University of Groningen

Output growth in German manufacturing, 1907–1936. A reinterpretation of time-series evidence

Veenstra, Joost

Published in:
Explorations in Economic History

DOI:
[10.1016/j.eeh.2015.03.001](https://doi.org/10.1016/j.eeh.2015.03.001)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Veenstra, J. (2015). Output growth in German manufacturing, 1907–1936. A reinterpretation of time-series evidence. *Explorations in Economic History*, 57(7), 38–49. <https://doi.org/10.1016/j.eeh.2015.03.001>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Output growth in German manufacturing, 1907–1936. A reinterpretation of time-series evidence

Joost Veenstra*

University of Groningen, Groningen, The Netherlands

Received 24 November 2014

Available online 17 March 2015

Abstract

This study uses state space time series analysis to estimate output growth in German manufacturing between 1907 and 1936. In the absence of net-output data, previous research has estimated output change using proxy variables. Various proxy-based output series are available that imply either a 25% German labor-productivity lead over Britain or a parity of performance in 1907. The conventional strategy to deal with this conundrum involves a choice between the series. With the correlation between the different proxies and output unknown, such a choice is inappropriate because it results in the loss of information provided by the rejected proxies. Instead, this study makes full and efficient use of all information by estimating a common component in the various output series. The new estimate of output change implies a German lead over Britain in manufacturing labor productivity of 4.4% in 1907. © 2015 Elsevier Inc. All rights reserved.

Keywords: Time series; state space models; comparative economic history; production; manufacturing; Germany

JEL classification: C320; L600; N140; N640; O140

1. Introduction

The data available to historical studies is frequently limited, especially for periods in the distant past. When a variable is unobserved, its behavior can be proxied by correlates. This strategy is widely applied in the field of economic history. One can think of the years of schooling as a proxy for human capital (e.g., [Goldin, 2001](#)) or installed horse power for technology levels (e.g., [Crafts, 2004](#)). Other, perhaps more daring examples are book production for literacy rates (e.g., [van Zanden and Baten, 2008](#)) or age heaping for numeracy rates (e.g., [Baten et al.,](#)

[2014](#)). Notwithstanding the usefulness of proxy variables, these correlates are too often implicitly, and mistakenly, treated as the true value of the unobserved variable of interest. This creates, first, a false sense of accuracy and, second, confusion when two or more proxies are available that display different dynamics. As it is impossible to measure the strength of the correlation between the variable of interest and the proxy when the former is not observed, only circumstantial evidence remains to assess the latter's accuracy.

In this paper I use state space time series analysis as an alternative method for estimating unobserved variables. I

* Nettelbosje 2, 9747 AE, Groningen, The Netherlands.

E-mail address: j.veenstra@rug.nl.

show that this alternative approach provides a way around the disadvantages discussed above by applying it to the debate on German industrial output growth in the early twentieth century. At stake in this debate is the labor-productivity leadership in Europe shortly before WWI. In most German industries data on net output (hereafter “output” unless expressly stated otherwise) is unobtainable for years prior to the first census of production of 1936. Output proxies, such as production volumes, are often available, however. For some industries two or even three different proxies are on hand. This becomes problematic when they show dissimilar dynamics. As it turns out, choosing between proxies leads to very different estimates of labor productivity in German manufacturing; before WWI Germany either performed on par with Britain or enjoyed a clear lead of about 25%.

The historical implications of such a choice sparked off a debate on the reliability of available proxies (Broadberry and Burhop, 2007, 2008a,b; Ritschl, 2004, 2008). I argue, however, that it is not appropriate to prefer one series over another. The proxies used in the debate can all be linked to output change on the basis of economic theory, but the correlation with the true unobserved change of output is never perfect. This means that all proxies carry information, yet none can be rightfully considered equivalent to true output change. Therefore, choosing one series over the others leads to a loss of information. The debate on German output change focused on the use of two different proxies for output change in metal processing, a large manufacturing industry. Both the wage bill and the value of sales can be related to output change on theoretical grounds. If labor is paid its marginal product and income shares are fixed, changes in the wage bill capture changes in net output. On the other hand, the value of sales equals gross output if the economy is closed and without inventories. These conditions do not apply in reality so that none of the proxies can be expected to capture output perfectly. At the same time, all proxies will correlate with output change and must partly be driven by common dynamics.

A technique is required that determines the shared dynamics between the proxy series. One way of doing this is by taking a simple average of the different output series, a technique used mainly in the national-accounts literature. A famous example is Feinstein’s (1972) ‘compromise estimate’ of U.K. national income between 1857 and 1948, calculated as the arithmetic average of three different national income estimates that are compiled with output, expenditure and income data (see also: Sefton and Weale, 1995). Burhop and Wolff

(2005) pursue the same strategy in an attempt to reconcile different national-income estimates for Germany in the late nineteenth century. The central idea in both studies is that each of the national income estimates ought to be equivalent by accounting restriction, but they deviate by an error from the true data (Weale, 1985). Taking the least-square estimate of the observed series provides a possible solution. If all observed series are unbiased estimates of the true data and they are equally reliable, then the simple average presents the least-square solution (Solomou and Weale, 1991).

While this may be a “pragmatic solution”, as Burhop and Wolff (2005, p. 635) call it, this strategy does not work in the case of the German debate on industrial output growth. None of the proxies used is an unbiased estimate of true output change and the accounting restriction that all estimates should be equivalent does not apply. What is needed is a method that relates the unobserved true output change to the observed proxy series by means other than assuming an accounting identity. For this I use state space analysis, the purpose of which is to infer the relevant properties of an unobserved variable from a knowledge of observed series (Durbin and Koopman, 2001). Because the state-space form provides a structural time series model, set up in terms of components which have a direct interpretation, it is very flexible and can handle a wide range of time series problems (Harvey, 1989). In the state-space form it is possible to model the observed German output series jointly, explicitly correlating the dynamics across the different series, which is appropriate given that they are subject to the same influence, namely output change in German industry.

My approach brings together two recent strands of research in the field of economic history. First, I build upon a literature that has used univariate state space time series models to estimate unobserved variables. Examples are studies on the interaction between economic and demographic variables, such as the work of Lee and Anderson (2002) and Crafts and Mills (2009) on Malthusian mechanisms in English history. Similar work for Germany is done by Pfister et al. (2012). Each of these studies uses the state-space form to estimate the dynamics of unobserved variables, such as technological change, the demand for labor or weather and disease prevalence. Second, my approach is closely linked to multivariate dynamic factor models that estimate co-movement across various observed series. This research avenue has been explored by Sarferaz and Uebele (2009) and Uebele (2011) in an attempt to track down business-cycle movements in pre-WWI Germany. To the best of my knowledge, this

study is the first in the field of economic history to combine the two by using the state-space form in a multivariate setting for estimating common components.

A second break with previous work is that I look at point as well as interval estimates. So far the debate has focused on point estimates of output only. Because most estimates in economic history are inaccurate to some extent, [Feinstein and Thomas \(2002\)](#) argued that any new statistical series should be accompanied by a guide to the associated margins of error. They observe that an indication of measurement error provides a criterion to assess the difference between available estimates. In response to their plea, I use confidence intervals to statically assess dissimilarities between my time series estimates and estimates previously presented in the literature. Without an indication of measurement error, the reliability of the proxy-based output series has previously been evaluated by comparison to benchmark estimates. The desired fit between both measures has been set at a 10% difference between both measures, but without statistical justification ([Broadberry and Burhop, 2007](#); [Broadberry, 1993](#)). The interval estimates presented here provide a statistical analysis of the difference between time series and benchmarks.

My new time-series estimate of output change in German manufacturing leads to output levels for the pre-WWI period that are significantly different from previous time-series estimates. Using my new output series to calculate German labor productivity levels in manufacturing leads to a performance level in 1907 that is 4.4% higher than that in British manufacturing. Taking into account the measurement error in my estimates, I can reconcile my time series projections with the benchmark of [Broadberry and Burhop \(2008a,b\)](#), but not of [Ritschl \(2008\)](#). This supports the conclusion that while Germany's lead over Britain in 1907 is statistically significant, the size of the gap was still quite small. At the same time, this finding underlines the German success in catching-up with Britain during the last decades of the nineteenth century. Moreover, in the years after 1907 until the outbreak of WWI this lead was extended to a margin of about 15%, which shows that Germany's development did not peter out after reaching British levels of labor productivity.

This paper argues that the debate on German industrial output change can be moved forward without presenting new data, yet it ought to be underlined that the analysis here is not an alternative to archival work. There is nothing in the above that attempts replacing the search for new archival material. In contrast, as at least two rivaling estimates are needed for this kind of exercise, it puts heavy demands on the data. The point of the paper is to explore a different way of interpreting the available data, such that

new figures are no longer seen as to replace old estimates. The state-space form offers a new way of combining new and old evidence. Figures drawn from historical documents can now be seen not to be either right or wrong, but may both be treated as carrying information. This makes the analysis introduced here complementary to the search for new historical sources.

The remainder of this paper is organized as follows. The debate on German output growth is described in greater detail first. The section thereafter introduces the state-space form and specifies the model used to study output growth in German industry. In the subsequent discussion of the results focus goes out to the new estimate of output growth. The second to last section describes the implications of the new output series for the debate on comparative German/British labor productivity before WWI. Lastly, in the conclusion I indicate what general methodological lessons can be drawn from this new approach to output growth in German industry.

2. The time-series debate

The reliability of the industrial output index estimated by [Hoffmann \(1965\)](#) in the reconstruction of the German Historical National Accounts has been called into question ([Fremdling, 1988, 2007](#)). [Hoffmann's \(1965\)](#) index fails to precisely capture the change in industrial output for various reasons. First, the data does not fully cover industrial output and essentially provides a sample of the output distribution. Second, there is an inconsistent use of various output proxies to measure output change. Third, to construct a compound series for total industry output shares of industries are necessary, but not available prior to 1936.

With respect to the measurement of output, the ideal measure is net output expressed in constant prices. In the absence of net-output data for the period prior to the first German census of industrial production of 1936, [Hoffmann \(1965\)](#) relies on a mix of net-output proxies, such as production volumes, gross output values and labor-income data. In particular the latter proxy has been contested in the literature. [Hoffmann \(1965\)](#) proxies output change in metal processing – a large manufacturing industry that contains machine building, shipbuilding and electrical engineering – by the change in the wage bill. This is an accurate proxy when the wage–productivity ratio remained roughly constant over time and income shares were fixed. However, [Borchardt \(1979\)](#) argued that after WWI wages rose as a consequence of labor unions' increased bargaining power, rather than due to labor productivity changes ([Ritschl, 1990](#); [Broadberry and Ritschl, 1995](#)). This would lead to a

clear upward bias in Hoffmann's (1965) output estimates for metal processing.

Ritschl (2004) observed that while Hoffmann (1965) used the wage bill as a proxy for output in metal processing, other, possibly less problematic, proxies were readily available. Production indexes of various metal processing industries were presented by Wagenführ (1933) of the *Institut für Konjunkturforschung* already during the interwar period. When Hoffmann's (1965) metal processing time series is compared to these official production data, deviations in output growth are apparent mainly for the machine-building industry. For this reason Ritschl (2004) used sales value data of the *Verband Deutscher Maschinen- und Anlagenbau* (the German machinery producers' association) to reassess output change in the machine-building industry and he records an output growth only half as large as suggested by Hoffmann.

Ritschl (2004) also adjusted the output shares used by Hoffmann (1965) for aggregating the industry series to a total for industry. To construct output shares Hoffmann (1965) multiplied labor-productivity levels in 1936 by employment in earlier years derived from the occupation censuses of 1933 and 1907. This assumes that relative levels of labor productivity across German industries remained unchanged over the period 1870–1938, which was not the case. Although a lack of data forced Ritschl (2004) to adopt the same procedure, he corrects the 1907 employment figures for Germany's territorial losses after WWI. Ritschl's (2004) changes to the output index for metal processing and the output shares both reduce the rate of output growth in German industry considerably, particularly for the period 1913–1925. The difference between the two series is clearly visible in Fig. 1 with Ritschl (2004) indicating an output level 13% lower than Hoffmann (1965) in 1925.

Ritschl's (2004) proposed changes to the index of output growth in German industry have implications for labor-productivity growth. A faster rate of output growth will, ceteris paribus, increase labor-productivity growth likewise. Combining Ritschl's (2004) output index with Hoffmann's (1965) employment series yields the change in labor productivity. The labor-productivity index can then be extrapolated from 1936 to earlier years. 1936 is the year of the first German industrial production census, which allows a precise measurement of labor-productivity levels for that year. Extrapolating the labor-productivity index based on Ritschl's (2004) output series from 1936 back to earlier years leads to an upward adjustment of German performance levels in the pre-WWI period of about 25%.

To check the accuracy of the time-series extrapolations, Broadberry and Burhop (2007) constructed a benchmark

of comparative German/British labor productivity in 1907 (see Table 1). The benchmark points at an equality in performance rather than a distinct German lead; a result seemingly at odds with the extrapolated labor-productivity levels obtained on the basis of Ritschl's (2004) output series. Hoffmann's (1965) original index of output, on the other hand, aligns nicely with the benchmark results, which Broadberry and Burhop (2007) interpret as proof of superior quality.

Because Ritschl's (2004) revised time-series estimate is rejected on the basis of the reconciliation between time series and benchmarks, the intrinsic quality of the modified index remains largely undiscussed in Broadberry and Burhop (2007). Consequently, Ritschl (2008) observed a trade-off in Broadberry and Burhop (2007) between the quality of the employed time series and the quality of its fit with the benchmark estimates. Therefore, Ritschl (2008) adhered to his revised output index. He did, however, accept the notion that the fit between benchmarks and time series provides a measure of the latter's accuracy. Ritschl (2008) consequently rejected Broadberry and Burhop's (2007) 1907 benchmark and modified it, which leads to results in line with his time-series extrapolations.

In the latest contribution to the debate Broadberry and Burhop (2008a,b) in turn rejected Ritschl's (2008) changes to their 1907 benchmark, but accepted an improved version of Ritschl's (2008) output index. This does not lead to reconciliation problems when Broadberry and Burhop (2008a,b) adopt a new employment series of Rainer Fremdling (2007) to backcast comparative labor productivity. The new employment figures bring the time-series estimate of comparative German/British labor productivity in 1907 down to 112.4%. They also presented a slightly upward adjusted version of their original benchmark for 1907 that indicates a German performance of 108.4% the level of Britain (see Table 1). The difference between their latest extrapolation results

Fig. 1. Time series of output in German manufacturing (1913 = 100).

Table 1

Benchmark estimates of comparative German/British labor productivity in 1907.

	Industry	Manufacturing
Broadberry and Burhop (2007)	101.8	105.0
Ritschl (2008)	124.5	128.0
Broadberry and Burhop (2008a,b)	104.7	108.4

Sources: see text.

and benchmark revision falls well within the 10% range that Broadberry and Burhop (2007); Broadberry(1993) use as a criterion for the fit between both measures. It should be noted, however, that while this goodness-of-fit measure is precisely defined, it is also without statistical justification.

The debate has been fueled largely by two conceptual issues. First, the output series are assumed to accurately measure the change in industrial production. However, as I have argued, output is not actually observed and its behavior is derived from proxy variables. Choosing between these proxy-based output series is inappropriate because none perfectly captures the true change in output. On the other hand, the different proxies are all clearly correlated with output. Therefore, the dynamic properties of the various proxy-based series must partly be the same, even if the strength of the correlation with output change differs. The appropriate course of action does not involve a choice between output proxies, as this would only lead to a loss of the information contained by the rejected series. Rather, an analysis is required that estimates the common component shared by the proxy-based output series.

A second reason why it proved difficult to reach consensus in the debate is the absence of a tool to assess the accuracy of time series. Without an indication of measurement error in the time series, only circumstantial evidence remains. The fit between time series and benchmark estimates is circumstantial evidence of the former's accuracy, because the literature has identified several causes for deviation between both measures (Summers and Heston, 1991; Heston et al., 2001; Deaton and Heston, 2010). For one, discrepancies can stem from methodological differences between time series and benchmarks. At the root of this inconsistency lies the difference between weight structures employed in bilateral benchmarks and time series (de Jong and Woltjer, 2009; Szilágyi, 1984). A further source of deviation stems from the coverage of industries or products, which frequently differs between both measures. So a perfect fit cannot be expected nor demanded.

What is needed then is a common-component analysis that makes full and efficient use of all available information and provides a measure of error in the estimates. Because

the purpose of state space time series analysis is to uncover the dynamic evolution of observations measured over time when the dynamic properties cannot be directly observed from the data, it provides an appropriate tool of analysis.

3. Methodology

With state space time series analysis I estimate the unobserved dynamic properties commonly shared by the various time series estimates of output change in German manufacturing. Apart from the use of different output proxies in metal processing, the dissimilarities between the time series of output change stem from two other sources. First, the lack of full data coverage in all industries, which subjects the output series to sampling error. Second, the use of different output shares to aggregate the industry-level data to total manufacturing. The analysis is therefore applied to the level of total manufacturing, and not to metal processing only. This allows me to address these three issues simultaneously, thus making full and efficient use of all information.¹

In this study I follow Harvey and Chung (2000), who use a bivariate model with common components based on the state-space form to improve the estimate of unemployment in Britain for the period 1984–1998. A first step in Harvey and Chung (2000) is the identification of a target series, which is the observed time series from which the unobserved variable of interest is estimated. The estimate based on the target series is then improved by also taking into account the dynamic properties of other series. In my case the target series is the output index of Broadberry and Burhop (2008a,b). This most recent contribution to the debate is essentially an update of Ritschl's (2004) output series and thus includes the latest insights. Subsequently, I estimate the unobserved change in output from the target series while also taking into account the dynamic properties of Hoffmann's (1965) and Wagenführ's (1933) output series.

Following the specifications of Harvey and Chung (2000, p. 315), I formulate a multivariate local linear trend model:

$$y_t = \mu_t + \varepsilon_t, \quad \varepsilon_t \sim \text{NID}(0, \Sigma_\varepsilon), \quad (1)$$

$$\mu_t = \mu_{t-1} + \beta_t + \xi_t, \quad \xi_t \sim \text{NID}(0, \Sigma_\xi), \quad (2)$$

$$\beta_t = \beta_{t-1} + \zeta_t, \quad \zeta_t \sim \text{NID}(0, \Sigma_\zeta), \quad (3)$$

¹ Although it is impossible to establish the accuracy of the German output index estimated in this paper, because the true values of output are unobserved, with Monte Carlo methods it can be shown that for the type of problem studied here the state-space form does indeed provide a more accurate estimate than those derived with methods explored before in this debate. See Online Appendix B for more detail.

where y_t and all other vectors are of 3×1 dimension. The model contains three equations. The observed series y_t – i.e. the output series of Broadberry and Burhop (2008a,b), Hoffmann (1965) and Wagenführ (1933) – are modeled by the measurement (or observation) Eq. (1). The observed series are modeled as a function of an unobserved dynamic process called the state μ_t plus a disturbance term ϵ_t . The state Eq. (2) models the state μ_t as a function of a level component μ_{t-1} , a slope component β_t and a disturbance term ξ_t . In state Eq. (3) the slope β_t is in turn determined by a level component β_{t-1} and a disturbance term ζ_t . The level and slope components are conceptually equivalent to the intercept and slope in standard regression models with the difference that in state space models these components are not fixed for all time points. Because the level and slope components are allowed to change, they apply only locally. Hence the model is referred to as the local linear trend model (Commandeur and Koopman, 2007). The measurement and state disturbances are assumed to be normally and independently distributed around a mean of zero with a variance–covariance matrices Σ_ϵ , Σ_ξ and Σ_ζ , respectively.

The dynamics of the state μ_t is determined by the disturbances ξ_t and ζ_t . The three different states $\mu_t^{(1)}$, $\mu_t^{(2)}$ and $\mu_t^{(3)}$ underlying, respectively, Broadberry and Burhop (2008a,b), Hoffmann (1965) and Wagenführ (1933) can thus be linked through the correlations of the state disturbances (Durbin and Koopman, 2001, p. 44). For this purpose the variance–covariance matrix of ξ_t can be written as:

$$\Sigma_\xi = \begin{pmatrix} \sigma_{\xi(1)}^2 & \rho_\xi \sigma_{\xi(1)} \sigma_{\xi(2)} & \rho_\xi \sigma_{\xi(1)} \sigma_{\xi(3)} \\ \rho_\xi \sigma_{\xi(1)} \sigma_{\xi(2)} & \sigma_{\xi(2)}^2 & \rho_\xi \sigma_{\xi(2)} \sigma_{\xi(3)} \\ \rho_\xi \sigma_{\xi(1)} \sigma_{\xi(3)} & \rho_\xi \sigma_{\xi(2)} \sigma_{\xi(3)} & \sigma_{\xi(3)}^2 \end{pmatrix} \quad (4)$$

and the variance–covariance matrix of ζ_t correspondingly as:

$$\Sigma_\zeta = \begin{pmatrix} \sigma_{\zeta(1)}^2 & \rho_\zeta \sigma_{\zeta(1)} \sigma_{\zeta(2)} & \rho_\zeta \sigma_{\zeta(1)} \sigma_{\zeta(3)} \\ \rho_\zeta \sigma_{\zeta(1)} \sigma_{\zeta(2)} & \sigma_{\zeta(2)}^2 & \rho_\zeta \sigma_{\zeta(2)} \sigma_{\zeta(3)} \\ \rho_\zeta \sigma_{\zeta(1)} \sigma_{\zeta(3)} & \rho_\zeta \sigma_{\zeta(2)} \sigma_{\zeta(3)} & \sigma_{\zeta(3)}^2 \end{pmatrix}, \quad (5)$$

where ρ_ξ and ρ_ζ are correlations. I can make full use of all available information to estimate $\mu_t^{(1)}$ – i.e. the unobserved state of the target series Broadberry and Burhop (2008a,b) – by restricting its dynamics to be identical to $\mu_t^{(2)}$ and $\mu_t^{(3)}$. This means that all three states share the exact same auto correlation function of the stationary form, which is called trend homogeneity (Harvey and Chung, 2000, p. 315). To this end I place two restrictions

on the variance–covariance matrices Σ_ξ and Σ_ζ . First, when the disturbances of the different series are perfectly correlated, i.e. $\rho_\xi \pm 1$ and $\rho_\zeta \pm 1$, the level and slope components are said to be common (Commandeur and Koopman, 2007, p. 112). This means that there is only one source of stochastic movement in the three state levels. Also, there is only one source of stochastic movement in the three state slopes. Second, the variance–covariance matrix Σ_ζ is restricted to be proportional to Σ_ξ (Harvey and Chung, 2000, p. 315).

The model is estimated using the log-likelihood function in Eviews, which corresponds to the definitions of Durbin and Koopman (2001, p. 138; Eviews 8 User Guide II, 2013, p. 603; Van den Bossche, 2011, p. 3). Estimation involves a numerical search procedure that starts by choosing a set of starting values for the unknown parameters and calculating the corresponding value of the log-likelihood function. Subsequently, the process is repeated, selecting different parameter values that improve the log-likelihood function. These iterations are executed up to the point that no further improvements are obtained and the log-likelihood function is optimized. However, due to the multivariate nature of the model, the optimization process may produce either a local optimum or no solution for particular starting values. To find the global optimum, I follow Van den Bossche (2011) by using a multiple random start procedure that runs the optimization algorithm repeatedly, each time starting from a different set of initial values for the unknown parameters. The solution with the highest log-likelihood value is used henceforth.²

4. The new output series

The estimated unobserved state of the target series Broadberry and Burhop (2008a,b) is plotted in Figs. 2 and 3 together with a 99% confidence interval and the original output indexes from which the unobserved state is estimated.³ My new estimate of output change provides a fresh view on the change of output between the pre-WWI period and 1936, which is the year of the first German census of industrial production.

Previously, the period 1914–1924 was problematic because the years covering WWI are not accounted for in any of the available output indexes and the years 1919–1924 only in Wagenführ (1933). The state space

² Eviews provides different optimization procedures, i.e. Marquardt and Berndt–Hall–Hausman. I used the former first derivative technique. The program is available upon request.

³ The values of the new output series are reported in Table 3 in Appendix A.

form can deal with missing observations with relative ease and provides an estimate of output change during WWI despite the lack of information. Of course, this estimate does not necessarily capture the development of industrial output in war time accurately. The state series takes into account the sharp output drop from 1913 to 1919 in [Wagenführ \(1933\)](#) and in the absence of current-year information for the period in between it shows a more-or-less continuous decline in production during WWI. Reality, however, may have been different with the decline in output confined to one or two years only, rather than evenly spread over the entire period 1914–1918. The expanded confidence limits in [Fig. 2](#) after 1913 capture the increased uncertainty in the estimates due to the lack of observations for these years.

The accumulated effect of four years of war on German production is captured by the drop in [Wagenführ's \(1933\)](#) long-run output index over the data gap of WWI (see [Fig. 2](#)). In 1919, when current-year data is available again, [Wagenführ \(1933\)](#) puts German manufacturing at 36% of the 1913 level. This output drop is followed by a strong recovery in the years afterward, indicating that after WWI “Germany’s economy was exhausted but not ruined” ([Ritschl, 2005, p. 66](#)). The state estimate also shows rapid growth since 1919 in line with [Wagenführ \(1933\)](#), but with consistently higher output levels. This is due to the level differences between [Wagenführ \(1933\)](#) and the other two output series in 1925, when [Broadberry and Burhop \(2008a,b\)](#) and [Hoffmann \(1965\)](#) provide information again. The latter two series put Germany at higher production levels in 1925 than [Wagenführ \(1933\)](#), lifting output levels in the state also for earlier years. As a result, the drop over WWI is less pronounced in the state with output in 1919 at 55% the level of 1913.

With regard to the change of output levels between 1913 and 1925, the period on which the time-series debate focused, the state series shows a growth rate

Fig. 2. The state and observed output series (1913 = 100).

Fig. 3. The state and observed output series (1936 = 100).

slower than [Hoffmann \(1965\)](#) but faster than [Broadberry and Burhop \(2008a,b\)](#) and [Wagenführ \(1933\)](#). With [Hoffmann \(1965\)](#) inside and [Broadberry and Burhop \(2008a,b\)](#) only just outside the confidence interval in 1925, the differences between these series are very small when measurement error is taken into account.

These findings lead to a new estimate of output change over the longer period 1907–1936. [Fig. 3](#) reproduces the results of [Fig. 2](#), with the difference that 1936 is set as the reference year. 1936 is the year of the first German industrial production census and serves in the remainder of this paper as the benchmark for all time series extrapolations.⁴ Going back from 1936, the estimated state lies in between [Hoffmann \(1965\)](#) and [Broadberry and Burhop \(2008a,b\)](#) for the pre-WWI period with no overlap between the original series and the state’s confidence interval. So with known output levels in 1936, the pre-WWI output levels indicated by the state series are significantly different from [Hoffmann \(1965\)](#) and [Broadberry and Burhop \(2008a,b\)](#).

5. Implications for labor productivity

The new estimate of output change can be used to calculate the change in German manufacturing labor productivity and, hence, the change in comparative German/British productivity. Extrapolating backward from a known level of comparative German/British productivity in 1936 yields levels of comparative performance in earlier years. All other things equal, the differences between the various German output series for years before 1936 (captured in [Fig. 3](#)) lead to the same differences in comparative German/British levels of labor productivity. The following discussion is

⁴ [Wagenführ's \(1933\)](#) series cannot be included in the backward extrapolations as it ends in 1931.

based on the working assumption that ‘all other things’, i.e. the British output and labor series and the German labor series, capture true values. Strictly speaking, this may not be the case. For instance, it can be plausibly argued that the available estimate of British output change is also based on proxies (Feinstein, 1972, p. 207). Even the employment counts, both in Germany and Britain, suffer from sampling errors and may not capture the true change in employment. Be that as it may, the purpose of this section is to illustrate the consequences of my findings on German industrial output for our understanding of German/British comparative performance and for now this is done best without muddying the water by simultaneously casting doubt on the accuracy of all other existing estimates.

Going back to the pre-WWI period, my new output estimate indicates a level of production in 1907 that is 6.0% higher than Hoffmann (1965) and 4.5% lower than Broadberry and Burhop (2008a,b). Using these output series to calculate comparative German/British performance yields the exact same differences. As Broadberry and Burhop (2008a,b) find a comparative German/British labor productivity level in 1907 of 112.4, it follows that the use of my new output series results in a comparative labor productivity level that is 4.5% lower at 107.4. This is in turn 6.0% higher than the comparative performance level of 101.3 calculated with Hoffmann’s (1965) output series. Columns one to three in Table 2 report these levels of comparative German/British labor productivity in manufacturing. Broadberry and Burhop (2008a,b) point at a clear German lead over Britain, Hoffmann (1965) indicates a parity of performance and in between these two lies my estimate. The new estimate is closest to Broadberry and Burhop (2008a,b) and confirms that Germany outperformed Britain in 1907, but by a small margin only.

The estimates of comparative performance presented so far measure output per worker levels in German manufacturing relative to those in Britain. This is in line with the labor-productivity definition employed previously in the debate on comparative German/British performance. Alternatively, one can measure levels of output per hour worked. It is important to measure labor input as total hours worked, because working days and working weeks shortened during the interwar period (Huberman and Minns, 2007).⁵ This affects the estimates of comparative

German/British labor-productivity levels. In 1907 the average German laborer worked longer than his British counterpart. The average worker in manufacturing put in 2722 h annually in Germany and 2619 in Britain. In 1936 the situation had reversed with workers in manufacturing working 2097 h in Germany and 2255 in Britain.

Columns four to six in Table 2 report levels of comparative German/British output per hour worked. The benchmark level of comparative performance in 1936 is upward adjusted from 102.0 to 110.8. This has a level effect on the backward extrapolations and increases Germany’s lead over Britain in 1907. However, the time series of comparative German/British performance change also. With the drop in working hours between 1907 and 1936 more pronounced in Germany than in Britain, the rate of German labor-productivity growth vis-à-vis Britain was substantially faster when labor productivity is measured as output per hour worked, rather than output per worker. This has a negative effect on comparative German/British labor-productivity levels in 1907 when the time series are extrapolated backward from 1936. With respect to levels of comparative performance in 1907, the downward adjustment due to changes in the time series is larger than the upward adjustment resulting from changes to the 1936 benchmark level. The level of comparative German/British performance consequently decreases from 107.4 to 104.4. Nevertheless, the conclusion does not alter; Germany enjoyed a small lead over Britain in 1907.

Fig. 4 depicts the level of comparative German/British labor productivity obtained through backward extrapolation from 1936 with the 99% confidence interval. With the lower bound estimate of comparative German/British performance at 104.6 when labor productivity is measured per worker and 101.7 when measured per hour worked, Germany’s lead is statistically significant.

The confidence interval can be used to assess the fit between my time-series estimate and the three benchmarks of comparative German/British labor productivity in 1907 presented in the literature (see Table 1). These

Table 2
Estimates of comparative German/British manufacturing labor productivity in 1936 and 1907.

	Output per worker			Output per hour worked		
	B&B (2008)	HM (1965)	This study	B&B (2008)	HM (1965)	This study
1936	102.0	102.0	102.0	110.8	110.8	110.8
1907	112.4	101.3	107.4	109.5	98.5	104.4

Sources: see text.

⁵ German annual working hours are calculated on the basis of Hoffmann (1965, p. 214) and Huberman and Minns (2007, p. 542). British annual hours worked are calculated on the basis of the British Labour Statistics: Historical Abstract of the Great Britain Department of Employment and Productivity (1971, p. 95) and Matthews et al. (1982, p. 566).

benchmarks were originally constructed to evaluate the quality of the different output series. However, because the benchmarks for 1907 report wildly different levels of comparative performance, they only introduced additional uncertainty to the debate. As the deviation between the various German output series no longer is a point of discussion in this study, I can turn the case around. That is, rather than choosing between time series on the basis of the fit with a benchmark for 1907, I let the confidence interval around my time series of comparative German/British labor productivity assess the accuracy of the three available pre-WWI benchmarks.

Fig. 4 shows that the time series results cannot be reconciled with Ritschl's (2008) benchmark of comparative German/British labor productivity in manufacturing. The fit with the estimates of Broadberry and Burhop's (2007, 2008a,b) is much better. When labor productivity is measured on a per worker basis, both of their benchmarks fall within the confidence interval of the time series extrapolation. After adjusting for differences in hours worked, their early benchmark estimate (Broadberry and Burhop, 2007) drops out of the confidence interval.⁶ Their latest estimate (Broadberry and Burhop, 2008a,b) does not, however, and with a comparative German/British labor productivity level of 104.3 comes very close the point estimate of 104.4 obtained by backward extrapolation. This leads me to a conclusion in line with Broadberry and Burhop's work. Although Germany enjoyed a statistically significant lead over Britain in 1907, the size of this lead in labor productivity was small and certainly did not amount to the large margin suggested by Ritschl (2008).

6. Conclusion

The dynamics of industrial output change in Germany during the early twentieth century have been the topic of intensive study and debate. In the absence of data on output change, production estimates are necessarily based on output proxies. Previously, the debate revolved around the search for the best proxy to estimate output change. Instead of presenting new data, this paper contributes to the debate by filtering a common component from various rivaling estimates. This is appropriate, because economic theory implies that all proxies correlate with output, even though this correlation is never perfect and none of the

Fig. 4. Benchmark and time series estimates of comparative German/British manufacturing labor productivity in 1907.

proxies truly captures the change in production. The rivaling estimates can be seen not to be either right or wrong, but may all be treated as carrying information about output change. With the help of state space time series analysis this paper captures that information by estimating a common dynamics that is shared between the rivaling estimates. By adopting a new approach to the data, rather than presenting yet another proxy, this paper breaks with the previous literature. Nevertheless, while no new data is presented, the explored methodology is by no means an attempt at replacing the arduous work of collecting historical data from archival sources. Rather, this paper provides a means to make full and efficient use of the historical evidence on hand and should be seen as complementary to data collection.

The application of state space time series analysis revealed a change of output levels between 1907 and 1936 that is slower than indicated by Hoffmann (1965) and faster than Broadberry and Burhop (2008a,b), the latter being the final revision of Ritschl's (2004) output index. When used to calculate comparative German/British labor productivity, the new estimate leads to a German advantage over Britain. My new output index suggests a level of output per hour worked in 1907 that is 4.4% higher in Germany than in Britain. This German

⁶ This happens because the benchmarks are only adjusted for German-Anglo differences in working hours in 1907, while the time series take into account the change in comparative workings hours over the period 1907–1936. In other words, the adjustment for hours worked is not the same for the benchmarks and the time series; a level versus a growth effect.

lead is statistically significant, but very unlikely to have extended to a margin of more than 7%. These findings align with the work of Broadberry and Burhop (2008a,b), but not of Ritschl (2008). However, the conclusions regarding German/British comparative performance levels may not be carved in stone yet. For the sake of convenience and to clearly illustrate the labor-productivity consequences of my new estimate of German output change, the German employment series are assumed to reflect the true change in employment. This need not be the case, as there has been debate on the change of employment in German industry (Fremdling, 2007). In similar vein, the British industrial output series are treated as to capture true output change. Even though there has not been a discussion about rivaling estimates, the British index is constructed with the aid of proxies, too (Feinstein, 1972, p. 207). Although convenient for now, this is against the spirit of the paper and invites further research.

The core contribution of this paper to the economic history literature is the new approach that is explored for estimating unobserved variables with common components. The potential of this approach is illustrated on the basis of the debate on German output growth, but the benefits of state space time series analysis are not necessarily limited to this particular case. There are a number of scenario's in which the variable of interest is unobserved and state space time series analysis may prove useful in the field of economic history. Especially in cases where two or more figures have been produced with different methods the state-space form may be used for estimating common components in order to make full and efficient use of all the data. Examples are historical national-account studies, such as the debate on British national income mentioned earlier, where various estimates aim to capture the dynamics of the same variable. Similarly, in the discussion on British/American comparative performance in the nineteenth century the use of different techniques to compute relative price levels has yielded discrepant levels of comparative labor productivity (Ward and Devereux, 2003; Broadberry and Irwin, 2006). Rather than choosing between one approach or another, the debate may benefit from a study that simultaneously considers all historical evidence on hand.

A related modern-day example concerns the revisions of the International Comparison Program (ICP). There is a growing discomfort with large ICP revisions that significantly change the course of history (Deaton and Aten, 2014). A sudden 'rewrite' of history begs the question if it is appropriate to reject all previously

known information and whether the latest figures always present the most accurate estimate (Johnson et al., 2013). Perhaps older vintages should be taken on board in calculating future vintages. This idea may apply more broadly than to the ICP alone. For instance, updates of the Maddison data set have modified many of the pre-1820 per capita income estimates (Bolt and van Zanden, 2014). If the revision involves improved data only, this is clearly a gain. When the new figures are derived from the use of alternative proxies, as in this paper, the situation may be more complex. One possible solution is explored by Rambaldini et al. (2005). Using modern-day panel data, they illustrate how the state-space form can be applied to estimate relative price levels over a time-span of three decades, making full and efficient use of all data on hand. The necessity to make optimal use of the full body of available evidence is arguably larger in historical studies where data obtainability is generally limited and a similar application of the state-space form may thus provide a helpful tool of analysis.

Acknowledgments

This research is supported by a grant from the Netherlands Organisation for Scientific Research (NWO Grant no.: 360-53-102). My particular thanks go to Jan P.A.M. Jacobs, Herman J. de Jong, Marcel P. Timmer, Stephen N. Broadberry, Albrecht Ritschl, Martin Uebele, Laurie S.M. Reijnders and Pieter J. Woltjer for their comments on earlier versions of the paper.

Appendix A. Data appendix

The unobserved change in output (last column in Table 3) is estimated from the behavior of the proxy-based output series reported in Broadberry and Burhop (2008a, b), Hoffmann (1965) and Wagenführ (1933) (columns one through three in Table 3). Alternatively, Ritschl's (2004) output series could have been used instead of Broadberry and Burhop (2008a,b). I have chosen to use the latter, because it already includes Ritschl's (2004) proposed changes to Hoffmann (1965) and in addition applies an improved weighting scheme, thus incorporating the latest insights in the debate.

The output series of Broadberry and Burhop (2008b, p. 26) and Hoffmann (Broadberry, 1997, p. 43–44) have been taken directly from the literature. The series of Wagenführ (1933, p. 58–61, 64) has been converted from an industry to a manufacturing base. The industry weights used for this adjustment are derived from Hoffmann (1965, p. 390).

Table 3

The observed proxy-based output series and the estimate of unobserved output change in German manufacturing, 1900–1938 (1913 = 100).

Year	Broadberry and Burhop (2008a,b)	Hoffmann (1965)	Wagenführ (1933)	Unobs. output (this study)
1900	58.5	60.1	65.0	60.9
1901	58.6	60.1	65.2	60.2
1902	59.7	61.3	69.2	60.0
1903	63.6	65.5	73.4	64.3
1904	66.2	68.2	78.3	66.1
1905	69.0	71.2	80.1	69.4
1906	71.3	73.8	84.8	71.0
1907	76.6	79.2	83.2	79.3
1908	77.6	80.1	78.8	81.6
1909	79.1	82.0	81.5	82.7
1910	81.1	85.2	89.1	83.9
1911	86.8	90.4	96.8	88.2
1912	94.8	97.2	99.3	96.9
1913	100.0	100.0	100.0	100.0
1914				94.2
1915				83.7
1916				72.5
1917				63.1
1918				57.2
1919			36.2	55.6
1920			54.2	60.1
1921			66.0	66.6
1922			71.8	72.3
1923			47.4	77.6
1924			70.9	89.2
1925	96.1	104.7	83.4	102.4
1926	88.1	106.2	79.8	105.6
1927	111.9	120.3	101.2	115.5
1928	114.2	121.6	103.3	115.3
1929	116.4	123.6	104.0	116.8
1930	97.4	109.3	91.6	100.2
1931	81.0	91.5	74.0	81.7
1932	70.5	79.7		73.9
1933	78.6	87.8		79.8
1934	95.6	105.8		96.1
1935	108.1	126.5		117.2
1936	121.4	139.6		131.8
1937	134.7	151.2		143.0
1938	151.1	168.5		157.4

Sources: see text above.

References

- Baten, J., Crayen, D., Voth, H.J., 2014. Numeracy and the impact of high food prices in industrializing Britain, 1780–1850. *Review of Economics and Statistics* 96 (3), 418–430. http://dx.doi.org/10.1162/REST_a_00403.
- Bolt, J., van Zanden, J.L., 2014. The Maddison Project: collaborativeresearch on historical nationalaccounts. *The Economic History Review* 67 (3), 627–651. <http://dx.doi.org/10.1111/1468-0289.12032>.
- Borchardt, K., 1979. Zwangslagen und Handlungsspielräume in der großen Wirtschaft der frühen dreißiger Jahre. *Jahrbuch der Bayerischen Akademie der Wissenschaften*, pp. 85–132.
- Broadberry, S.N., 1993. Manufacturing and the convergence hypothesis: what the long-run data show. *The Journal of Economic History* 53 (4), 772–795. <http://dx.doi.org/10.1017/S0022050700051317>.
- Broadberry, S.N., 1997. *The Productivity Race: British Manufacturing in International Perspective, 1850–1990*. Cambridge University Press, Cambridge, pp. 1–404. <http://dx.doi.org/10.1017/CBO9780511584770>.
- Broadberry, S.N., Burhop, C., 2007. Comparative productivity in British and German manufacturing before World War II: reconciling direct benchmark estimates and time series projections. *The Journal of Economic History* 67 (2), 315–349. <http://dx.doi.org/10.1017/S0022050707000125>.
- Broadberry, S.N., Burhop, C., 2008a. Resolving the Anglo-German Industrial productivity puzzle, 1895–1935: a response to Professor Ritschl. *The Journal of Economic History* 68 (3), 930–934. <http://dx.doi.org/10.1017/S0022050708000685>.
- Broadberry, S.N., Burhop, C., 2008b. Resolving the Anglo-German industrial productivity puzzle, 1895–1935: a response to Professor Ritschl. Max Planck Institute research paper 848, pp. 1–28. <http://dx.doi.org/10.2139/ssrn.1161106>.
- Broadberry, S.N., Irwin, D., 2006. Labor productivity in the United States and the United Kingdom during the nineteenth century. *Explorations in Economic History* 43 (2), 257–279. <http://dx.doi.org/10.1016/j.eeh.2005.02.003>.
- Broadberry, S.N., Ritschl, A., 1995. Real wages, productivity, and unemployment in Britain and Germany during the 1920s. *Explorations in Economic History* 32 (3), 327–349. <http://dx.doi.org/10.1006/exeh.1995.1014.20>.
- Burhop, C., Wolff, G.B., 2005. A compromise estimate of German Net National Product, 1851–1913, and its implications for growth and business cycles. *The Journal of Economic History* 65 (3), 613–657. <http://dx.doi.org/10.1017/S0022050705000239>.
- Commandeur, J., Koopman, S.J., 2007. *An Introduction to State Space Time Series Analysis*. Oxford University Press Inc.
- Crafts, N., 2004. Steam as a general purpose technology: a growth accounting perspective. *The Economic Journal* 114 (495), 338–351. <http://dx.doi.org/10.1111/j.1468-0297.2003.00200.x>.
- Crafts, N., Mills, T., 2009. From Malthus to Solow: how did the Malthusian economy really evolve? *Journal of Macroeconomics* 31 (1), 68–93. <http://dx.doi.org/10.1016/j.jmacro.2007.08.007>.
- de Jong, H.J., Woltjer, P.J., 2009. A comparison of real output and productivity for British and American manufacturing in 1935. Groningen Growth and Development Centre Memorandum, no. 108, pp. 1–34 (<http://irs.ub.rug.nl/ppn/317810332>).
- Deaton, A., Aten, B., 2014. Trying to understand the PPPs in ICP2011: why are the results so different? NBER Working Papers, no. 20244, pp. 1–34. <http://dx.doi.org/10.3386/w20244>.
- Deaton, A., Heston, A., 2010. Understanding PPPs and PPP-based national accounts. *American Economic Journal: Macroeconomics* 2 (4), 1–35. <http://dx.doi.org/10.1257/mac.2.4.1>.
- Durbin, J., Koopman, S.J., 2001. *Time series analysis by state space methods*. Oxford Statistical Science SeriesOxford University Press Inc.
- Feinstein, C.H., 1972. *National Income, Expenditure and Output of the United Kingdom, 1855–1965*. Cambridge University Press, Cambridge, pp. 1–244.
- Feinstein, C.H., Thomas, M., 2002. A plea for errors. *Historical Methods* 35 (4), 155–165. <http://dx.doi.org/10.1080/01615440209601205>.
- Fremdling, R., 1988. German national accounts for the 19th and early 20th century: a critical assessment. *Vierteljahrschrift für Sozial-*

- und Wirtschaftsgeschichte 75 (3), 339–357 (<http://www.jstor.org/stable/20735178>).
- Fremdling, R., 2007. German industrial employment 1925, 1933, 1936 and 1939. A new benchmark for 1936 and a note on Hoffmann's tales. *Jahrbuch für Wirtschaftsgeschichte* 48 (2), 171–195. <http://dx.doi.org/10.1524/jbwg.2007.48.2.171>.
- Goldin, C., 2001. The human-capital century and American leadership: virtues of the past. *The Journal of Economic History* 61 (2), 263–292. <http://dx.doi.org/10.1017/S0022050701028017>.
- Great Britain Department of Employment and Productivity, 1971. *British Labour Statistics: Historical Abstract 1886–1968*. H.M. Stationary Office, London.
- Harvey, A., 1989. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press. <http://dx.doi.org/10.1017/CBO9781107049994>.
- Harvey, A., Chung, C., 2000. Estimating the underlying change in unemployment in the UK. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 163 (3), 303–339. <http://dx.doi.org/10.1111/1467-985X.00171>.
- Heston, A., Summers, R., Aten, B., 2001. Price structure, the quality factor and chaining. *Statistical Journal of the United Nations* 18, 77–101 (<https://pwt.sas.upenn.edu/papers/sju00475.pdf>).
- Hoffmann, W.G., 1965. *Das Wachstum der Deutschen Wirtschaft Seit der Mitte des 19. Jahrhunderts*. Springer-Verlag, Berlin. <http://dx.doi.org/10.1007/978-3-642-86088-1>.
- Huberman, M., Minns, C., 2007. The times they are not changin': days and hours of work in old and new worlds, 1870–2000. *Explorations in Economic History* 44 (4), 538–567. <http://dx.doi.org/10.1016/j.eeh.2007.03.002>.
- IHS Global Inc., 2013. *Eviews 8 User's Guide II*.
- Johnson, S., Larson, W., Papageorgiou, C., Subramanian, A., 2013. Is newer better? Penn world table revisions and their impact on growth estimates. *Journal of Monetary Economics* 60 (2), 255–274. <http://dx.doi.org/10.1016/j.jmoneco.2012.10.022>.
- Lee, R., Anderson, M., 2002. Malthus in state space: macro economic–demographic relations in English history. *Journal of Population Economics* 15 (2), 195–220. <http://dx.doi.org/10.1007/s001480100091>.
- Matthews, R., Feinstein, C.H., Odling-Smee, J.C., 1982. *British Economic Growth, 1856–1973*. Clarendon Press, Oxford, pp. 1–712. <http://dx.doi.org/10.1093/0198284535.001.0001>.
- Pfister, U., Riedel, J., Uebele, M., 2012. Real wages and the origins of modern economic growth in Germany, 16th to 19th centuries. *EHES Working Papers in Economic History*, no. 17, pp. 1–27 (http://www.ehes.org/EHES_No17.pdf).
- Rambaldini, A.N., Prasada Rao, D.S., Doran, H.E., 2005. Constructing internationally comparable real income aggregates by combining sparse benchmark data with annual national accounts data. A state-space approach. *Centre for Efficiency and Productivity Analysis: Working Paper Series*, no. 3, pp. 1–22 (<http://www.uq.edu.au/economics/cepa/docs/WP/WP032005.pdf>).
- Ritschl, A., 1990. Zu hohe Löhne in der Weimarer Republik? Eine Auseinandersetzung mit Holtfrichs Berechnungen zur Lohnposition der Arbeiterschaft 1925–1932. *Geschichte und Gesellschaft* 16, 375–402.
- Ritschl, A., 2004. Spurious growth in German output data, 1913–1938. *European Review of Economic History* 8 (2), 201–223. <http://dx.doi.org/10.1017/S1361491604001157>.
- Ritschl, A., 2005. The economics of World War I. In: Broadberry, S.N., Harrison, M. (Eds.), *Chap. The Pity of Peace: Germany's Economy at War, 1914–1918*. Cambridge University Press, Cambridge, pp. 41–76. <http://dx.doi.org/10.1017/CBO9780511497339>.
- Ritschl, A., 2008. The Anglo-German industrial productivity puzzle, 1895–1935: a restatement and a possible resolution. *Journal of Economic History* 68 (2), 535–565. <http://dx.doi.org/10.1017/S0022050708000399>.
- Sarferaz, S., Uebele, M., 2009. Tracking down the business cycle: a dynamic factor model for Germany, 1820–1913. *Explorations in Economic History* 46 (3), 368–387. <http://dx.doi.org/10.1016/j.eeh.2009.04.002>.
- Sefton, J., Weale, M.R., 1995. *Reconciliation of National Income and Expenditure: Balanced Estimates of National Income for the United Kingdom, 1920–1990*. Cambridge University Press, Cambridge.
- Solomou, S., Weale, M.R., 1991. Balanced estimates of UK GDP 1870–1913. *Explorations in Economic History* 28 (1), 54–63. [http://dx.doi.org/10.1016/0014-4983\(91\)90024-D](http://dx.doi.org/10.1016/0014-4983(91)90024-D).
- Summers, R., Heston, A., 1991. The Penn World Table (mark 5): an expanded set of international comparisons, 1950–1988. *The Quarterly Journal of Economics* 106 (2), 327–368. <http://dx.doi.org/10.2307/2937941>.
- Szilágyi, G., 1984. Procedures for updating the results of international comparisons. *Review of Income and Wealth* 30 (2), 153–165. <http://dx.doi.org/10.1111/j.1475-4991.1984.tb00483.x>.
- Uebele, M., 2011. National and international market integration in the 19th century: evidence from comovement. *Explorations in Economic History* 48 (2), 226–242. <http://dx.doi.org/10.1016/j.eeh.2010.12.003>.
- van den Bossche, F.A.M., 2011. Fitting state space models with Eviews. *Journal of Statistical Software* 41 (8), 1–16 (<http://www.jstatsoft.org/v41/i08>).
- van Zanden, J.L., Baten, J., 2008. Book production and the onset of modern economic growth. *Journal of Economic History* 13 (3), 217–235. <http://dx.doi.org/10.1007/s10887-008-9031-9>.
- Wagenführ, R., 1933. *Die Industriewirtschaft. Entwicklungstendenzen der deutschen und internationalen Industrieproduktion 1860 bis 1932*. Vierteljahrshefte zur Konjunkturforschung. Verlag von Reimar Hobbing, Berlin.
- Ward, M., Devereux, J., 2003. Measuring British decline: direct versus long-span income measures. *The Journal of Economic History* 63 (3), 826–851. <http://dx.doi.org/10.1017/S0022050703542000>.
- Weale, M.R., 1985. Testing linear hypotheses on national account data. *The Review of Economics and Statistics* 67 (4), 685–689. <http://dx.doi.org/10.2307/1924815>.