interval to each estimate, but the error is so small that it is hardly noticeable. Overall, it can be noted that productivity increases across all regions over time, and that this development is fairly consistent. However, the different vertical positions of each of the curves demonstrates how it was not at all the same across the different regions. Again, a clear core-periphery pattern emerges.

The data does not describe the entire set of creameries in Denmark at the time. Of particular concern is that missing creameries might be so because of some process that is correlated with regions. The most likely root of problems like this could be that less well-performing creameries might have underreported or that some creameries might even have reported fake positively skewed data and that this issue would vary across regions. The typical MB ratio was well-known and might have provided an incentive for well-performing creameries to submit their impressive reports. On the other hand, underperforming creameries might have reported results that were better than the reality. If this is the case, we expect it to be the case in generally underperforming regions, as the potential signaling gain is greater, and this would thus generate a smaller difference between regions than was actually the case. Consequently, our estimates of regional heterogeneity might be considered a lower bound.

Finally, we have controlled directly for creamery size, but production scale is not a trivial matter. Specifically, there might be external economies of scale. Individual creameries were dependent on a sufficient supply of milk in order to ensure the efficient use of the machinery, given the technology of the time. At the other end of the spectrum, if supplies became too large to handle, they might have needed time to invest in an additional centrifuge.

Fire Insurance

The database contains the nominal amount spent on fire insurance. In mainly coal-powered creameries, fires were frequent, and how best to insure against this was an important decision for the board of a creamery to make. [Figure 2](#) provides a plot of the average amount (in Danish kroner, DKK) spent on fire insurance.

The amount spent on fire insurance roughly follows inflation for the first twenty years, but this gradually increases while inflation falls. To better understand the actual insurance behavior of the creameries, we must take inflation into account. As insurance does not follow the general inflation patterns throughout the period, it might be misleading to simply deflate using the consumer price estimates from Statistics Denmark. Our data offers a natural way forward: We can use a combination of variables, to estimate each creamery's turnover and thereby estimate how much each creamery spent on insurance as a share of their total budget. This gives a sense of how much weight the individual creamery placed on insurance.

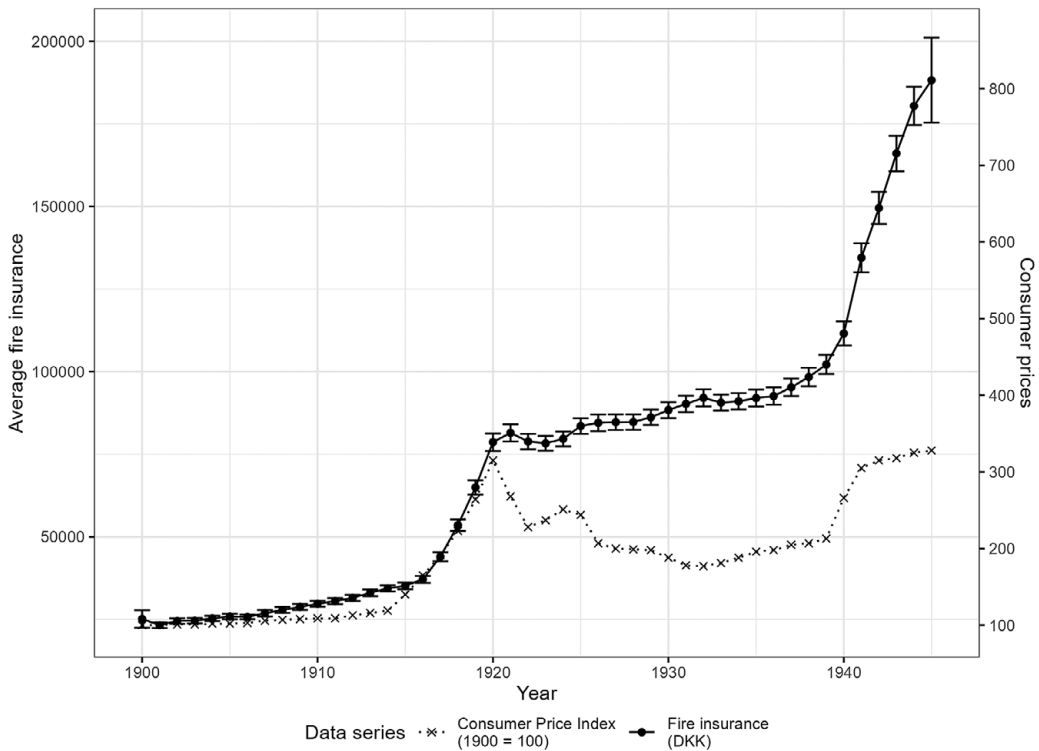


Figure 2. Average amount paid to fire insurance together with consumer price index.

Source: MDS and CPI from Statistics Denmark (data series PRIS8). Error bars represent 95 percent CI based on empirical standard error within each year.

Unfortunately, however, the reports did not contain estimates of total turnover. However, we can calculate that 93.6 percent of all milk was used for butter production.⁶⁴ Therefore, a reasonable proxy for the total turnover is the amount of butter produced multiplied by the price of butter. However, this leaves out an average of 93.6 percent of the production. A simple correction would be to multiply any turnover estimate by $\frac{100\%}{93.6\%}$ to obtain an estimate of the full turnover. However, we can do better than this. Our rich data allows for this estimate to be at the individual level.⁶⁵ The variable *Milk_for_butter_alt* contains the amount of milk used for butter production as can be inferred from the Milk input and the MB ratio.⁶⁶ For each creamery

⁶⁴ We have calculated the amount of milk used for butter production as $Butter = MB_ratio$. For the years 1908–1918, we also have it directly reported in the source as *Milk_for_butter*. Using this, we can estimate that 98.7 percent of milk was used for butter production in the period 1908–1918.

⁶⁵ Another motivation for this approach follows simple economic theory: If the creameries were able to earn more producing something else, they would. Therefore, the price of other goods produced must be equivalent. Is this realistic? Yes, it is sufficiently realistic for this application. Those that produced the data were ultimately motivated by optimizing every aspect of dairy production. See section 3.

⁶⁶ $Milk_for_butter_alt = MB_ratio \times Butter$.

i , in year t we can estimate that they used x percent of milk for butter. We label this the $Butter_share$. Next, we can define

$$\begin{aligned} Turnover\ raw_{it} &= Butter\ price_{it} \times Butter_{it} \\ Turnover\ corrected_{it} &= Butter\ price_{it} \times Butter_{it} \times Butter\ share_{it}^{-1} \end{aligned} \tag{4}$$

This can be reduced to

$$Turnover\ corrected_{it} = Butter\ price_{it} \times \frac{Milk_{it}}{MB\ ratio_{it}} \tag{5}$$

Figure 3 contains turnover and the corrected turnover. The turnover and corrected turnover are almost the same, except for the last few years of the database. This can be explained by creameries in the Copenhagen region starting to produce products other than butter to a much larger extent. Making the same plot by region (not included here) shows that the turnover and corrected turnover only strongly diverge in the København og Frederiksberg region, which covers the greater Copenhagen area.

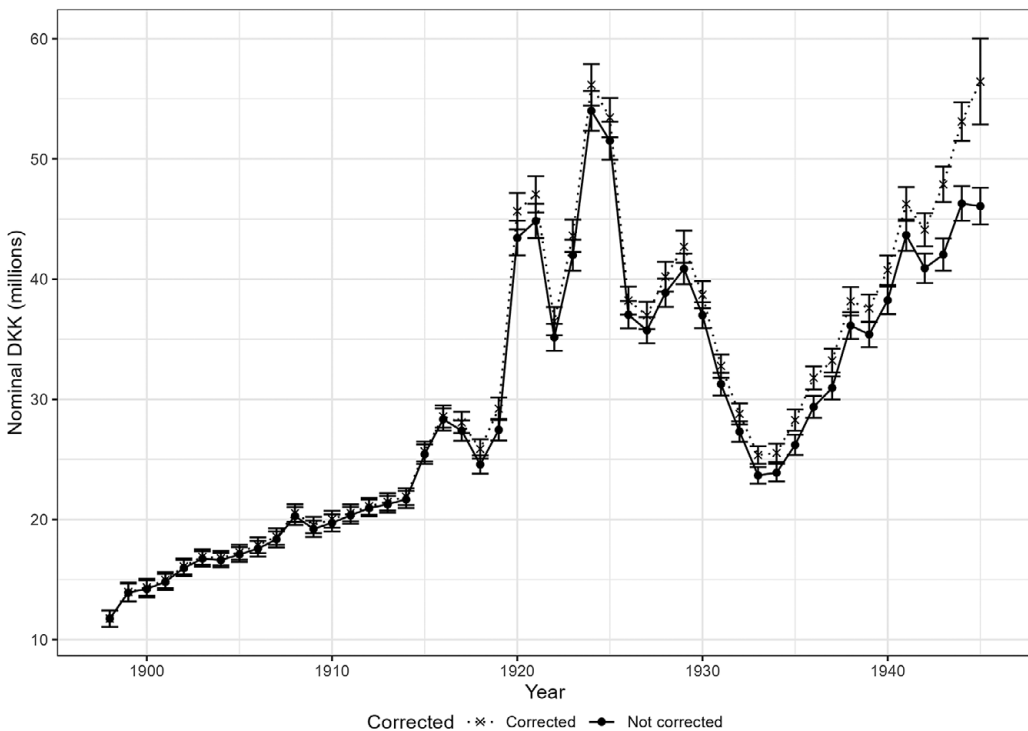


Figure 3. Average turnover and corrected average turnover.

Source: MDS. Error bars represent 95 percent CI based on empirical standard error within each year.

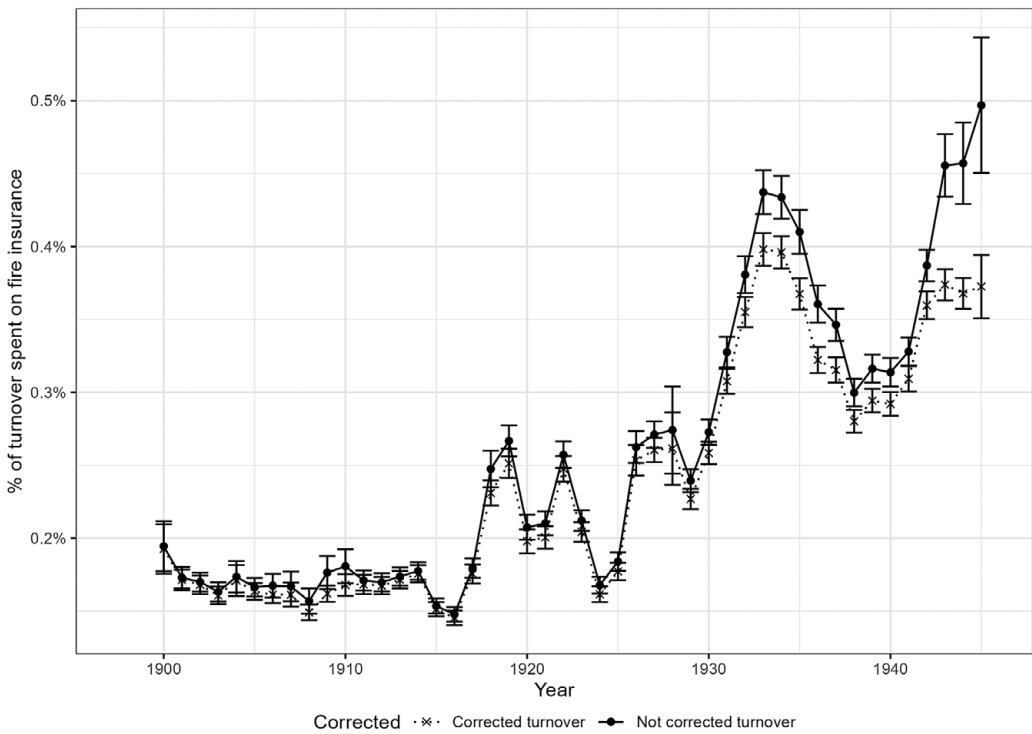


Figure 4. Insurance share over time.

Source: MDS. Error bars represent 95 percent CI based on empirical standard error within each year.

How much was spent on insurance? This is now a simple matter of dividing one by the other:

$$Corrected\ insurance\ share = \frac{Fire\ insurance}{Turnover\ corrected} \tag{6}$$

Figure 4 shows the insurance share of turnover over time.

What explains insurance at the level of the individual creamery? We can apply the classical tools from the econometrics toolbox. One example is the robust negative relationship between size and how much is spent on fire insurance. We estimate the following equation:

$$Cor.insurance\ share_{it} = \ln(Shareholders_{it})\beta_1 + z'_{it}\gamma + \varepsilon_{it} \tag{7}$$

Cor.insurance share is the share of turnover spent on fire insurance using the corrected turnover estimates. Here β_1 represents the extra percentage points spent on fire insurance per shareholder. We scale the shareholders variable to count hundreds of shareholders, as this makes the units more comparable. $z'_{it}\gamma$ contains a set of controls such as fixed effects.

From Table 2 it should be noted that we find a robust negative relationship between size and the share of turnover spent on insurance. Across specifications we find a negative relationship, and for the comparable specifications (1, 2, 3, 6) this relationship is very similar:

Table 2. Regression table for fire insurance

	Cor. insurance share			ΔCor. insurance share	Cor. insurance share	Raw insurance share
	(1) Simple	(2) FE	(3) Only butter producers	(4) First difference	(5) Alternative size	(6) Non-corrected
log(Shareholders_ now)	-0.117 (0.010)***	-0.135 (0.010)***	-0.125 (0.019)***		-0.068 (0.013)***	-0.166 (0.016)***
Δlog(Shareholders)				*** -0.071 (0.014)		
log(Cows now)					-0.093 (0.014)***	
Year FE	No	Yes	Yes	Yes	Yes	Yes
Creamery FE	No	Yes	Yes	Yes	Yes	Yes
R ²	0.139	0.025	0.038	0.003	0.035	0.015
Num. obs.	30027	30027	13734	28095	25994	30220

*** $p < 0.001$;

** $p < 0.01$;

* $p < 0.05$.

Clustered standard error. P-values based on Satterhwaite-corrected t-statistics

Note: (1) Is a simple pooled estimate. (2) Introduces Year and Creamery level fixed effects (3) Includes only creameries using 99% or more of their Milk for butter production (4) Is a first difference estimate. (5) Introduces an alternative size measure: The number of cows associated with the creamery. (6) Uses the raw insurance share instead of the corrected one.

-0.166 percent and -0.117 percent less is spent on insurance when the creameries grow 1 percent in size. Moreover, we see that this relationship is replicated when measuring the size in terms of cows. Last, a similar negative relationship is found for a first difference specification (i.e., when measuring the relationship between a change in insurance share and a change in number of shareholders for the same creamery). This is in line with studies using modern data (see e.g., Mayers and Smith Jr, “On the Corporate Demand”), and a fuller analysis could take into account some (although of course, because of data constraints, not all) of the other factors found relevant in the literature, such as geographic concentration.⁶⁷ Such relationships would be very hard to pick up using classical micro-historical methods of business history. As for causality, we can speculate that creameries that were more risk averse would grow less, and we could speculate that larger creameries were able to negotiate better contracts. The introduction of this industry-wide data allows for further quantitative and/or qualitative research to explore such questions and more.

Survival Analysis

A classical survival analysis would involve a proportional hazards model or similar, and an assessment of whether assumptions for such a model were fulfilled. However, this is beyond what is necessary for the current application, with which we simply want to get a more informed grasp of missing data. We instead focus on a simple adaption of Kaplan-Meier curves.⁶⁸

67. Mayers and Smith Jr, “On the Corporate Demand.”

68. Kaplan and Meier, “Nonparametric Estimation.”

A Kaplan-Meier curve illustrates how many—as a share of the original number of units (in this case creameries)—are still observed at any point in time. This was originally developed in the context of medicine, in which one might be interested in actual survival; however, what we are interested in is seeing how many creameries stop reporting. In the classical medical setting, you cannot have units that are reported dead in one period turning up later. However, this is not the case for MDS, in which creameries might report once, then drop out, and later report again. This motivates a slight adaptation to allow for this. We thus count all the creameries in 1898 and then we see how many of those creameries report in 1899, 1900, 1901, and so on until 1945. This gives us a single pseudo survival curve for the creameries of 1898. We then repeat this exercise for 1899: We count the number of creameries in 1899 and then see how many of them show up later. This process is repeated until we have a survival curve based on the cross section for every single year. All of these survival curves are illustrated in Figure 5, which should be interpreted as the share of units retained after the number of years on the x-axis; for example, after ten years, around 80 percent of the creameries observed in time 0 (1898 or 1899 or ...) are still observed.

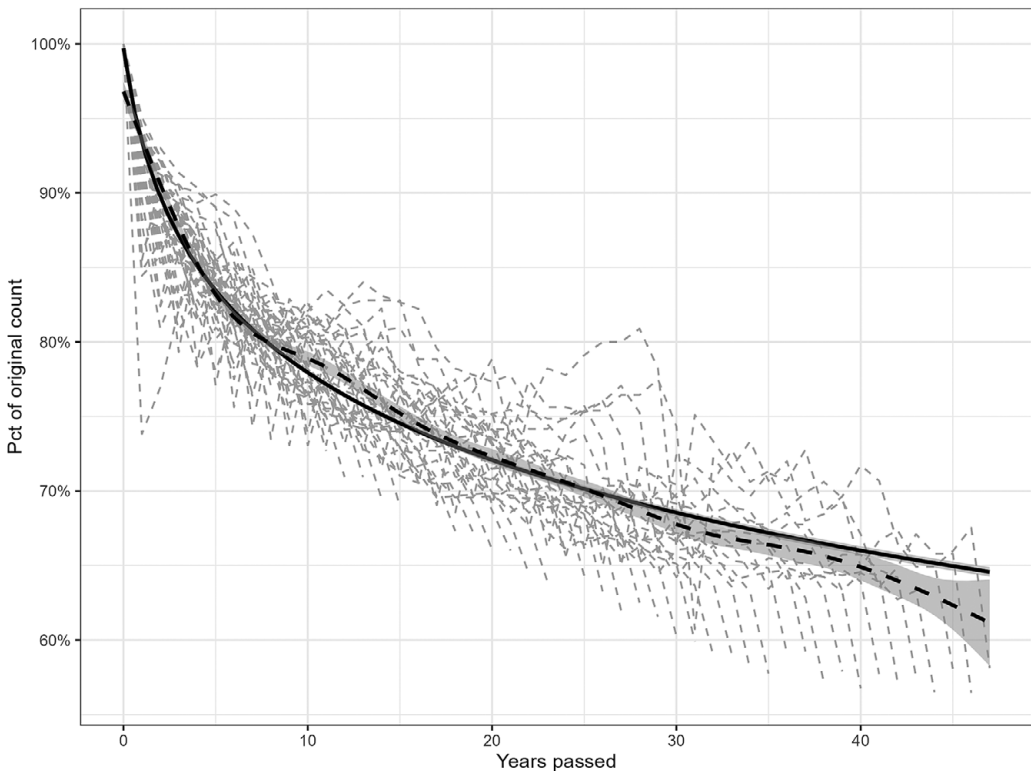


Figure 5. Pseudo Kaplan-Meier curves for creameries in MDS.

Source: MDS.

There is quite a persistent pattern. The dashed curve is a best fit line, with no model structure imposed,⁶⁹ and this follows rather well a curve with a model structure given by the solid line. The model structure is

$$Pct\ of\ original = 1 - \beta \log(\text{Years passed} + 1) + \varepsilon \quad (8)$$

This is something we can estimate, and the β then gives a sense of the rate of loss of observations per percentage increase in time. The estimates we obtain range from -0.06 to -0.11, depending on which survival curve we run it on (we tried with origin in every single year from 1898 to 1940). Therefore, we roughly know that, if we increase the time passed by 1 percent, an additional 6–11 percent of the creameries cease reporting. It is important to note that this is not a formally identified survival analysis. Any such analysis would have to be case specific to the use of the data. However, it does indicate very specifically the type of dropout to expect, which contrasts with the alternative of starting with extant firms and working backward.

For comparison, we can do this counterfactual exercise. What if we started with just a list of creameries that existed in 1945 and then used the common approach of working backward. Given this approach, we would have 36.1 percent of all creameries in 1898 and more than 40 percent for some other years, and it turns out that already ten years prior to 1935 we would have missed 24.8 percent of the creameries by relying on the 1945 list (see the full table in [appendix F](#)). This is hard evidence for the selection problem from which any analysis based on surviving firms might suffer.

Discussion and Conclusion

We reflected on the methodological debate within business history, and argue that business historians, economic historians, economists, and others have much to learn from one another. In part to avoid this being mere platitudes, and in the spirit of de Jong’s “bizhismetrics,” we presented a new microlevel longitudinal database of cooperative creameries in Denmark for the period 1898–1945.⁷⁰ We hope that this provides inspiration for future work within business history: collecting an industry-wide picture that includes the smallest and largest of firms, with less survivor bias, and offering a more quantitative approach to assess drivers of productivity across both time and space. Each individual creamery was very small, employing perhaps three workers, and has left little trace beyond the data we present here. Nevertheless, the industry as a whole was central to Danish economic and business history.

We presented a number of applications of the database, and as one of the first uses of the data, in contrast to the implicit assumption often given in the literature that the creameries were a relatively homogeneous institution, we demonstrate considerable heterogeneity between the regions of Denmark in terms of the productivity as measured by the milk/butter ratio. Our econometric results should, of course, not be given a direct causal interpretation. However, the value of this finding is that it offers an insight into this previously unseen

69. Specifically, a nonparametric estimate using the GAM estimator of the *mgcv* R-package. See Wood, *Generalized Additive Models*.

70. de Jong, “Research in Business History.”

heterogeneity, which might provide inspiration for future research. For now, we note that the productivity of the modern dairies did not spread uniformly, that the regional heterogeneity we observe seems to be robust and attributable to factors other than the quality of the milk input and the production scale. Finally, we observe a clear core-periphery pattern.

Why might this be? Because the technology exhibited economies of scale, anything that restricted the amount of milk available would have an impact on productivity.⁷¹ Thus, one possibility is institutional. Harry Haue documents how the new-pietist Inner Mission movement and its associated pressure to keep creameries closed on Sundays caused a loss of productivity.⁷² Because they were stronger in certain regions, this might be one possibility. Another possibility is geography, and access to fuel in particular. During World War I and well into the 1920s, coal supplies were limited, creating uncertainty for creamery managers. During this period, access to coal might have differed according to location, and moreover, alternatives, principally peat and firewood, were also mostly available in certain locations. Geography might have also played a role in terms of endowments in infrastructure. Because the main export market was the UK, there might have been an advantage for creameries closest to the coast. Finally, geographical spillover could also be considered in this context. To some degree the pattern might be caused by agglomeration and external economies of scale in urban regions. Also, scale might have been a driver of productivity, although greater productivity might have also encouraged expansion, thus driving a spurious correlation. We leave all these potential determinants, and others, to future work, which is now made possible by opening the “black box” of the Danish creameries through the creation of a new database.

PAUL SHARP, is Professor at the Historical Economics and Development Group (HEDG) in the Department of Economics at the University of Southern Denmark (SDU). He is an economic historian working mostly on agricultural history and the economic history of Denmark. Contact information: Department of Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. E-mail: pauls@sam.sdu.dk

SOFIA TEIVES HENRIQUES, is an assistant professor at the Faculty of Economics of the University of Porto, and a researcher at cefUP. She holds a PhD in Economic History from Lund University. Her research lies at the intersection between energy, technological change, long-term economic growth and environmental sustainability with a focus on Portugal, Denmark and Europe. Contact information: Centre for Economics and Finance at University of Porto, Rua Dr Roberto Frias S/N, 4200-464 Porto, Portugal. Email: shenriques@fep.up.pt.

EIN MCLAUGHLIN, is a professor of economics at Edinburgh Business School, Heriot-Watt University Gait 11, The Avenue, Currie EH14 4AS, United Kingdom. Email: e.mclaughlin@hw.ac.uk

XANTHI TSOUKLI, is a postdoctoral research associate at the the Chair of Empirical Economics in the Department of Economics at the University of Bamberg. Xanthi completed her PhD in Economics at the University of Southern Denmark in September 2021, where she was working closely with HEDG – the Historical Economics and Development Group based in the

71. Henriksen, Lampe, and Sharp, “Role of Technology.”

72. Haue, “Mejerikrigen.”

Department of Economics. Her work is primarily concentrated in the fields of applied economics and economic history, with specific focuses on development, gender and labor. Contact information: Room F21/02.08, Feldkirchenstraße 21, 96045 Bamberg, Germany. Email: xanthi.tsoukli@uni-bamberg.de.

CHRISTIAN VEDEL, is a PhD candidate at the Historical Economics and Development Group (HEDG) in the Department of Economics at the University of Southern Denmark (SDU). In his research, he employs techniques from causal inference and machine learning on large historical georeferenced data to expand our understanding of the intersection of economics, history and geography. Contact information: Department of Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. E-mail: christian-vs@sam.sdu.dk

Supplementary Materials

To view supplementary material for this article, please visit <http://doi.org/10.1017/eso.2023.5>.

Bibliography of Works Cited

Books

- Abbott, Andrew. *Chaos of Disciplines*. Chicago: University of Chicago Press, 2001.
- Berle, Adolf A., and Gardiner C. Means. *The Modern Corporation and Private Property*. New Brunswick, NJ: Transaction Publishers, 1932.
- Bjørn, Claus. *Dansk mejeribrug 1882–1914*. Aarhus: De Danske Mejeriers Fællesorganisation, 1982.
- Bryman, Alan, and Emma Bell. *Business Research Methods*. 3rd ed. Oxford: Oxford University Press, 2011.
- Burnham, James. *The Managerial Revolution: What Is Happening in the World?* New York: John Day Company, 1941.
- Cassis, Youssef. *Big Business: The European Experience in the Twentieth Century*. Oxford: Oxford University Press, 1997.
- Chandler, Alfred D. *Scale and Scope: The Dynamics of Industrial Capitalism*. Cambridge, MA: Belknap Press of Harvard University Press, 1990.
- . *Strategy and Structure: Chapters in the History of the Industrial Empire*. Cambridge, MA: Harvard University Press, 1962.
- . *The Visible Hand: The Managerial Revolution in American Business*. Cambridge, MA: Harvard University Press, 1977.
- Galbraith, John Kenneth. *The New Industrial State*. Boston: Houghton Mifflin, 1967.
- Lampe, Markus, and Paul Sharp. *A Land of Milk and Butter: How Elites Created the Modern Danish Dairy Industry*. Chicago: University of Chicago Press, 2018.
- Waters, Donald. *Quantitative Methods for Business*. London: Prentice Hall Financial Times, 2008.
- Wilson, John F., Anthony Webster, and Rachael Vorberg-Rugh. *Building Co-operation: A Business History of the Co-operative Group, 1863–2013*. Oxford: Oxford University Press, 2013.
- Wood, Simon N. *Generalized Additive Models: An Introduction with R*. 2nd ed. Boca Raton, FL: Chapman and Hall/CRC, 2017.

Articles and Chapters in Books

- Akerlof, George A. “Sins of Omission and the Practice of Economics.” *Journal of Economic Literature* 58, no. 2 (2020): 405–418.
- Anglin, Aaron H., and Aaron F. McKenny, and Jeremy C. Short. “The Impact of Collective Optimism on New Venture Creation and Growth: A Social Contagion Perspective.” *Entrepreneurship Theory and Practice* 42, no. 3 (2018): 390–425.
- Battese, George E., and Timothy James Coelli. “A Model for Technical Inefficiency Effects in a Stochastic Frontier Production Function for Panel Data.” *Empirical Economics* 20, no. 2 (1995): 325–332.
- Battilani, Patrizia, and Vera Zamagni. “The Managerial Transformation of Italian Co-operative Enterprises 1946–2010.” *Business History* 54, no. 6 (2012): 964–985.
- Bloom, Nicholas, Steven J. Davis, Lucia Foster, Brian Lucking, Scott Ohlmacher, and Itay Saporta-Eksten. “*Business-Level Expectations and Uncertainty*.” NBER Working Paper no. 28259, National Bureau of Economic Research, Cambridge, MA, 2021.
- Boberg-Fazlic, Nina, and Peter S. Jensen, Markus Lampe, Paul R. Sharp, and Christian Skovsgaard. “*Getting to Denmark: The Role of Elites for Development*.” Discussion Papers on Economics no. 5, University of Southern Denmark, Odense, DK, 2020.
- Boberg-Fazlic, Nina, and Paul Richard Sharp. “*Immigrant Communities and Knowledge Spillovers: Danish-Americans and the Development of the Dairy Industry in the United States*.” CEPR Discussion Paper no. DP13757, Centre for Economic Policy Research, London, 2019.
- Brynjolfsson, Erik, and Lorin Hitt. “Paradox Lost? Firm-Level Evidence on the Returns to Information Systems Spending.” *Management Science* 42, no. 4 (1996): 541–558.
- Chow, Melissa C., Teresa C. Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elizabeth Ruth Perlman, Martha Stinson, and T. Kirk White. “*Redesigning the Longitudinal Business Database*.” NBER Working Paper no. 28839, National Bureau of Economic Research, Cambridge, MA, 2021.
- Colli, Andrea, and Michelangelo Vasta. “Large and Entangled: Italian Business Groups in the Long Run.” *Business History* 57, no. 1 (2015): 64–96.
- Decker, Stephanie, Matthias Kipping, and R. Daniel Wadhvani. “New Business Histories! Plurality in Business History Research Methods.” *Business History* 57, no. 1 (2015): 30–40.
- de Jong, Abe. “Research in Business History: From Theorising to Bizhismetrics.” *Australian Economic History Review* 62, no. 1 (2022): 66–79.
- de Jong, Abe, David Michael Higgins, and Hugo van Driel. “Towards a New Business History?” *Business History* 57, no. 1 (2015): 5–29.
- Delmar, Frederic, Jonas Wallin, and Ahmed Maged Nofal. “Modeling New-Firm Growth and Survival with Panel Data Using Event Magnitude Regression.” *Journal of Business Venturing* 37, no. 5 (2022) <https://doi.org/10.1016/j.jbusvent.2022.106245>.
- Doshi, Hitesh, Praveen Kumar, and Vijay Yerramilli. “Uncertainty, Capital Investment, and Risk Management.” *Management Science* 64, no. 12 (2018): 5769–5786.
- Eloranta, Jari, Jari Ojala, and Heli Valtonen. “Quantitative Methods in Business History: An Impossible Equation?” *Management and Organizational History* 5, no. 1 (2010): 79–107.
- Fernández, Eva. “Selling Agricultural Products: Farmers’ Co-operatives in Production and Marketing, 1880–1930.” *Business History* 56, no. 4 (2014): 547–568.
- Garrido, Samuel. “Cooperatives, Opportunism and Quality Product: Why the Early Spanish Cooperative Wineries Produced Ordinary Wine.” *Business History* 64, no. 1 (2022): 118–133.
- Georgallis, Panayiotis, Glen Dowell, and Rudolphe Durand. “Shine on Me: Industry Coherence and Policy Support for Emerging Industries.” *Administrative Science Quarterly* 64, no. 3 (2019): 503–541.
- Glover, J. D. “Comment: Approaches to Business History.” *Business History Review* 36, no. 1 (1962): 70–75.

- Haue, Harry. "Mejerikrigen - En undersøgelse af de søndagshvilende mejeriers udvikling i Danmark." *Fortid Og Nutid* 1, no. 17 (1978): 359–390. <https://tidsskrift.dk/fortidognutid/article/view/73600>.
- Henriksen, Ingrid. "Avoiding Lock-In: Cooperative Creameries in Denmark, 1882–1903." *European Review of Economic History* 3, no. 1 (1999): 57–78.
- . "The Transformation of Danish Agriculture." In *The Economic Development of Denmark and Norway Since 1870*, edited by Karl G. Persson. Aldershot, UK: Edward Elgar, 1993: 153–178.
- Henriksen, Ingrid, and Morten Hviid. "Diffusion of New Technology and Complementary Best Practice: A Case Study." *European Review of Economic History* 9, no. 3 (2005): 365–397.
- Henriksen, Ingrid, Morten Hviid, and Paul Sharp. "Law and Peace: Contracts and the Success of the Danish Dairy Cooperatives." *Journal of Economic History* 72, no.1 (2012): 197–224.
- Henriksen, Ingrid, Markus Lampe, and Paul Sharp. "The Role of Technology and Institutions for Growth: Danish Creameries in the Late Nineteenth Century." *European Review of Economic History* 15, no. 3 (2011): 475–493.
- . "The Strange Birth of Liberal Denmark: Danish Trade Protection and the Growth of the Dairy Industry Since the Mid-Nineteenth Century." *Economic History Review* 65, no. 2 (2012): 770–788.
- Henriksen, Ingrid, Eoin McLaughlin, and Paul Sharp. "Contracts and Cooperation: The Relative Failure of the Irish Dairy Industry in the Late Nineteenth Century Reconsidered." *European Review of Economic History* 19, no. 4 (2015): 412–431.
- Henriksen, Ingrid, and Kevin H. O'Rourke. "Incentives, Technology and the Shift to Year-Round Dairying in Late Nineteenth-Century Denmark." *Economic History Review* 58, no. 3 (2005): 520–554.
- Henriques, Sofia Teives, and Paul Sharp. "The Danish Agricultural Revolution in an Energy Perspective: A Case of Development with Few Domestic Energy Sources." *Economic History Review* 69, no. 3 (2016): 844–869.
- Hidy, Ralph W. "Business History: Present Status and Future Needs." *Business History Review* 44, no. 4 (1970): 483–497.
- Jones, Geoffrey. "Debating Methodology in Business History." *Business History Review* 91, no. 3 (2017): 443–455.
- Jones, Geoffrey, and Walter A. Friedman. "Business History: Time for Debate." *Business History Review* 85, no. 1 (2011): 1–8.
- Jones, Geoffrey, and Jonathan Zeitlin, eds. Introduction to *The Oxford Handbook of Business History*. Oxford: Oxford University Press, 2008.
- Kalay, Alon, Suresh Nallareddy, and Gil Sadka. "Uncertainty and Sectoral Shifts: The Interaction Between Firm-Level and Aggregate-Level Shocks, and Macroeconomic Activity." *Management Science* 64, no. 1 (2018): 198–214.
- Kaplan, E. L., and Paul Meier. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53, no. 282 (1958): 457–481.
- Khaustova, Ekaterina, and Paul Sharp. "A Note on Danish Living Standards Through Historical Wage Series, 1731–1913." *Journal of European Economic History* 44, no. 3 (2015): 143–172.
- Kipping, Matthias, Takafumi Kurosawa, and R. Daniel Wadhvani. "A Revisionist Historiography of Business History: A Richer Past for a Richer Future." In *The Routledge Companion to Business History*, edited by John F. Wilson, Steven Toms, Abe de Jong, and Emily Buchnea. London: Routledge, 2017: 19–35.
- Korchmina, Elena, and Paul Sharp. "Denmark and Russia: What Can We Learn from the Historical Comparison of Two Great Arctic Agricultural Empires?" *Russian Journal of Economics* (forthcoming 2020).

- Kumbhakar, Subal C., Christopher F. Parmeter, and Valentin Zelenyuk. “Stochastic Frontier Analysis: Foundations and Advances I.” In *Handbook of Production Economics*, edited by Subhash C. Ray, Robert G. Chambers, and Subal C. Kumbhakar, 331–370. Singapore: Springer, 2022.
- Lampe, Markus, and Paul Sharp. “Accounting for the Wealth of Denmark: A Case Study of Smithian Growth Using the Emergence of Modern Accounting in Danish Dairying.” *European Journal of the History of Economic Thought* 26, no. 4 (2019): 659–697.
- . “Just Add Milk: A Productivity Analysis of the Revolutionary Changes in Nineteenth-Century Danish Dairying.” *Economic History Review* 68, no. 4 (2015): 1132–1153.
- . “A Quest for Useful Knowledge: The Early Development of Agricultural Accounting in Denmark and Northern Germany.” *Accounting History Review* 27, no. 1 (2017): 73–99.
- Leifeld, Philip. “texreg: Conversion of Statistical Model Output in R to LaTeX and HTML Tables.” *Journal of Statistical Software* 55, no. 8 (2013): 1–24.
- Lo, Fang-Yi, Andrea Rey-Martí, and Dolores Botella-Carrubi. “Research Methods in Business: Quantitative and Qualitative Comparative Analysis.” *Journal of Business Research* 115, (2020): 221–224.
- Margo, Robert A. “The Integration of Economic History into Economics.” *Cliometrica* 12, no. 3 (2018): 377–406.
- Mayers, David, and Clifford W. Smith Jr. “On the Corporate Demand for Insurance: Evidence from the Reinsurance Market.” *Journal of Business* 63, no. 1 (1990): 19–40.
- McLaughlin, Eoin, and Paul Sharp. “Competition Between Organisational Forms in Danish and Irish Dairying Around the Turn of the Twentieth Century.” *Business History* 63, no. 2 (2019): 1–28.
- Mordhorst, Mads. “Arla and Danish National Identity—Business History as Cultural History.” *Business History* 56, no. 1 (2014): 116–133.
- Nielson, Bo Bernhard, Catherine Welch, Agnieszka Chidlow, Stewart Robert Miller, Roberta Aguzzoli, Emma Gardner, Maria Karafyllia, and Diletta Pegoraro. “Fifty Years of Methodological Trends in JIBS: Why Future IB Research Needs More Triangulation.” *Journal of International Business Studies* 51, no. 9 (2020): 1478–1499.
- O’Rourke, Kevin H. “Culture, Conflict and Cooperation: Irish Dairying Before the Great War.” *Economic Journal* 117, no. 523 (2007): 1357–1379.
- . “Late 19th Century Denmark in an Irish Mirror: Land Tenure, Homogeneity and the Roots of Danish Success.” In *The State of Denmark: Small States, Corporatism and the Varieties of Capitalism*, edited by John L. Campbell, John A. Hall, and Ove K. Pedersen, 159–196. Montreal: McGill-Queen’s University Press, 2006.
- Panza, Laura, Simon Ville, and David Merrett. “The Drivers of Firm Longevity: Age, Size, Profitability and Survivorship of Australian Corporations, 1901–1930.” *Business History* 60, no. 2 (2018): 157–177.
- Planas, Jordi. “The Emergence of Winemaking Cooperatives in Catalonia.” *Business History* 58, no. 2 (2016): 264–282.
- Popp, Andrew. “Making Choices in Time.” *Enterprise & Society* 14, no. 3 (2013): 467–474.
- Raff, Daniel M. G. “Business History and the Problem of Action.” *Enterprise & Society* 21, no. 3 (2020): 561–591.
- . “How to Do Things with Time.” *Enterprise & Society* 4, no. 3 (2013): 435–466.
- Ranestad, Kristin, and Paul Richard Sharp. “Success Through Failure? Four Centuries of Searching for Danish Coal.” EHE Working Paper no. 183, European Historical Economics Society, Vienna, 2020.
- Redlich, Fritz. “Approaches to Business History.” *Business History Review* 36, no. 1 (1962): 61–70.
- Rosen, Christine Meisner. “What Is Business History?” *Enterprise & Society* 14, no. 3 (2013): 475–485.
- Scranton, Philip. “Foundations and Futures: Visions of Business History.” *Enterprise & Society* 21, no. 3 (2020): 638–647.

- Shepherd, Dean A., Vangelis Souitaris, and Marc Gruber. "Creating New Ventures: A Review and Research Agenda." *Journal of Management* 47, no. 1 (2020): 11–42.
- Sterk, Vincent, Petr Sedláček, and Benjamin Pugsley. "The Nature of Firm Growth." *American Economic Review* 111, no. 2 (2021): 547–579.
- Strang, David, and Kyle Siler. "From 'Just the Facts' to 'More Theory and Methods, Please': The Evolution of the Research Article in *Administrative Science Quarterly*, 1956–2008." *Administrative Science Quarterly* 47, no. 4 (2017): 528–555.
- Syverson, Chad. "Product Substitutability and Productivity Dispersion." *Review of Economics and Statistics* 86, no. 2 (2004): 534–550.
- . "What Determines Productivity?" *Journal of Economic Literature* 49, no. 2 (2011): 326–365.
- Toms, Steven, and John F. Wilson. "Business History: Agendas, Historiography, and Debates." In *The Routledge Companion to Business History*, edited by John F. Wilson, Steven Toms, Abe de Jong, and Emily Buchnea. London: Routledge, 2017.
- Van Lieshout, Carry, Robert J. Bennett, and Harry Smith. "The British Business Census of Entrepreneurs and Firm-Size, 1851–1881: New Data for Economic and Business Historians." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 54, no. 3 (2021): 129–150.
- Wennberg, Karl, Frederik Delmar, and Alexander McKelvie. "Variable Risk Preferences in New Firm Growth and Survival." *Journal of Business Venturing* 31, no. 4 (2016): 408–427.
- Wilson, John F., Steven Toms, Abe de Jong, and Emily Buchnea, eds. Introduction to *The Routledge Companion to Business History*. London: Routledge, 2017.

Data Sources

Danmarks Mejeri-Drifts-Statistik, 1897/9–1945/6. The first year refers to the data covered by the data in that volume, and the second year refers to the publication date. All volumes are held in the collection of the National Library of Denmark.

Cite this article: Sharp, Paul, Sofia Henriques, Eoin McLaughlin, Xanthi Tsoukli, and Christian Vedel. "A Microlevel Analysis of Danish Dairy Cooperatives: Opportunities for Large Data in Business History." *Enterprise & Society* (2023): 1–29.