HOW NOT TO BACKCAST TIME SERIES DATA, OR WHY BRITAIN'S POST-WAR NATIONAL ACCOUNTS COULD STILL LEAD YOU ASTRAY

Bill Martin

WP 543
June 2025

HOW NOT TO BACKCAST TIME SERIES DATA, OR WHY BRITAIN'S POST-WAR NATIONAL ACCOUNTS COULD STILL LEAD YOU ASTRAY

Centre for Business Research, University of Cambridge
Working Paper no. 543

Bill Martin
Centre for Business Research
E-mail: bm342@cam.ac.uk

June 2025

Abstract

Eleven years ago, the Office for National Statistics, Britain's main statistics agency, attempted to merge two sets of time series data in order to backcast a long history of the country's capital investment. The restored investment figures became the series for total investment in the official 'historic' national accounts between 1948 and 1996. The ONS chose to merge old investment data with new national accounts data using splicing, a common technique, and did so 'bottom up': by adding up detailed backcast figures to derive the total. This paper, a sequel to one published in September 2024, argues that the ONS methodology was a mistake. Extraordinary differences between some of the old and new data series should have alerted the ONS to the deficiency of its approach. And the agency failed comprehensively to sense check its results. The implausible re-write of Britain's economic past that resulted is still embedded in the national accounts, but might be corrected as a result of a new ONS investigation begun in response to the earlier paper. Official correction is not assured, however. As matters stand, those wishing to draw lessons from Britain's economic past are still advised not to rely on the 'historic' national accounts.

Keywords: national accounts, capital investment, UK post-war economic history, backcasting, splicing.

JEL Codes: C82, E01, N1.

Acknowledgements
I am most grateful to those who have set aside more important things to read earlier drafts. Dr Anne Harrison commented extensively. Professor Robert Rowthorn read the technical appendix. Mr Simon Briscoe and Mr Ryland Thomas made helpful observations. But I am responsible for the outcome, warts and all.

Further information about the Centre for Business Research can be found at:
https://www.jbs.cam.ac.uk/centres/business-research-cbr/

Introduction

Eleven years ago, Britain's main statistics agency, the Office for National Statistics, revised its data for capital investment. It did so for good reasons. A new accounting standard required the ONS to add research and development to the investment figures, making the national accounts more relevant to an increasingly knowledge-based economy.[1] The ONS had also to fill-in previous investment history lest it be lost – a process of 'backcasting' – by joining together old and new sets of data unfortunately separated by a host of discontinuities. The official revisions to the capital investment figures were not a success, however. Far from it: without good reason they artificially re-wrote major aspects of Britain's economic development after the Second World War. The revised ONS data, which enjoyed the official imprimatur of being 'national statistics', implausibly uplifted Britain's investment and output growth record in the 1950s, expunged a couple of well-known recessions, altered the cumulative scale and timing of the Barber Boom and Bust in the first half of the 1970s (the revised Boom began before Chancellor Anthony Barber's expansionary U-turn), and erased a much-debated thirty-year decline in company profit share.[2] The ONS itself documented weird revisions to the investment data over the years of Chancellor Nigel Lawson's Boom and Bust that began in the second half of the 1980s. These large, implausible and arbitrary re-writings of Britain's economic history are still embedded in the national accounts.

In an earlier paper, published in September 2024, I demonstrated that the offending revisions to the investment data had nothing to do with the inclusion of R&D. Robert Jump, an economics lecturer in quantitative methods, had paved the way to this conclusion in his exploration of the peculiar ONS revisions to company profits data, affected by the investment revisions, in the late-1960s.[3] He also noted '[u]nfortunately, however, the ONS has no record of the exact manner in which the 2014 gross fixed capital formation [investment] revisions were calculated […]'.[4] As my earlier paper also demonstrated, the offending revisions had everything to do with the old investment data that the ONS chose to use for its backcasting project and the way in which the ONS merged old and new investment datasets. The method of merger was described in a now long-forgotten ONS 'advisory' published in 2014.[5]

This paper returns to it, and considers the advisory through a lessons-possibly-to-be-learnt lens. My earlier paper was written with non-specialists in mind; this one too, but with the caution that it involves a more technical account of the way the old and new investment datasets were merged by the ONS. The reason for the technicalities is that there appears to be an omission in the specialist literature, which in this context provides practical but little analytical comment.

The paper is in three main sections. Section one reconsiders in detail the ONS backcasting project of 2014 that lay behind the revisions to the investment data in the national accounts. These revisions were made over the 'historic' period – an ONS term – that stretches from 1948 to 1996. Section two offers an analytical framework, further developed in Appendix A. Section three uses the framework to interpret what went wrong with the ONS revisions as a result of its method of merging old and new datasets. This interpretation is confined to a shorter period, the only one feasible given publicly-available information, beginning in 1970. The concluding section suggests three lessons and meditates on 'user engagement' with the ONS. As of this writing, the bottom-line remains unchanged from my earlier paper: 'Those wishing to draw lessons from Britain's economic past are advised not to rely on the 'historic' national accounts'.

Section 1: The ONS backcasting project reconsidered

In 2014, the ONS faced the unenviable task of restoring the official estimates of stocks of capital investment, estimates that had been suspended three years earlier during a perfect pickle of data management troubles. Estimates of stocks of capital require as inputs very long-run time series of data on investment expenditure at a fine level of industry and capital asset detail. The investment data are in nominal or 'current price' terms and, along with related price indices, are contained within a database known now, if not in 2014, as the 'PIM inputs' database, the acronym standing for the Perpetual Inventory Method. This method is used by the ONS to calculate the surviving level of capital assets that have resulted from cumulative past investments with a history that stretches back to the early-nineteenth century. In 2014, the ONS tried to restore the suspended capital stocks data in the face of major discontinuities in the available investment series. Somehow, the ONS had to stitch together old investment data with a very long history in the equivalent of the PIM inputs dataset with the national accounts dataset with, by comparison, a very short history in order to have a sufficient back-run of investment data, and in sufficient detail, in order to re-calculate the capital stocks figures. It is perhaps best to realise that the ONS was focussed on the revival of the capital stocks figures themselves; the new backcast investment figures that emerged from this project were just a step, as the title of the ONS advisory emphasises, to enable 'the estimation of capital stocks and consumption of fixed capital.' But it was these backcast investment figures, a mere step in the backcasting capital stocks project, that ended up as the 'historic' total investment series in the main national accounts.

In 2014, the old PIM data – by which is meant the current price investment data, not other series in this copious dataset – were those last used for the 2010 capital stocks publication, a year before the suspension. These data were organised around a classification of industries known as standard industrial classification (SIC) 2003. The new investment data, on the other hand, were part of the national accounts dataset most actively maintained by the ONS on its 'central processing system'. As now, these data began in 1997 – a cut-off date that explains the ahistoric, but handy, ONS description of national accounts data before 1997 as 'historic'. These investment data had been converted three years earlier to the latest standard industrial classification – known as SIC 2007 - over the 'contemporary period' from 1997 until the point at which the ONS could rely on investment data collected on a SIC 2007 basis.[6] The first challenges facing the ONS were therefore to re-create the old PIM investment data of 2010 vintage – not so simple a task as these were not maintained in current price form – and then to convert them to the 2007 classification of industries.

There were additional difficulties. The ONS had to take into account a host of revisions to the investment data that had occurred since 2010. One was the reversal of a much-trumpeted innovation introduced in the 2013 National Accounts ‘Blue Book’ publication (‘blue’ being a reference to the colour of the cover) that had led to foreseen but, it turned out, unacceptable volatility in quarterly current price investment series.[7] Other notable revisions were the introduction of new sources of information for investment in computer software and in original works of art and the reclassification of that element of software investment that had been subsumed in a machinery and equipment category; this software then joined software investment already classified as intangible.[8] To add to the complexities, the ONS had discovered material problems in the way it had converted the national accounts investment figures from the old to the new standard industrial classifications, and corrections had to be applied. All told, the ONS faced a tall order. To create a long history of investment data that might be regarded as consistently defined, ‘to account for methods changes’,[9] the ONS had to link old, out-of-date PIM investment data that were compliant with the 2003 standard industrial classification to a new national accounts investment dataset affected by many subsequent revisions. The necessary calculations were to be made before taking account of the inclusion of R&D and other changes that were introduced in the 2014 Blue Book. But the calculations had to take account of the reversal of the innovation in the 2013 Blue Book that led to excessive volatility in the investment data and of the adjustments to the national accounts figures designed to correct problems caused by the conversion three years earlier from the 2003 to the 2007 classification of industries.[10]

To recover the capital stocks data, a new team was created under the aegis of the national accounts ‘gross capital formation branch’ but comprising members with different skills and experience. One member of what can be conveniently called the ‘capital stock team’ was an economics undergraduate on an ‘industrial placement’. If the social networking platform ‘LinkedIn’ is any guide, he later became a ‘data scientist and engineer’ in the private sector.[11] Another had been an assistant economist at HM Treasury. He later became the ‘Lead Data Scientist’ at the ONS Data Science Campus, an organisation set up following a recommendation by Professor Sir Charles Bean in his review, published in 2016, of UK economic statistics.[12] The presence of economists on the capital stock team pre-dates Bean’s recommendation to embed people with economics expertise in the ONS teams responsible for data production, encouraging them, Bean argued, to be ‘more self-critical’. Bean added, ‘it may also yield some quick returns in moving to a smarter approach to quality assurance and a reduction in the frequency of unnecessary errors’.[13] Another member of the capital stock team was the highly experienced ONS lead for national accounts investment data. With a background in astrophysics, she later became the ONS Chief Data Scientist.[14] The contact point for the ONS advisory that described the team’s results was an

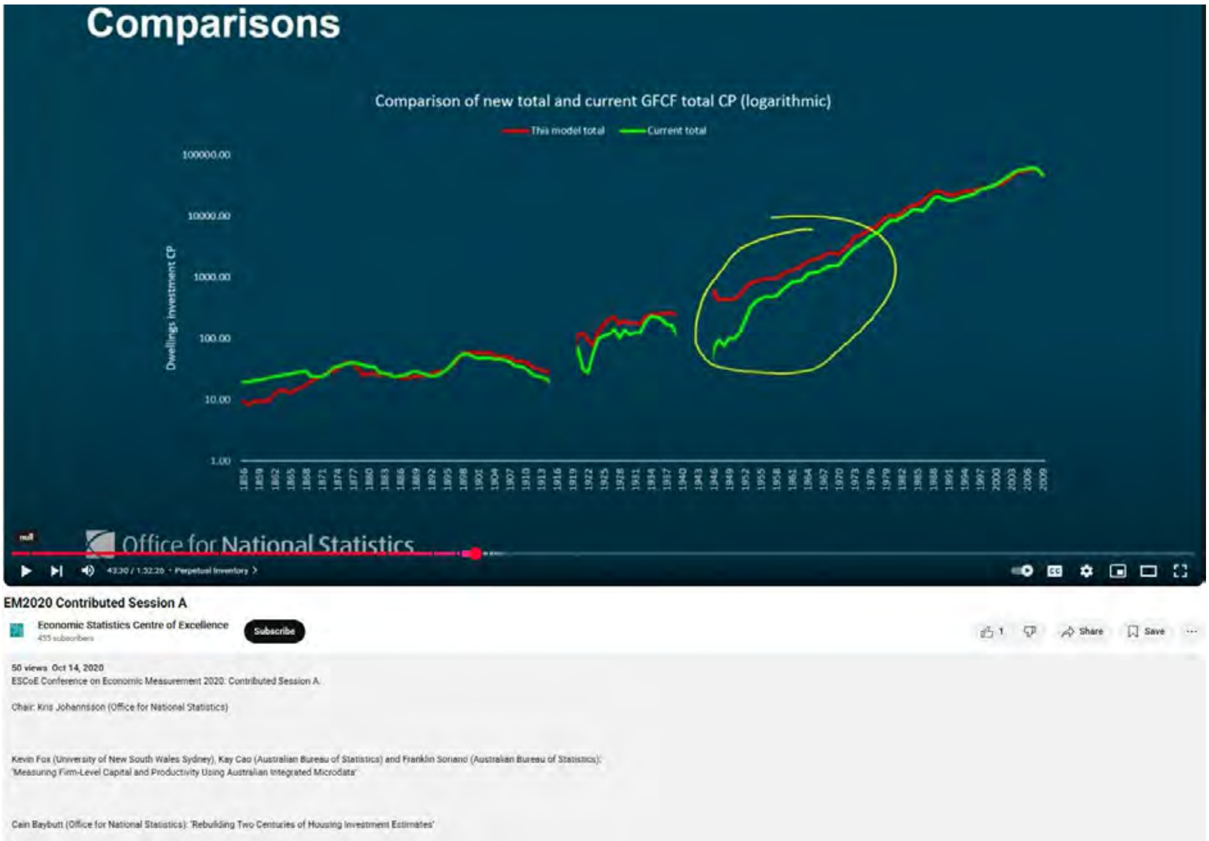
ONS statistician with specialist knowledge of capital stocks data.[15] It would be reasonable to conclude that the capital stock team had the inter-disciplinary skills and experience to undertake the difficult task of recovering the capital stocks data and of backcasting the investment data required for the national accounts. The more telling then that the team inadvertently re-wrote Britain's economic history. Two aspects of the team's approach produced this unfortunate result: the quality of the investment data that the team used and the method that the team chose to backcast the history of investment before 1997.

The old 2010 PIM dataset used by the team recorded the investment expenditures of many industries in fine detail.[16] The quality problems were of two broad types. First, the PIM data of 2010 vintage were unlikely to have been fully consistent with the national accounts data. Although a preliminary search has not uncovered any official qualification issued at the time of the 2010 capital stocks publication, the likelihood of inconsistency can be inferred from the warning that currently accompanies PIM datasets that, unlike the one in 2010, have usefully been released by the ONS in more recent years.[17] The latest PIM team's user guide states:

> This [the PIM dataset] contains current price estimates of GFCF [gross fixed capital formation] from 1828 to the latest period, broken down by sector, industry, and asset. Please note that currently these estimates can differ from published estimates of GFCF, but we are working to align these estimates.[18]

By comparing the December 2023 version of the PIM dataset with the 2010 Blue Book, my earlier paper identified a large anomaly in the PIM record of new dwellings, specifically the record of government investment in what was formerly called council housing. In 1948, the shortfall in the PIM figure for dwellings relative to the national accounts estimate was equivalent to 2¼ per cent of current estimates of the nominal gross domestic product.[19] Although not the only possible explanation, an implication is that the capital stock team may have used a dataset that greatly understated investment in dwellings, a mistake that made its way to the revised national accounts estimates for total investment (no asset detail was provided) that first appeared in the 2014 Blue Book. This inference can also be drawn from a September 2020 conference presentation by an ONS statistician who had investigated the PIM investment data for dwellings. The conference, hosted by the Economics Statistics Centre of Excellence (ESCoE), is available on-line.[20] Reporting his findings, the statistician noted a 'significant gap' between his proposed estimates and those in then PIM dataset for investment in dwellings data 'before 1985'. As Exhibit 1 shows, the 'gap' was especially marked in the 1940s and 1950s.[21]

Exhibit 1: ONS presentation in 2020 on PIM dwellings investment data



Sources: Baybutt (September 2020); ESCoE. **Notes:** On a logarithmic scale, the ONS presentation chart compares the then PIM estimate of investment in new dwellings in current prices (labelled ‘Current total’) with the estimates proposed (labelled ‘This model total’) by the ONS statistician who had investigated the PIM dataset. Permission to reproduce the screenshot was kindly granted by ESCoE’s Operations Director (e-mail correspondence, 23 June 2025).

In the session’s closing questions and answers, I noted the probable connection between the ‘gap’ in the presented estimates of dwellings investment and the too-low level of the national accounts total investment series.[22] The ONS session chair offered the prospect of further investigation. However, as noted, a material anomaly in the record of housing investment was still present in the December 2023 version of the PIM dataset used for the calculations in my earlier paper, and remains unchanged as of this writing.[23]

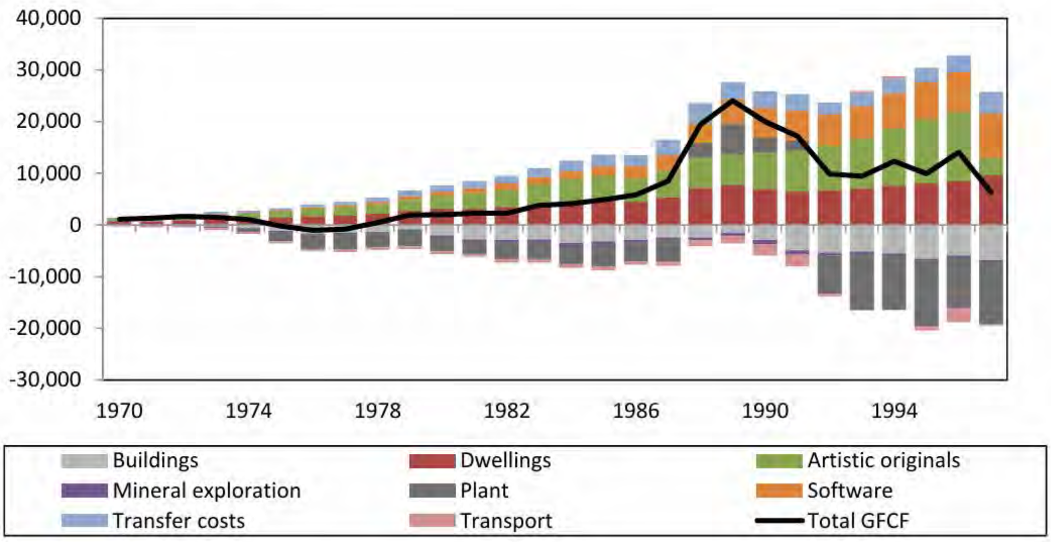
The second type of quality problem concerns the micro-level nature of the PIM dataset. In general, it is well established that micro-data can be of poor quality, being plagued by allocation errors – in this case, the attribution of investment data to industry B that in fact belong to industry A. If industry A is small and industry B is big, the error will affect A's data record disproportionately. This general point was recognised by the ONS in 2014. An advisory – a companion of the one presented by the capital stock team - pointed out that '[t]he more detailed the estimate breakdowns, the higher the uncertainty, as a proportion of the series'.[24] Forty years before, econometricians were noting the advantages of using aggregate data in which errors at a more detailed level naturally cancel out: the sum of the errors in the records of industries A and B - one a minus, the other a plus - is zero.[25] Allocation errors can occur for many reasons: reporting or sampling error, method discontinuities and industrial reclassifications gone wrong. National accounts expert Anne Harrison has highlighted the possible impact of an inconsistent treatment of leasing in the official accounts. The convention between 1975 and 1988 was to attribute leased capital assets to the legal owner, that is to the financial sector. Subsequently, in accordance with revisions to commercial accounting practice, the distinction was made between the economic owner and the legal owner with a financial liability from the former to the latter. Harrison notes that this new distinction was not carried all the way back consistently in the detailed official series of asset by industry – the level at which the capital stock team was working - although it was at an intermediate level of disaggregation.[26] In addition to such inconsistencies, the conversion in 2014 of the investment data, from the 2003 to the 2007 standard industrial classifications, introduced errors in addition to those that would have existed in the 2010 PIM dataset.

The conversion was based not on investment data, which were regarded as too volatile, but on business turnover, and at one brief point in time (December 2009) when data on the old and new industrial classifications co-existed.[27] One effect was to misallocate capital assets across industries: even if by chance the correct amount of investment was re-allocated from industry A to industry B, it was all too likely that the asset composition of the reallocation would have been incorrect. A re-allocation of investment data across industries should not have affected the economy-wide totals for different types of capital asset, but it did.[28] For the contemporary national accounts period beginning in 1997, the ONS discovered that the record of investment in buildings was the category most affected by asset allocation errors. Next in line was the record of investment in plant. The ONS applied corrections to the national accounts data in the interval 1997 to 2010 - they were especially large in the 2007 to 2010 period - but a close reading of the ONS advisories leaves it unclear whether comparable corrections were applied to the PIM investment data before 1997 following conversion to the new industrial classification. These data are likely to have been affected by misallocations between industries for the same reason.[29] If no corrections were applied, there would be an additional reason to doubt the reliability of the capital stock team's backcasting results.

In addition to the poor quality of the historic PIM dataset, it was the team's choice of method that explains its unjustified re-writing of Britain's economic record. The team chose to 'splice' the investment data together using a frequently-used technique. Each of the old 2010 PIM investment series after the conversion to the 2007 standard industrial classification were re-scaled throughout the entirety of the back-run to align the 1997 value of each series with the 1997 value of the corresponding series in the national accounts dataset. The team referred to the quotient of the 1997 value of a new national accounts series to the 1997 value of the old PIM series as the 'linking factor'. The effect of the method was to equate the rates of growth of each of the backcast micro-series with those of the old PIM series. The revised historic figure for total investment that emerged was the summation of all the spliced micro-series, an approach to splicing that can be conveniently called 'bottom up'.

Exhibit 2: Impact of ONS splicing: Capital stock team's advisory 'Figure 2'

Figure 2: difference in the asset composition of total GFCF for the pre-linked and post-linked series (CP, £m)



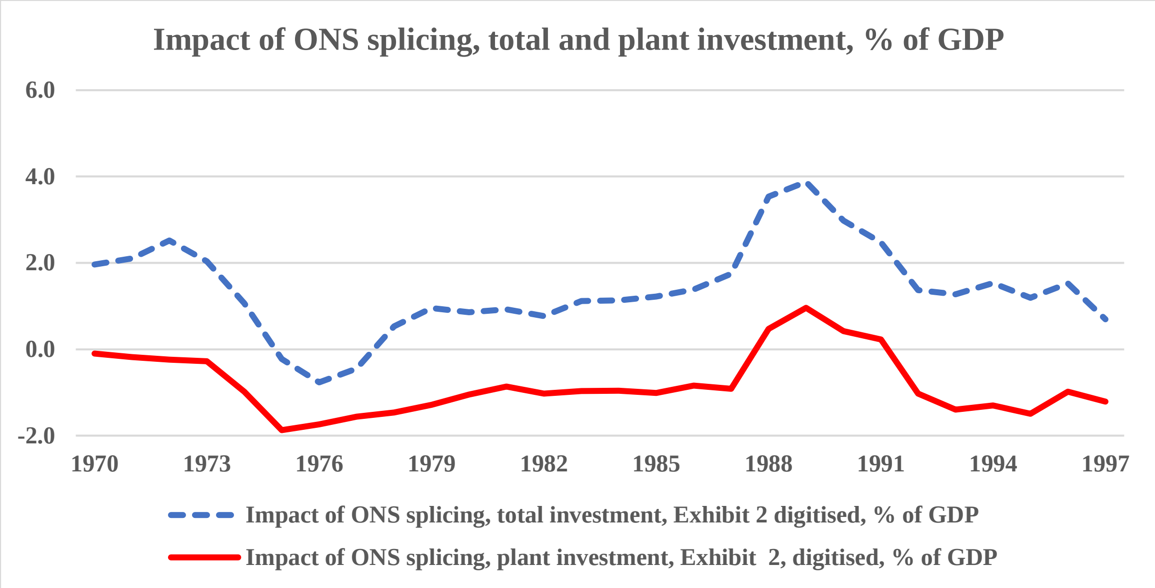
Source: Mosquera et al. (2 July 2014), p. 9. **Notes**: Mosquera et al., p. 8 defines the 'pre-linked' series as the 'version of the GFCF [gross fixed capital formation] used in the 2010 publication of capital stocks. For the CP [current price] data, it has been converted to SIC [standard industrial classification] 2007' and the 'post-linked' series as 'the current input into the PIM […] It includes the […] conversion to SIC [standard industrial classification] 2007, and linking the pre-1997 series to the post-1996 series […].'. Mosquera et al. (2 July 2014) does not precisely define the capital asset classes shown in 'Figure 2'. Most are identifiable: Buildings (Other buildings and structures), Dwellings (new), Artistic Originals (intellectual property classified under 'entertainment'), Mineral Exploration (exploration for petroleum and natural gas), Software (computer software), Transfer Costs (cost of ownership transfer), Transport (equipment). The residual asset class 'plant' can be inferred largely to comprise information and communications technology equipment (ICT) and non-ICT machinery and equipment (which at the time excluded military weapons systems but included some computer software expenditure otherwise classified under intangible investment – see Oulton and Srinivasan (2003), p. 78). 'Plant' can be taken to correspond to the 2010 Blue Book asset category 'other machinery and equipment and cultivated assets' and the 2010 Capital Stocks publication asset category 'plant and machinery'. Appendix B provides further details.

The impact of the splicing is reported by the team in the form of a colour-coded combination stacked-bar and line chart: the advisory's 'Figure 2' reproduced here as Exhibit 2.[30] The results are traced as the difference between the investment data before and after splicing (or 'linking' – the terms are interchangeable) for the period 1970 to 1997. The data are expressed in current prices, with the unfortunate effect that the differences in the 1970s inevitably appear far smaller than those that occur later simply as a result of inflation. Over the 1970 to 1997 period, capital investment prices rose some 600 per cent.[31] Even so, it is clear from the 'figure' that the splicing had a marked effect on the record of total investment during the period of the Lawson Boom and Bust, between the late-1980s and early-1990s. The ONS 'figure' also shows the main cause of this roller-coaster: the switch in the impact of splicing on plant (and machinery) investment from negative to positive and back again.

To appreciate the effect of the splicing on the wider economy, the results displayed in the team's 'Figure 2' are better re-expressed as shares of current estimates of nominal GDP.[32] As an extensive search has not uncovered the underlying data, it is necessary to extract them by a process of digitisation, the software equivalent of a ruler and pencil, that estimates what the points on the 'Figure' imply given the scales of the vertical and horizontal axes. The results of the digitisation are subject to error, the more so for the extraction of the impact of splicing on plant investment depicted as a stacked bar. Invisibility of the early-1970s impacts necessitates some inspired interpolation. Nevertheless, the digitisation (and interpolation) provides a sense of orders of magnitude which accord reasonably well with the brief description provided in the capital stock team's advisory.[33]

Chart 1 shows the extracted data, as a per cent of GDP, tracing the impact of the capital stock team's splicing on investment, both in aggregate and for the plant category. Compared with the team's 'Figure 2', the feature added by the re-expression of the results can be seen in the first half of the 1970s. Whereas before little impact was visible, re-expressed as a per cent of GDP the impact of the splicing on total investment now appears material. The impact sharply declines from an addition worth over 2 per cent of GDP in 1972 to a decrement worth over ½ per cent of GDP in 1976. This marked change is clearly related to the oscillations in ONS revisions to the volume of GDP that took effect in the 2014 Blue Book, which altered the timing and cumulative scale of the Barber Boom and Bust.[34] The splicing impact on plant investment played a significant role, accounting for nearly half of the downward swing in the impact on total investment between 1972 and 1976.

Chart 1: Impact of the 2014 ONS splicing calculations



Sources: Mosquera et al. (2 July 2014), 'Figure 2', p. 9; UK Economic Accounts release, 28 March 2025 – for nominal GDP. **Notes**: The digitisation of the advisory's 'Figure 2' is necessarily approximate, especially for plant investment in the 1970s. The too-small-to-see plant data before 1973 are interpolated back to zero. Repeated trial attempts at digitisation produced similar results. The software used was: 'Plotdigitizer: 3.3.9, 2024, Plotditizer.com'.

The impact of splicing on plant investment played an even more decisive role in the upward and downward swings in the splicing impact on total investment that began in the second half of the 1980s. The impact on total investment moved from about 1½ per cent of GDP in 1986 to 4 per cent of GDP in 1989, returning to 1½ per cent of GDP in 1992. Over the same period, the impact on plant investment as a per cent of GDP swung from a negative near-1 per cent to a positive 1 per cent at the peak in 1989, then falling back to a negative 1 per cent three years later. The up and down swings in the impact on plant investment account for around three-quarters of the up and down swings in the impact on total investment. The related impacts on investment, GDP and other national accounts aggregates as they appear in the national accounts are less easily discerned, however. It is possible that the ONS introduced balancing adjustments of one form or another in 2014, on to which have come later data revisions. The result is that the distorting impact of the ONS splicing on the history of the Lawson Boom and Bust is buried deep in the undergrowth of the national accounts.[35]

The implied criticism that arises from this account of the historic re-writing is not focussed on the use of splicing as a backcasting method. In the absence of a clear alternative, splicing offers the opportunity to recover long-run datasets when other methods are either impractical or not necessarily better. Professor Sir Charles Bean gave his qualified approval of splicing '[w]here consistent historical series are not yet available'.[36] Alternative econometric methods, on which the specialist literature focuses, usually depend on the presence of a period in which old and new data co-exist – an overlapping period – of sufficient length and representativeness that econometric estimates can be used confidently without running the risk of small-sample or other prediction errors.[37] The longer the period of backcast, the greater that risk. Splicing may fare no worse. Perhaps because of its simplicity, it is the backcasting method was commonly used by national accountants, and is widely used by other researchers.[38]

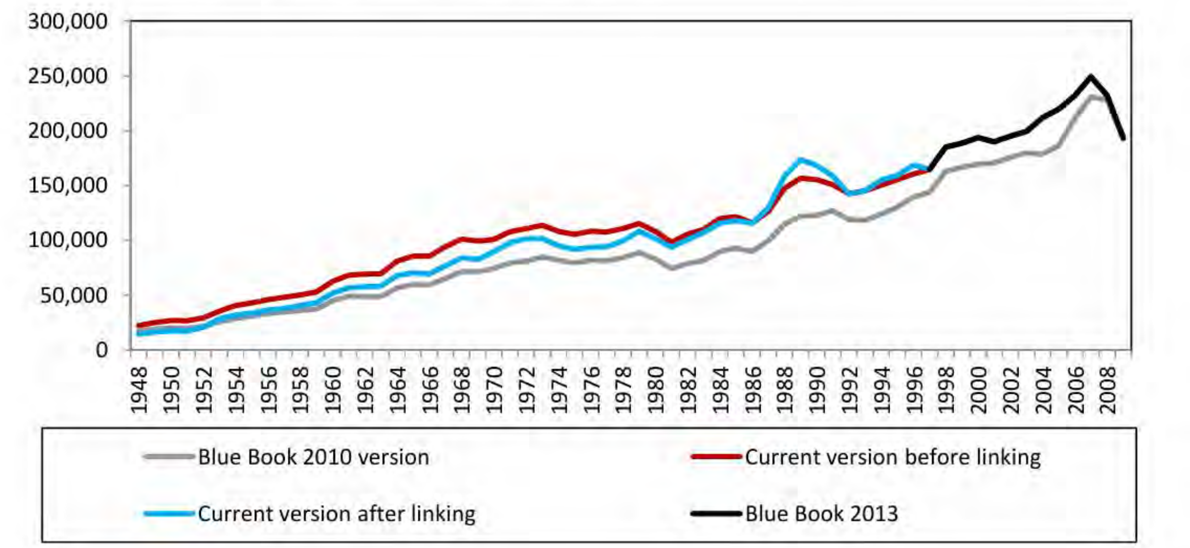
But there are strong provisos. The risk that splicing can lead the data compiler inadvertently to over-write well-established economic history is emphasised in practical guides. In 2014, around the time of the ONS exercise, a United Nations meeting of expert practitioners advised national accountants who had to backcast data to ensure that 'key known economic events are not altered by the splice', and to deploy the skills and knowledge of 'an economic historian' when checking the backcast results for plausibility.[39] International national-accounting handbooks and manuals similarly stressed the need when backcasting to minimise 'changes in the economic history of a country'.[40] A 2018 United Nations draft 'Handbook on Backcasting' addressed the specific problem faced by the ONS capital stock team four years earlier. The handbook warned that 'a change in classification which simply re-uses old data mapped to a new classification […] should not change the level or trend of GDP or the aggregate to which the reclassification applies'.[41] So what went wrong in 2014? It will be argued that the capital stock team's splicing methods were marred by three errors of judgement, if not also of omission.

First, at a general level, the capital stock team appears not to have thoroughly 'sense checked' the results of its calculations for plausibility – beyond, that is, its close scrutiny of the impact on the capital stocks data – and seemed strangely incurious about the impacts on the history of investment that the team did detect. The capital stock team advisory provides a chart, 'Figure 9', reproduced below as Exhibit 3, that compares, amongst other series, volume measures of total investment before and after linking. Presenting this comparison for the period 1948 to 1997, the advisory notes:

A further examination of this [matter] shows that growth in total GFCF [gross fixed capital formation] in the linked data diverges from the pre-linked estimates (red series) at times where detailed industry by asset series experience sharp changes in investment up to and including 1997. […] Users should be reminded that, because of the methodological changes to GFCF since 2010, changes both in the levels and growth rates of GFCF (pre- and post-1997) are to be expected.[42]

Exhibit 3: Long-run comparisons: Capital stock team's advisory 'Figure 9'

Figure 9: comparison of total GFCF under different methodologies. 'Blue Book 2013' estimates that start in 1997 are chained volume measures (CVM, reference year=2006), while all other series are KP (reference year=2006)

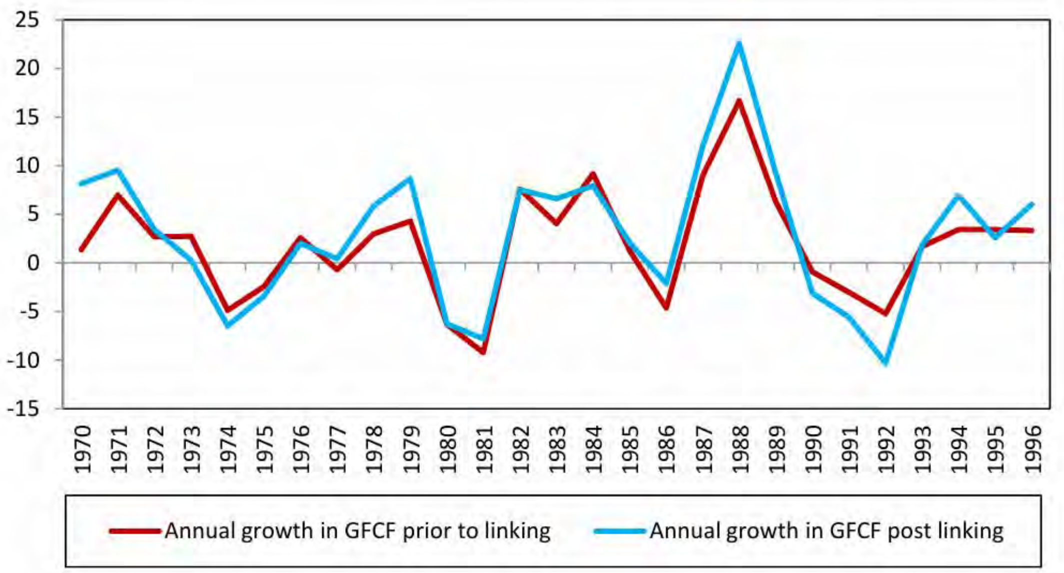


Source: Mosquera et al. (2 July 2014), p. 13.

The advisory's defence that such changes were only to be expected skirts over the clearly-to-be-seen departure of the levels of the post-linked from the pre-linked volume investment series (the blue and red lines in 'Figure 9') in the period of the Lawson Boom and Bust and in the period before the late-1970s. What cannot be seen clearly is the too-low-level of investment in the immediate post-war years that came with the 2014 Blue Book revisions. The impact of the team's splicing on growth rates is more clearly depicted in the advisory's next 'Figure 10', reproduced as Exhibit 4. This 'Figure' is limited to the period 1970 to 1996. For reasons that are unclear, the team appears unfazed by the large impact of its splicing on the volume growth of investment in the early-1970s, the late-1970s and over the Lawson Boom and Bust. What is also missing is a closer examination by the team of the wider impact of these splicing impacts on the historic record of GDP and related data, such as profits. The need for such scrutiny, sometimes known as a ‘sniff test’, was understood by at least one other branch of the ONS at the time and was later repeatedly stressed in the Bean Report.[43]

Exhibit 4: Investment growth comparisons: Capital stock team’s advisory ‘Figure 10’

Figure 10: annual growth in total GFCF time series prior to linking and post-linking (KP, %)



Source: Mosquera et al. (2 July 2014), p. 14.

A second error of judgement or of omission concerns the need to pre-test the data to be spliced, to determine whether there are signs of instability in the relationships between the old and new series during an overlapping period, which in this case ran from 1997 to 2009. It is unclear whether the team conducted thorough pre-tests, or, if it did, heeded inauspicious results. The brief statement that was used by the team to justify its decision to splice the data using 1997 values of the national accounts dataset rather than over the longer overlapping period available can be interpreted to imply that there were instabilities.[44] A similar inference can be drawn from the team’s discovery of spectacularly large linking factors for a few of the series. If an old series differs so much from the new series to which the old series is to be spliced, a question naturally arises about the reliability of the old series to help represent the missing history of the new series.

But even if the pre-test had been applied and no instabilities found, and even if there were no questions left begging about the representativeness of the old series, the method that the capital stock team decided to use was destined erroneously to re-write economic history. The team’s third error of judgement was to persist with its chosen method after it had discovered the spectacularly large linking factors.

Some – the advisory may not have reported all such cases – are documented in a series of ‘figures’ in the capital stock team’s advisory. These show the linking factors ranging from 1.4 in the case of investment in building by real estate activities, including those of housing associations, to 17.9 in the case of investment in original works of art produced by those involved in radio and television programming and broadcasting.[45] These linking factors mean, approximately, that the 1997 value of the national accounts series were respectively 40 per cent and 1,690 per cent greater than the equivalent PIM series. Table 1, taken from my earlier working paper, gives the details. This small sample of linking factors was not selected at random by the capital stock team; the linking factors it chose to report were precisely those that had a material impact on the team’s results. It can be reasonably inferred that these linking factors were not only large but aberrantly so: outliers that were different from the majority of the remaining linking factors that the capital stock team chose not to report.

Table 1: Linking factors given in the capital stock team’s advisory

Industry (SIC 2007 division number)	Investment Asset	1997 linking factor	Difference in 1997, %
Construction of buildings (41)	Buildings	1.9	90
Telecommunications (61)	Plant	7.8	680
Real estate activities (68)	Buildings	1.4	40
Publishing activities (58)	Artistic originals	8.8	780
Film & TV production, etc. (59)	Artistic originals	8.2	720
TV programming, etc. (60)	Artistic originals	17.9	1,690

Sources: Mosquera and others (2 July 2014), ‘Figures’ 3 to 8. **Notes**: ‘SIC 2007’ stands for the 2007 version of the Standard Industrial Classification. It is assumed that the title to ‘Figure 4’ in the ONS advisory is mislabelled. The title refers to industry SIC 2007 division ‘61’, the legend to ‘Figure 4’ states ‘industry 68’. The final column in the table re-expresses the ONS linking factors in 1997 as the difference in 1997 between what the advisory refers to as the ‘current version’ and the ‘Blue Book 2010 version’ of each series expressed as a per cent of the ‘Blue Book 2010 version’, calculated as 100 multiplied by (the linking factor minus 1). The results will be subject to rounding error. The linking factors for original works of art (‘artistic originals’) are affected by ‘improved’ methods of estimation introduced in the 2013 Blue Book.

It is not self-evident that the capital stock team understood the implications. With regard to its ‘Figure 2’ (Exhibit 2 here), which displayed the impact of splicing on the aggregate series for total investment and its capital asset composition, the advisory states:

> As described above, the linking process is inherently uncertain. In the absence of better information, a single point-in-time factor has been applied to the whole back series. This can have the effect of inflating (or deflating) levels in the historic estimates.[46]

This statement is misleading. The ‘inflating’ or ‘deflating’ impact of the splicing on these historic aggregate series owed little to the use of constant linking factors. The inflation and deflation of the levels of the historic totals arose because the capital stock team chose to splice the series at the micro-level and to derive the totals, the aggregate series, by adding up the micro-results: a ‘bottom-up’ approach. The likelihood of the bottom-up splicing having an inflating and deflating effect on the aggregate series was signalled by the presence of outliers amongst the range of linking factors. In these circumstances, the splicing effect, of inflating and deflating, would have been avoided had the team instead chosen to begin the calculations by splicing the aggregates directly: a ‘top-down’ approach.

It is an omission of the existing specialist literature that there appears to be no formal analysis of this aspect of the splicing method. Most of the literature focuses on practical advice. As far as can be ascertained, very few contributions are analytical and, of these, none deal with the problems that can be caused when splicing is performed ‘bottom up’ on micro-data that, possibly because of pervasive data error, are characterised by aberrant linking factors.[47] The next section attempts to fill this gap by providing a simple analytical framework, which can be applied to the peculiar impact on plant investment of the capital stock team’s bottom-up splicing.

Section 2: Splicing bottom up and top down – an analytical framework

Consideration is given to a basic case in which an aggregate time series – taking the example of the economy-wide total for investment in plant – comprises investments by just two industries: the telecommunications industry and, taken as a group, all the remaining industries. To keep things simple, the remaining industries are assumed to be identical. The analysis proceeds by presenting symbolically the basic splicing method, using this set-up to illuminate why bottom-up splicing can inflate or deflate the historic total and, finally, moves to a comparison of bottom-up and top-down methods of splicing. Appendix A presents a general case of bottom up versus top down in which there are multiple industries, encompassing the two-industry example as a special, but nevertheless instructive, case.

A ‘new’ aggregate time series – equivalent to the national accounts series – and an ‘old’ aggregate time series – equivalent to the PIM series – are denoted by the symbols X_t and Y_t respectively. The subscript *t* denotes the year, so that X_t, for example, is the value of the new series in year *t*. The new series runs from $t = \ell$ to some later year, ℓ being the linking year; the old series runs from $t = 0$ to $t = \ell$. In this example, there is no overlapping period. The linking factors, designated by the Greek letter phi, are formed by dividing the values of the new series by the corresponding values of the old series at the linking year, thus:

For telecommunications (*TC*):

$$\Phi_{TC} = \frac{X_{TC,t=\ell}}{Y_{TC,t=\ell}} \tag{1}$$

For the other (assumed identical) industries (*OI*):

$$\Phi_{OI} = \frac{X_{OI,t=\ell}}{Y_{OI,t=\ell}} \tag{2}$$

Since the data in mind refer to current price investment, economy-wide plant investment is the simple sum of the investments in plant by the telecommunications industry and the other industries. In the case of the old series, total plant investment is described by the equation:

$$Y_t = Y_{TC,t} + Y_{OI,t} \tag{3}$$

The equivalent equation for the new series is:

$$X_t = X_{TC,t} + X_{OI,t} \tag{4}$$

When the telecommunications industry and other industries investment data are spliced, they form backcast series which are estimates of what the new series would have looked like had they existed in the historic period from $t = 0$ to $t = \ell - 1$. This interval is the equivalent of the ONS 'historic' period that begins in 1948 and ends in 1996. The backcast series for plant investment by the telecommunications industry and the remaining industries are denoted $\hat{X}_{TC,t}$ and $\hat{X}_{OI,t}$ respectively and formed by splicing the historic series to the related new series at the linking year:

For telecommunications:

$$\hat{X}_{TC,t} = \Phi_{TC} Y_{TC,t} \tag{5}$$

For the other (assumed identical) industries:

$$\hat{X}_{OI,t} = \Phi_{OI} Y_{OI,t} \tag{6}$$

Two basic properties of the backcast historic series emerge. First, if extended to the linking year, they are by construction equal to the new series at the linking year. For example, in the case of the telecommunications industry, the combination of equations (1) and (5) gives:

$$\hat{X}_{TC,t=l} = \Phi_{TC} Y_{TC,t=l} = \frac{X_{TC,t=l}}{Y_{TC,t=\ell}} Y_{TC,t=\ell} = X_{TC,t=l} \tag{7}$$

The second property is that the proportionate change in the backcast industry-level investment series from one year to the next is equal to the proportionate change in the corresponding old series. For example, in the case of the telecommunications industry, the proportionate change, denoted $G(\hat{X}_{TC,t})$, can be written:

$$G(\hat{X}_{TC,t}) = \frac{\hat{X}_{TC,t}}{\hat{X}_{TC,t-1}} = \frac{\Phi_{TC} Y_{TC,t}}{\Phi_{TC} Y_{TC,t-l}} = \frac{Y_{TC,t}}{Y_{TC,t-1}} = G(Y_{TC,t}) \tag{8}$$

The capital stock team backcast total plant investment by adding up the backcast series for all the industries. In the two-industry case, this bottom-up backcast total, denoted $BU(\hat{X}_t)$, is represented by the equation:

$$BU(\hat{X}_t) = \hat{X}_{TC,t} + \hat{X}_{OI,t} \tag{9}$$

Using the backcasting formulae (5) and (6), the bottom-up backcast total can also be written:

$$BU(\hat{X}_t) = \Phi_{TC} Y_{TC,t} + \Phi_{OI} Y_{OI,t} \tag{10}$$

It can be noted in passing that any allocation errors that affect the records of the two industries with equal and opposite effect will now probably not sum to zero. In equation (10), the allocation errors are multiplied by linking factors which, if different, will render the sum of the scaled allocation errors either positive or negative.[48]

Returning to the main derivations, it will be found that unlike the backcast series for telecommunications and the other industries, the proportionate change in the bottom-up backcast total is not equal to the proportionate change in the old total. The proportionate change in the old total is a weighted average of the proportionate changes in the old series for plant investment by the telecommunications industry and by the other industries. The weights, denoted respectively $v_{TC,t}$ and $v_{OI,t}$, are the shares of the telecommunications industry's plant investment and other industries' plant investment in the old total for plant investment. These weights can vary over time. The formulae for the weights are:

For telecommunications:

$$v_{TC,t} = \frac{Y_{TC,t}}{Y_t} \tag{11}$$

For the other (assumed identical) industries:

$$v_{OI,t} = \frac{Y_{OI,t}}{Y_t} \tag{12}$$

The proportionate change in the old total series is represented by the equation:

$$G(Y_t) = \frac{Y_{TC,t} + Y_{OI,t}}{Y_{t-1}} = G(Y_{TC,t}) v_{TC,t-1} + G(Y_{OI,t}) v_{OI,t-1} \tag{13}$$

(The derivation in, for example, the case of telecommunications, comes from the relationship:

$\frac{Y_{TC,t}}{Y_{t-1}} = \frac{Y_{TC,t}}{Y_{YC,t-1}} \frac{Y_{TC,t-1}}{Y_{t-1}} = G(Y_{TC,t}) v_{TC,t-1}$ where from equation (11) $v_{TC,t-1} = \frac{Y_{TC,t-1}}{Y_{t-1}}$.)

By contrast to equation (13) that relates to the old series, the proportionate change in the bottom-up backcast total is given by:

$$G(BU(\hat{X}_t)) = \frac{\Phi_{TC}Y_{TC,t} + \Phi_{OI}Y_{OI,t}}{BU(\hat{X}_{t-1})} = G(Y_{TC,t})\frac{\hat{X}_{TC,t-1}}{BU(\hat{X}_{t-1})} + G(Y_{OI,t})\frac{\hat{X}_{OI,t-1}}{BU(\hat{X}_{t-1})} \quad (14)$$

(The derivation in, for example, the case of telecommunications comes from the relationship:

$$\frac{\Phi_{TC}Y_{TC,t}}{BU(\hat{X}_{t-1})} = \frac{\Phi_{TC}Y_{TC,t}}{\Phi_{TC}Y_{YC,t-1}}\frac{\Phi_{TC}Y_{YC,t-1}}{BU(\hat{X}_{t-1})} = G(Y_{TC,t})\frac{\hat{X}_{TC,t-1}}{BU(\hat{X}_{t-1})}.$$

Note this can also be written:

$$\frac{\Phi_{TC}Y_{TC,t}}{BU(\hat{X}_{t-1})} = G(Y_{TC,t})\frac{\Phi_{TC}Y_{TC,t-1}}{\Phi_{TC}Y_{TC,t-1} + \Phi_{OI}Y_{OI,t-1}} = G(Y_{TC,t})\frac{\Phi_{TC}v_{TC,t-1}}{\Phi_{TC}v_{TC,t-1} + \Phi_{OI}v_{OI,t-1}}$$

which collapses to $\frac{\Phi_{TC}Y_{TC,t}}{BU(\hat{X}_{t-1})} \Rightarrow G(Y_{TC,t})\frac{v_{TC,t-1}}{v_{TC,t-1} + v_{OI,t-1}} = G(Y_{TC,t})v_{TC,t-1}$

if the linking factors are the same.)

Equations (13) and (14) differ because of the weights. The proportionate change in the total historic series is a weighted average of the proportionate changes of each of the industry investments recorded in the old series, where the weights are the shares of the industry investments in the old total. By contrast, the proportionate change in the bottom-up backcast total is a weighted average of, as before, each of the proportionate changes of the industry investments recorded in the old series, but with different weights; the weights are the shares of the backcast industry investments in the bottom-up backcast total. The weights differ because the linking factors affect the weighting of the bottom-up backcast total and will diverge from the weights that govern the proportionate change in the old total save in the exceptional case in which the linking factors are the same. Because of the different weights, the bottom-up backcast total can part company with the history previously recorded by the old total series. This is what happened in the ONS backcasting project.

The scale of the historical re-write can be represented by subtracting the old series total from the bottom-up backcast total. To set this difference in context, bearing in mind that the values may inflate over time, the degree of historical re-write at time *t* due to the bottom-up method (denoted $\overset{BU}{HR}_t$) is expressed below as the difference between the two totals as a proportion of the old total:

$$\overset{BU}{HR}_t = \frac{BU(\hat{X}_t)}{Y_t} - 1 \tag{15}$$

Equation (15) can be combined with equation (10), which describes the bottom-up backcast total in terms of the linking factors and the old series:

$$\overset{BU}{HR}_t = \frac{\Phi_{TC} Y_{TC,t} + \Phi_{OI} Y_{OI,i}}{Y_t} - 1 \tag{16}$$

Equations (11) and (12), which describe the investment shares of the two industries in the old series total, can be used to substitute for the terms $\frac{Y_{TC,t}}{Y_t}$ and $\frac{Y_{OI,t}}{Y_t}$ in equation (16):

$$\overset{BU}{HR}_t = (\Phi_{TC} - 1) v_{TC,t} + (\Phi_{OI} - 1) v_{OI,t} \tag{17}$$

(The derivation uses the fact that $v_{TC,t} + v_{OI,t} = 1$.)

Equation (17) says that the extent to which the bottom-up splicing inflates or deflates the level of the old aggregate series is equal to the weighted sum of the difference of each linking factor from unity, the weights being the shares of the two industry's investments in the old series for total investment. The same formula can be used to derive an expression for the change over time in the degree of inflation or deflation. Denoting a change in a variable with the prefix Δ so that, for example, $\Delta v_{TC,t} = v_{TC,t} - v_{TC,t-1}$, and noting that of all the terms in equation (17) only the weights vary over time, it follows:

$$\Delta \overset{BU}{HR}_t = \Phi_{TC} \Delta v_{TC,t} + \Phi_{OI} \Delta v_{OI,t} \tag{18}$$

Equation (18) can be further simplified by noting that in the special two-industry case, the changes in the weights are equal and opposite:
$\Delta v_{OI,t} = -\Delta v_{TC,t}$.

If one chooses to substitute for the weight of the other industries, equation (18) becomes:

$$\Delta \overset{BU}{HR}_t = (\Phi_{TC} - \Phi_{OI})\Delta v_{TC,t} \tag{19}$$

This result means that the change in the degree of inflation or deflation of the level of the historic total is equal to the product of the difference between the linking factors and the change in the weight of the telecommunications industry in the old total. If the difference between the linking factors is very large, comparatively small changes in the investment share of the telecommunications industry combined with the method of bottom-up splicing would produce a marked inflation or deflation of the historic total.

The comparison can be made with the alternative splicing method in which the old total series is spliced to the new total series directly, leaving the detail of the industries investments to be filled in by other methods: a top-down approach. The other methods, not explored here, take the top-down spliced series as a benchmark total to which the sum of the industry investments must equate. Comparable to equations (5) and (6), which show how the splicing works for the two industries, the top-down backcast of total investment can be written:

$$\hat{X}_t = \Phi Y_t \tag{20}$$

where the total linking factor is the quotient of the values of the new and the old total series at the linking year:

$$\Phi = \frac{X_{t=\ell}}{Y_{t=\ell}} \tag{21}$$

The new and old series totals are the summations of the investments of the two industries, the telecommunications industry and the other industries, the latter taken as a group of identical industries, as described by equations (3) and (4). Using these equations, the total linking factor can be shown to be a weighted average of the linking factors of the two industries, where the weights are the investment shares of the industries in the old series evaluated at the linking year. These shares are constants in the splicing formulae and can be distinguished from the weights that vary over the historic period by dropping the subscript *t*:

For telecommunications:

$$\bar{v}_{TC} \equiv v_{TC,t=\ell} = \frac{Y_{TC,t=\ell}}{Y_{t=\ell}} \tag{22}$$

For the other (assumed identical) industries:

$$\bar{v}_{OI} \equiv v_{OI,t=\ell} = \frac{Y_{OI,t=\ell}}{Y_{t=\ell}} \tag{23}$$

By combining equation (4) with equation (21), the top-down linking factor can be written:

$$\Phi = \frac{X_{TC,t=\ell} + X_{OI,t=\ell}}{Y_{t=\ell}} = \Phi_{TC}\bar{v}_{TC} + \Phi_{OI}\bar{v}_{OI} \tag{24}$$

(The derivation in, for example, the case of telecommunications, comes from the relationship:

$\frac{X_{TC,t=\ell}}{Y_{t=\ell}} = \frac{X_{TC,t=\ell}}{Y_{TC,t=\ell}}\frac{Y_{TC,t=\ell}}{Y_{t=\ell}} = \Phi_{TC}\bar{v}_{TC}$, where, from equation (1), $\Phi_{TC} = \frac{X_{TC,t=\ell}}{Y_{TC,t=\ell}}$.)

Nothing can be said in principle about the likely size of the total linking factor in circumstances in which one of the industry linking factors is an aberrant outlier. If the industry in question is dominant, with a large linking-year weight, the total linking factor is itself likely to be large. If, on the other hand, aberrant linking factors are a peculiar feature of micro-data, perhaps as a result of pervasive measurement error, the total linking factor may remain small, the affected industry having a low linking-year weight. Top-down splicing may then inflate or deflate the historic total, and will do so in direct proportion to the total linking factor, as formalised by equation (20).

What can be said in principle, however, is the inflation or deflation impact of top-down splicing will not vary over time, and so will not create the ups and downs in the impact of splicing on the total series seen in the capital stock team's results. Put more formally, the top-down degree of historical re-write, denoted $\overset{TD}{HR_t}$,the equivalent of the bottom-up $\overset{BU}{HR_t}$ is simply the difference between the total linking factor and one:

$$\overset{TD}{HR_t} = \frac{\hat{X}_t}{Y_t} - 1 = \frac{\Phi Y_t}{Y_t} - 1 = \Phi - 1 \tag{25}$$

This measure of inflation or deflation of the historic record is a constant, and does not vary over time unlike the equivalent expressions (equations (18) or (19)) for the bottom-up backcast total:

$$\Delta \overset{TD}{HR_t} = 0 \tag{26}$$

This result springs directly from the fact that the backcast of any individual series created by splicing replicates the proportionate change of the old series, as equation (8) showed in the case of the telecommunications industry. However, this equality does not hold when a total series is formed as the bottom-up summation of spliced micro-series. As shown in the comparison of equations (13) and (14), the proportionate change in the bottom-up total has an implicit weighting of the proportionate changes of each individual series that differs from the weighting underlying the proportionate change in the old total.

It remains to formalise the degree to which a bottom-up backcast total can differ from a top-down backcast total. This difference at any point in time, denoted D_t, is defined as:

$$D_t = \frac{BU(\hat{X}_t)}{\hat{X}_t} - 1 \tag{27}$$

Equations (10) and (21), which express the backcast series in terms of the linking factors and old series, can be used to restate this difference between bottom up and top down:

$$D_t = \frac{\Phi_{TC} Y_{TC,t} + \Phi_{OI} Y_{OI,t}}{\Phi Y_t} - 1 \tag{28}$$

Equation (28) can be further simplified by using equations (11) and (12), which describe the investment shares of the two industries in the old series total, to substitute for the terms $\frac{Y_{TC,t}}{Y_t}$ and $\frac{Y_{OI,t}}{Y_t}$:

$$D_t = \frac{\Phi_{TC} v_{TC,t} + \Phi_{OI} v_{OI}}{\Phi} - 1 \tag{29}$$

Further simplification is possible. Equation (24) for the top-down linking factor can be combined with equation (29) to produce:

$$D_t = \frac{\Phi_{TC} v_{TC,t} + \Phi_{OI} v_{OI}}{\Phi_{TC} \bar{v}_{TC} + \Phi_{OI} \bar{v}_{OI}} - 1 \qquad (30)$$

Advantage can then be taken of the complementarity of the weights in equation (30), namely $v_{OI,t} = 1 - v_{TC,t}$ and $\bar{v}_{OI,t} = 1 - \bar{v}_{TC,t}$, to substitute for one of the two industries. If one chooses to substitute for the weight of the other industries, equation (30) becomes:

$$D_t = \frac{(\Phi_{TC} - \Phi_{OI}) v_{TC,t} + \Phi_{OI}}{(\Phi_{TC} - \Phi_{OI}) \bar{v}_{TC,t} + \Phi_{OI}} - 1 \qquad (31)$$

which can be further simplified:

$$D_t = \frac{\Phi_{TC} - \Phi_{OI}}{\Phi} (v_{TC,t} - \bar{v}_{TC}) \qquad (32)$$

(In the derivation, the single terms in the other industries' linking factor cancel out.)

This equation states that the difference between the bottom-up backcast total and the top-down backcast total, as a proportion of the top-down backcast, at any point in time depends on two things: the difference between the industry linking factors and the change in the weight of the telecommunications industry between that point in time and the linking year. Since $v_{TC,t}$ is the only term in this equation that varies over time, it follows that the change from one year to the next in the difference between the bottom-up and top-down totals depends on the same differences in the linking factors and the change from one year to the next in the weight of the telecommunications industry:

$$\Delta D_t = \frac{\Phi_{TC} - \Phi_{OI}}{\Phi} \Delta v_{TC,t} \qquad (33)$$

This equation is the natural complement to equation (19) which describes the change in the degree of inflation of deflation of the old total series brought about by bottom-up splicing.

The analytical framework described above can be used to understand why the capital stock team arbitrarily re-wrote the history of plant investment, and the way that re-write could have been avoided if the team had not rejected the top-down route.[49]

Section 3: An interpretation the capital stock team's re-write of history

The impact of the capital stock team's bottom-up splicing on the history of plant investment emerged from a highly detailed, multi-industry set of computations. An attempt analytically to decompose those results faces a number of practical difficulties. There is, first, a problem of data. The team's original dataset developed using specialised software has been mislaid, and other details of the team's advisory raise doubts about the faithfulness of the data the team re-created to the 2010-vintage investment data that it sought to emulate.[50] Appendix B provides more details of the data that can be reconstructed using published information. The second problem, which arises from the first, is the confinement of the analysis to just two industries – the telecommunications industry with all the remaining industries treated as a group – and so not the large number of individual industries upon which the capital stock team developed its backcasting project. Save for the telecommunications industry, for which the team reported a linking factor, the other industry-level linking factors for investment in plant are unknown. Within a two-industry framework, the linking factors for the other industries are represented by a single linking factor which is unlikely to equate with a properly weighted average of the true linking factors. In short, the limitations of the available data mean that the analysis has to proceed using two industries to represent many, with therefore only approximate results. Nevertheless, the 2-industry approximation, as it may be conveniently called, identifies why the aberrant linking factor for the telecommunications industry caused the capital stock team arbitrarily to re-write economic history.

Chart 2 shows the extent of the team's re-write of the history of plant investment and the 2-industry approximation. The blue tramline in the chart traces the impact of the team's bottom-up splicing expressed, not as in Chart 1 as a share of nominal GDP, but as a share of the historic series for total plant investment. This trace of the team's results is compared with the two-industry approximation traced by the solid red line. The latter conforms to two formulae shown before but repeated here for convenience:

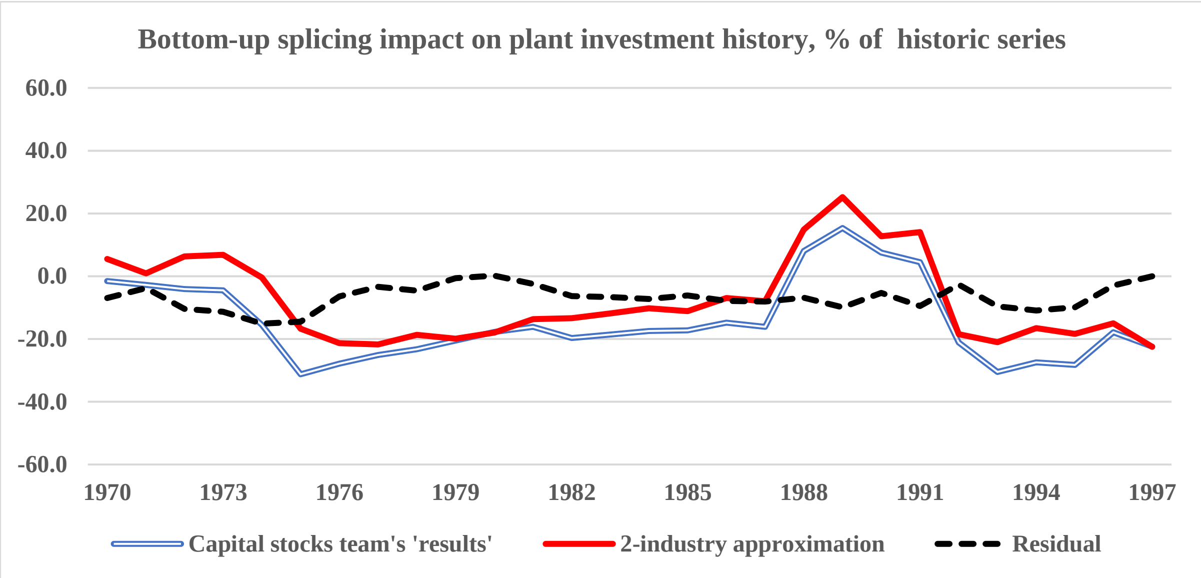
$$\overset{BU}{HR_t} = \frac{BU(\hat{X}_t)}{Y_t} - 1 \qquad (15)$$

and its alternative form:

$$\overset{BU}{HR_t} = (\Phi_{TC} - 1)v_{TC,t} + (\Phi_{OI} - 1)v_{OI,t} \qquad (17).$$

The chart shows that the two-industry results replicate quite closely the ups and downs in the capital stock team's results, but generally exceed them, save in 1997 when the two series are necessarily equal. This visual impression is confirmed by standard measures of association.[51]. The gap between the two is large but limited in its variation compared with the variation seen in the series themselves.

Chart 2: Inflation and deflation of the history of plant investment, % of historic series



Sources: See Chart 1; Appendix B. **Notes**: The Chart shows the impact of bottom-up splicing on the historic record of total plant investment expressed as a per cent of the historic total series. The 2-industry approximation uses equations (15) or (17). The residual equals the traced capital stock team's 'results' minus the 2-industry approximation and can be shown to equal $\left(\Omega_{OI,t}-\Phi_{OI}\right)v_{OI,t}$ where $\Omega_{OI,t}\equiv\dfrac{\Phi_{OI_1}Y_{OI_1,t}+\Phi_{OI_2}Y_{OI_2,t}+\Phi_{OI_3}Y_{OI_3,t}+...+\Phi_{OI_n}Y_{OI_n,t}}{Y_{OI,t}}$ for the full number n of other industries. The numerator of this expression is the capital stock team's post-linked plant investment series for industries other than the telecommunications industry and can be inferred residually from the team's splicing impact data (Exhibit 2), its post-linked telecommunications series (supplemented with a 1997 figure taken from ONS, Industry by Asset, (9 September 2016)) and the historic total taken from the 2010 Blue Book.

Table 2: Inflation and deflation of the history of plant investment, % point changes

Interval	Capital stock team's 'results'	2-industry approx.	Residual	Residual due to	
				$\Delta\Omega_{OI,t}v_{OI,t-1}$	$\left(\Omega_{OI,t}-\Phi_{OI}\right)\Delta v_{OI,t}$
1972 - 76	-23.7	-27.7	4.0	4.2	-0.3
1986 - 89	30.2	32.2	-2.0	-2.5	0.5
1989 - 92	-36.5	-43.7	7.2	7.3	-0.2

Sources: See Chart 2. **Notes**: See Chart 2. The first two data columns show percentage point changes in the impact of bottom-up splicing expressed as a per cent of the 'old' historic series for total plant investment. The two components of the residual decomposition, which is not unique, sum to the residual, subject to rounding error, and are given by the formulae $\Delta\Omega_{OI,t}v_{OI,t-1}$ and $\left(\Omega_{OI,t}-\Phi_{OI}\right)\Delta v_{OI,t}$ respectively. The sum of the two components is identically equal to $\Delta\left\{\left(\Omega_{OI,t}-\Phi_{OI}\right)v_{OI,t}\right\}$ where $\left(\Omega_{OI,t}-\Phi_{OI}\right)v_{OI,t}$ is the level of the residual. $\Omega_{OI,t}$ can be defined as $\Phi_{OI_1}v_{OI_1,t}+\Phi_{OI_2}v_{OI_2,t}+\Phi_{OI_3}v_{OI_3,t}+...+\Phi_{OI_n}v_{OI_n,t}$ so changes in it occur as a result of changes in the individual investment weights in the old total series, $Y_{OI,t}$

Table 2 examines how these data changed over the intervals first described in connection with Chart 1: the downswing between 1972 and 1976 that affected the record of the Barber Boom and Bust; the upswing between 1986 and 1989, associated with the Lawson Boom, and the subsequent downswing between 1989 and 1992, associated with the Lawson Bust. The 2-industry approximation, shown in the second data column, fairly closely replicates the swings in the capital stock team's results, shown in the first data column. The residual formed by subtracting the 2-industry approximation from the team's results is, by comparison with these swings, relatively small and contained. The final two data columns provide a decomposition of this residual. The columns show that the immediate cause of the disparity between the team's results and the 2-industry approximation is the changing investment shares of individual industries other than telecommunications, shares which would be measurable had the original data been publicly available. In sum, the 2-industry approximation offers insight into the sources of the ups and downs in the capital stock team's re-write of investment history, and the remaining discrepancy can be fully attributed to the necessity to approximate a far more complex set of calculations.

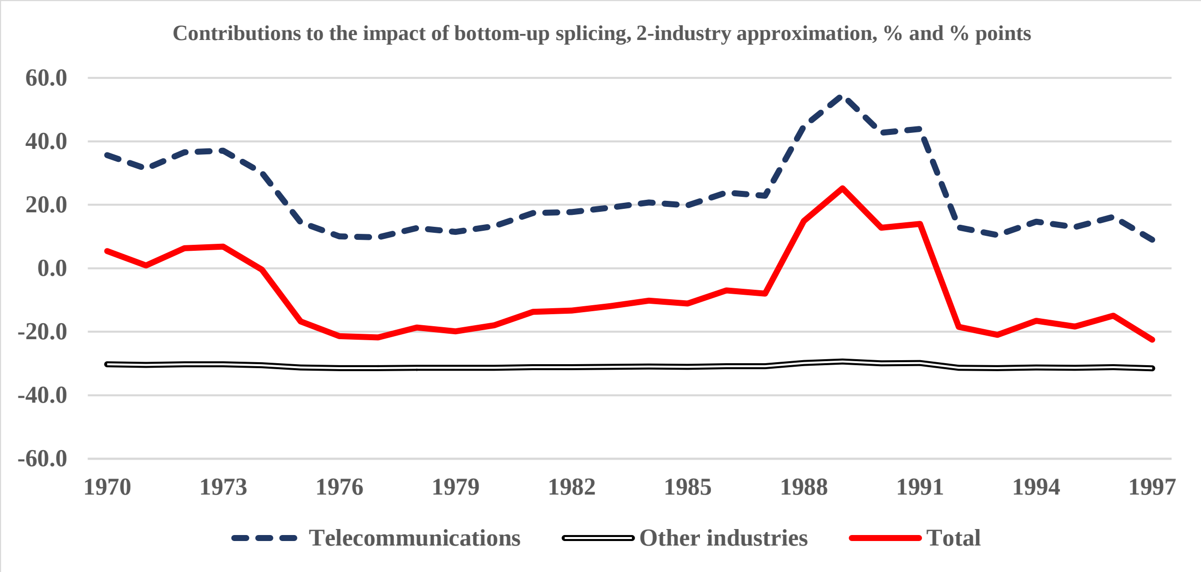
Table 3: Inflation and deflation of history – role of linking factors and changing weights

Interval	2-industry approximation			Residual
Formulae	$\Delta \overset{BU}{HR_t}$	$\Delta v_{TC,t}$	$\Phi_{TC} - \Phi_{OI}$	$\Delta\left\{\left(\Omega_{OI,t} - \Phi_{OI}\right)v_{OI,t}\right\}$
1972 - 76	-27.7	-3.90	7.1	4.0
1986 - 89	32.2	4.53	7.1	-2.0
1989 - 92	-43.7	-6.15	7.1	7.2

Sources: See Chart 2. **Notes**: See Chart 2, Table 2; Data columns record percentage point changes except for the difference in the linking factors, which is a constant. The formulae refer to equation 19. The changes in the historic share of plant investment by the telecommunications industry, $\Delta v_{TC,t}$, is shown to two decimal places so that it can be easily verified that $\Delta \overset{BU}{HR_t} = (\Phi_{TC} - \Phi_{OI})\Delta v_{TC,t}$ where Φ_{TC} = 7.8 (as in Mosquera et al (2 July 2014), 'Figure 3', p. 10) and Φ_{OI} = 0.7 (own estimate). A Φ_{OI} below one fits with the impact of the capital stock team's splicing, which was generally to deflate the historic record of total plant investment while inflating the corresponding record of the telecommunications industry.

Table 3 decomposes the 2-industry approximation of the changes in the degree of inflation and deflation of the historic record. The relevant equation 19, repeated here for convenience, shows that these changes can be represented as the product of the difference between the linking factors for the telecommunications industry and the other industries treated as a group, $\Phi_{TC} - \Phi_{OI}$, and the change in the investment share of the telecommunications industry $\Delta v_{TC,t}$ such that $\Delta \overset{BU}{HR_t} = (\Phi_{TC} - \Phi_{OI})\Delta v_{TC,t}$ (19). For example, the 2-industry approximation of the inflation of the historic record, expressed as a per cent of that record, that occurred between 1986 and 1989, coinciding with the Lawson Boom, is 32 percentage points. The approximation overstates the capital stock team's 'result' by 2 percentage points. Over this three-year period, the investment share of the telecommunications industry rose by just 4½ percentage points. But the impact of this small change in investment share is magnified seven-fold as a result of the spectacularly large and aberrant linking factor (7.8) that the capital stock team reported for the telecommunications industry. It is estimated that the linking factor for the other industries treated as a group was just 0.7. The change in the investment share is therefore multiplied by 7.1, leading to the gyrations in the impact of splicing on the historic record of the telecommunications industry seen in Charts 3 and 4. By comparison, the impact on the record of the other industries is stable.

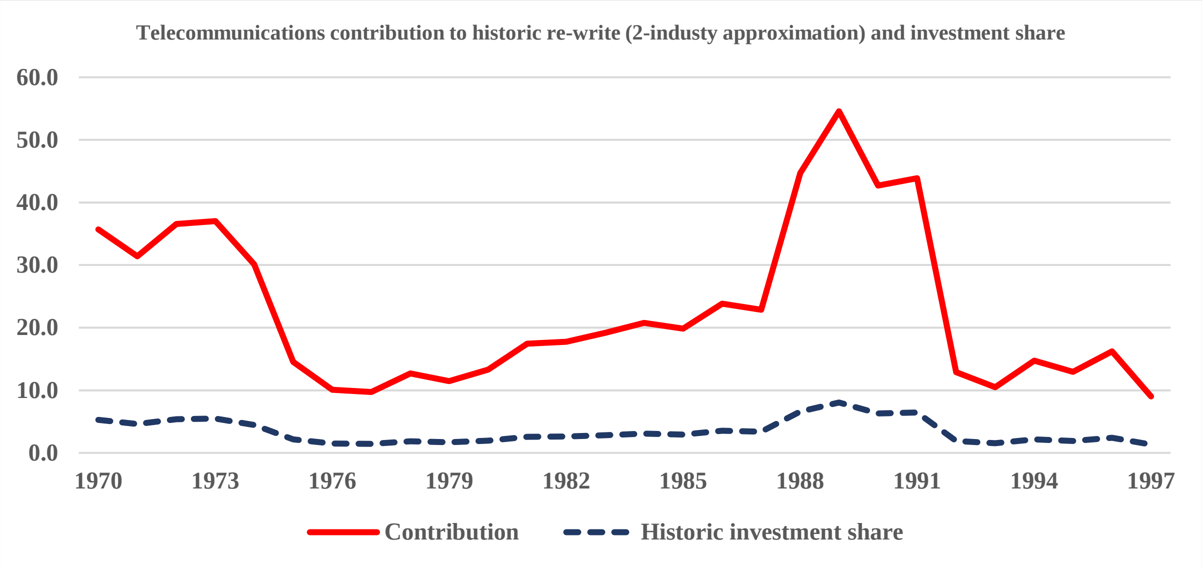
Chart 3: Contributions to the historical re-write, % of historic plant investment series



Sources and notes: see Chart 2; Table 2. The chart traces the left-hand side and separately the two components of the righthand side of formula 17:

$$\overset{BU}{HR_t} = (\Phi_{TC} - 1)v_{TC,t} + (\Phi_{OI} - 1)v_{OI,t} \,.$$

Chart 4: Telecommunications industry contribution to the historical re-write, % of historic plant investment series



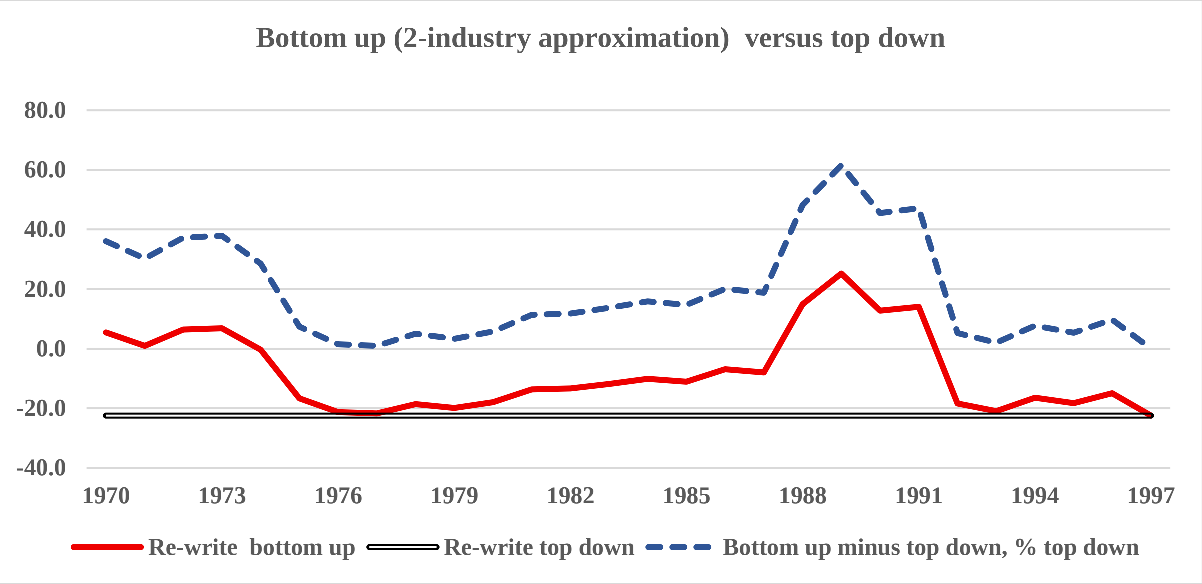
Sources and notes: see Chart 2; Table 2. The chart traces the telecommunications component of the righthand side of formula 17 and telecommunications investment share, $(\Phi_{TC} - 1)v_{TC,t}$ and $v_{TC,t}$ respectively. Note the vertical scale differs from that used in both Charts 2 and 3.

The capital stock team could have avoided this implausible re-write of economic history had it chosen to splice the data at a higher level of aggregation, starting with a total series for, say, the asset class of investment in plant and working top down to the detail that the team required for its capital stock calculations. Ways of in-filling such detail while using a total series as a controlling benchmark are commonplace, ranging from simple proportional methods to more sophisticated balancing techniques.[52] Unlike the variable impact of bottom-up splicing on the historic total series, a top-down splice has a constant impact, governed by a single linking factor. In the case of the ONS backcast of plant investment, this linking factor is estimated to have been about 0.8, implying a backcast history that is always around 20 per cent below the old total series. In the 2-industry approximation, the difference between this backcast of history and the bottom-up version varies as a magnified version of the changes in the investment share of one of the two industries. Equation (33), repeated here for convenience, presented this result using the telecommunication industry share:

$$\Delta D_t = \frac{\Phi_{TC} - \Phi_{OI}}{\Phi} \Delta v_{TC,t} \tag{33}$$

The magnification factor, $\frac{\Phi_{TC} - \Phi_{OI}}{\Phi}$, has a value of 9.2 in this example, again the result of the aberrant linking factor that the capital stock team reported for the telecommunications industry. The gyrations in this measure of avoidable distortion are clearly visible in Chart 5.

Chart 5: Re-writing history - bottom up versus top down, % of history and top down



Sources and notes: see Chart 2, Chart 4; Table 2. The solid red line and the black tramline trace the degree of historic re-write using equations (17) and (25) respectively: $\overset{BU}{HR_t} = (\Phi_{TC} - 1)v_{TC,t} + (\Phi_{OI} - 1)v_{OI,t}$ and $\overset{TD}{HR_t} = \Phi - 1$. The dashed blue line traces the difference between the bottom-up and top-down backcast totals, as a per cent of the top-down backcast, using equation (32): $D_t = \frac{\Phi_{TC} - \Phi_{OI}}{\Phi}(v_{TC,t} - \bar{v}_{TC})$. Note the vertical scale differs from that used in Chart 4.

.

Conclusion

Eleven years ago, the ONS attempted to merge two sets of UK time series data that were separated by discontinuities. The merged data describe a major macro-economic concept – total capital investment – over a period, known by the ONS as the 'historic' period, between 1948 and 1996. The capital stock team responsible for this project used old data that were prone to measurement error, quite possibly severely understating the surge in council house building that occurred after the Second World War. Other data problems call the results into question. The team also chose to join the old data to new national accounts data using splicing, a common technique, and did so at a fine level of detail. Total investment was backcast 'bottom up', by adding up the detailed backcast figures. Extraordinary differences between some of the old and new data series should have alerted the capital stock team to the deficiency of its chosen approach, but the team's aim lay elsewhere. It was intent on reconstructing suspended data for capital stocks; the backcast investment figures were a means to that end. Yet the backcast investment data also became the series for total investment in the main historic national accounts. The result was an implausible re-write of Britain's economic history that has been embedded in the national accounts ever since.

Some lessons might be usefully learnt. Three suggestions relate to splicing:

1. Sense check the results of splicing for plausibility, noting that revisions due to splicing are not the result of better sources of information or of improved statistical methods. If the splicing distorts economic history, find another way.
2. Pre-test the old and new data to establish whether there exist stable relationships over a period in which old and new data co-exist. If instability is found, find another way.
3. If aberrant 'linking factors' are found, perhaps as a result of pervasive measurement error in micro-series, do not backcast bottom up. Try splicing top down from a larger aggregate, and if the results do not seem implausible fill in the details another way.

A more general lesson concerns what in statistical circles is often referred to as 'user engagement'. A recent well-attended conference, the January 2025 UK Statistics Assembly, considered the way the ONS engaged with a wide variety of 'users', including the likes of this author and independent scholar Anne Harrison.[53] A number of contributors, notably the statistics consultant Simon Briscoe, were highly critical of the way the ONS responded to informed comment by outsiders.[54] ONS engagement on the capital investment question provides an additional case study worthy of consideration. A potted history will suffice.[55]

It was Harrison who first alerted the ONS that the capital investment figures appeared far too low in the immediate post-war period. She did so in October 2018.[56] A year later, after much correspondence from me and an ONS conference attended by Harrison, a senior ONS statistician remarked on seeing the figures, and after consultation with an ONS colleague, that there was 'clearly a problem here that we will certainly resolve in BB20 [Blue Book 2020]'.[57] The ONS began an investigation. Its report in March 2020 proposed an adjustment to the investment data in the immediate post-war years to be followed by a more comprehensive review.[58] In the September of that year, I drew the attention of another branch of the ONS to the likelihood that the too-low-level of total investment recorded in the national accounts in the immediate post-war years arose from a probable mistake in its perpetual inventory record of investment in dwellings. An investigation was promised, but the dwellings figures were not revised. In July 2021, the main ONS investigation of the historic investment data was abandoned. Brief notification was delivered verbally, with regret, but not in writing.

It became increasingly apparent that even expert users were flummoxed by the peculiar official data, or, alternatively, were wholly unaware of its deficiencies. As Robert Jump found, and discussed with me in 2023, experts in the camp of the bemused (my expression, not Jump's) included the ONS statisticians with whom he had consulted about the odd profits data in the late-1960s. Several authoritative articles on long-run investment trends were written without any awareness of the data problems, unsurprisingly so because the ONS had not publicly informed the expert authors, or anyone else.[59] It seems likely that users of the national accounts, seasoned as well occasional, were instead reassured by the official statistical kitemark of trustworthiness and quality. As previously noted, the suspect investment and related macroeconomic data enjoyed the imprimatur of being 'National Statistics'.[60]

My earlier paper was intended to raise awareness. Published in September 2024, it was sent to the ONS with little expectation of a positive response. Instead, to its credit, the ONS planned a new investigation, which began in January 2025. Following a question from me, the ONS regulator, the Office for Statistics Regulation, began its own 'casework' enquiry. The ONS has since notified users of its business investment time series of its investigation and of its intention to 'provide further updates by 30 September 2025 as we consolidate our [the ONS's] investigations'. Appendix C provides the full ONS notification available at the time of writing. The notification would be easily missed by general users of the national accounts, but it is better than the limited, if not complete absence of, public notification that the ONS had in mind prior to the intervention of the OSR.

It seems reasonable to assume that the ONS investigation will not result in actual remedy of the investment data until after September 2025. If so, rectification could arrive more than seven years after Harrison's original alert. Partial rectification allied to the troublesome dwellings investment figures is a possibility: the ONS notification highlights dwellings data as an area 'for further investigation'. But there are indications that the ONS does not regard as a mistake the capital stock team's bottom-up splicing method, or accept the need to sense check the results of splicing to ensure that well-established economic history is not arbitrarily re-written. No effective remedy is also a possible outcome. The traverse from the March 2020 good intentions to the July 2021 abandonment provides a precedent.

An 'independent report' on the UK Statistics Assembly suggested that the first of four high-level priorities for the UK governing body, the UK Statistics Authority, should be to '[r]e-invigorate sustained and effective user engagement'.[61] Any lesson that might be learnt from the history reviewed here should perhaps begin by asking why the ONS omitted to inform users of the problems with the investment data that, in October 2019, provoked the 'clearly a problem here' reaction of a senior ONS official. The absence of public warning can be regarded as a breach of the Code of Practice that governs the expected behaviour of the agency.[62] A second, more perplexing, question is this. Why has 'user engagement' over six years and counting failed yet to deliver a set of historic UK national accounts on which economic historians, policy makers and other researchers can confidently rely?

Appendix A: Splicing top down and bottom up – a formal description

A new aggregate time series, X_t, has a time span $t = \ell,\ \ell+1,\ \ell+2\ ...\ \Omega\ ...\ T$, denoted $t \in \{\ell,\ ...\ \Omega,\ ...\ T\}$. The related old aggregate series, Y_t, has a time span $t \in \{0,\ ...\ \ell,\ ...\ \Omega\}$. New and old series comprise *n* sub-series such that

$$X_t \equiv \sum_{i=1}^{i=n} X_{i,t} \qquad \text{(A1)}$$

$$Y_t \equiv \sum_{i=1}^{i=n} Y_{i,t} \qquad \text{(A2)}$$

These summations mean that the data in mind are additive current price measures rather than volume measures, such as constant price or chained volume measures (CVM), in which the aggregate is not usually the simple sum of sub-series.[63] The limitations imposed by definitions (A1) and (A2) are not as restrictive as they may appear, however. The backcasting of volume series may rely first on the backcasting of equivalent current price series, which are then deflated using backcast price deflators. The 2014 ONS exercise was of this type.

The old and new aggregate series and sub-series overlap in the interval $t = \ell$ through to $t = \Omega$, an interval that can be used to help ascertain whether there exist stable relationships between the old and new data. If there is no evidence of instability, the splicing of the old and new data can be entertained as a means to reconstruct how the new series would have appeared had it existed in the 'historic' period $t \in \{0,\ ...\ \ell-1\}$. Consider the identity $X_t \equiv \Phi_t Y_t$, where X_t is the 'true' backcast version of the new aggregate series and Φ_t is the ratio of the true backcast series to the old series: $\Phi_t \equiv \dfrac{X_t}{Y_t}$, $t \in \{0,\ ...\ \ell-1\}$. Time series $\{\Phi_t\}$ is not known, however. The splicing solution replaces $\{\Phi_t\}$ with a constant, the quotient of the new and old series at time $t = \ell$. For the aggregate series: $\Phi_t \Rightarrow \Phi = \dfrac{X_{t=\ell}}{Y_{t=\ell}}$ .This quotient is sometimes known as the 'linking factor' and the time $t = \ell$ as the 'linking point'. Equivalent quotients can be calculated for the *n* sub-series: $\Phi_{i,t} \Rightarrow \Phi_i = \dfrac{X_{i,t=\ell}}{Y_{i,t=\ell}}$.

The resulting top-down *estimate* of the 'true' backcast aggregate series, $\hat{X}_t$, over the historic period is defined as:

$$\hat{X}_t = \Phi Y_t\ ,\ t \in \{0,\ ...\ \ell-1\} \qquad \text{(A3)}$$

The equivalent bottom-up estimate of the aggregate over the historic period is defined as the sum of the backcast sub-series:

$$\sum_{i=1}^{i=n} \hat{X}_{i,t} = \sum_{i=1}^{i=n} \Phi_i Y_{i,t} \text{ , } t \in \{0, \ldots \ell - 1\} \tag{A4}$$

Two simple properties follow. First, as a result of the linking factor definitions, the levels of the backcast series extended to the linking point are equal to the corresponding levels of the new series:

Top-down aggregate: $\hat{X}_{t=\ell} = \Phi Y_{t=\ell} = \left(\frac{X_{t=\ell}}{Y_{t=\ell}} \right) Y_{t=\ell} = X_{t=\ell}$;

Bottom-up aggregate: $\sum_{i=1}^{i=n} \hat{X}_{i,t=\ell} = \sum_{i=1}^{i=n} \Phi_i Y_{t=\ell} = \sum_{i=1}^{i=n} \left(\frac{X_{i,t=\ell}}{Y_{i,t=\ell}} \right) Y_{i,t=\ell} = \sum_{i=1}^{i=n} X_{i,t=\ell} = X_{t=l}$.

Second, the growth rates of the backcast aggregate and of the backcast sub-series are equal to the growth rates of the corresponding old series:

Top-down aggregate, ratio to previous period: $\frac{\hat{X}_t}{\hat{X}_{t-1}} = \frac{\Phi Y_t}{\Phi Y_{t-1}} = \frac{Y_t}{Y_{t-1}}$;

Sub-series, ratio to previous period: $\frac{\hat{X}_{i,t}}{\hat{X}_{i,t-1}} = \frac{\Phi_i Y_{i\,t}}{\Phi_i Y_{i,t-1}} = \frac{Y_{i,t}}{Y_{i,t-1}}$.

However, the growth rate of the backcast bottom-up aggregate is not so simply stated. Its growth rate is a variable weighted average of the growth rates of the old sub-series, where the weights are the shares of the backcast, spliced sub-series in the bottom-up aggregate:

Bottom-up aggregate, ratio to previous period:

$$\frac{\sum_{i=1}^{i=n} \hat{X}_{i,t}}{\sum_{i=1}^{i=n} \hat{X}_{i,t-1}} = \frac{\sum_{i=1}^{i=n} \Phi_i Y_{i,t}}{\sum_{i=1}^{i=n} \Phi_i Y_{i,t-1}} = \sum_{i=1}^{i=n} \frac{Y_{i,t}}{Y_{i,t-1}} \frac{\Phi_i Y_{i,t-1}}{\sum_{i=1}^{i=n} \Phi_i Y_{i,t-1}} .$$

In the historic time period before the linking point, the top-down and bottom-up methods of deriving the backcast aggregate series are likely to give different results. A formal measure of this divergence can be constructed by subtracting the top-down backcast aggregate (equation A3) from the bottom-up backcast aggregate (equation A4) and expressing the difference as a proportion of the top-down aggregate, so to set the size of the difference in context. The divergence measure, D_t , is defined as:

$$D_t = \frac{\sum_{i=1}^{i=n} \hat{X}_{i,t}}{\hat{X}_t} - 1 \qquad \text{(A5)}$$

Equation (A5) can be re-formulated in terms of the linking factors and the old series using equations (A3) and (A4):

$$D_t = \frac{\sum_{i=1}^{i=n} \Phi_i Y_{i,t}}{\Phi Y_t} - 1 \qquad \text{(A6)}$$

which can be further simplified as:

$$D_t = \sum_{i=1}^{i=n} \left(\frac{\Phi_i - \Phi}{\Phi} \right) v_{i,t} \qquad \text{(A7)}$$

where $v_{i,t} \equiv \frac{Y_{i,t}}{Y_t}$ and $\sum_{i=1}^{i=n} v_i = 1$.

Equation (A7) expresses the proportionate divergence between top-down and bottom-up backcast aggregate series in terms of the dispersion of each sub-series linking factor around the aggregate linking factor and the varying shares (or weights), $v_{i,t}$, of the old sub-series in the aggregate old series. It follows that there would be no divergence if all the sub-series linking factors were the same. This is so because the aggregate linking factor is itself a weighted average of the sub-series linking factors, where the weights, which sum to one, are the shares of the old sub-series in the old aggregate series evaluated at the linking point:

$$\Phi = \sum_{i=1}^{i=n} \Phi_i \overline{v}_i \qquad \text{(A8)}$$

where $\overline{v}_i \equiv v_{i,t=\ell} = \frac{Y_{i,t=\ell}}{Y_{t=\ell}}$.[64] Were all the sub-series linking factors to take the same value, say $\tilde{\Phi}$, it would follow that $\Phi \Rightarrow \tilde{\Phi} \sum_{i=1}^{i=n} \overline{v}_i = \tilde{\Phi}$; so that $\Phi_i - \Phi \Rightarrow \tilde{\Phi} - \tilde{\Phi} = 0$, $i \in \{1, \ldots n\}$.

It is possible to re-express the divergence formalised in equation(A7) to bring out more precisely the roles played by the dispersion of the linking factors and the changing weights of the old sub-series in the historic period. This derivation relies on the property that the weights sum to unity, making it possible to treat one of the sub-series as a numeraire. The choice of numeraire is arbitrary, and does not affect the calculated divergence between top-down and bottom-up backcast aggregations, but does affect the form the equation takes.

The reformulation involves two steps. The first step re-writes equation (A5) using equation (A6) and the expression for the aggregate linking factor, equation (A8):

$$D_t = \frac{\sum_{i=1}^{i=n} \Phi_i v_{i,t}}{\sum_{i=1}^{i=n} \Phi_i \overline{v}_i} - 1 \qquad \text{(A9)}$$

The summations can then be re-arranged by using the weight properties: $v_{j,t} = 1 - \sum_{i \neq j} v_{i,t}$ and $\overline{v}_j = 1 - \sum_{i \neq j} \overline{v}_i$ where j denotes the sub-series arbitrarily chosen to be the numeraire:

$$D_t = \frac{\sum_{i \neq j} \Phi_i v_{i,t} + \Phi_j (1 - \sum_{i \neq j} v_{i,t})}{\sum_{i \neq j} \Phi_i \overline{v}_i + \Phi_j (1 - \sum_{i \neq j} \overline{v}_i)} - 1 \qquad \text{(A10)}$$

Simplified, equation (A10) becomes:

$$D_t = \sum_{i \neq j} \left(\frac{\Phi_i - \Phi_j}{\Phi} \right) (v_{i,t} - \overline{v}_i) \qquad \text{(A11)}$$

The proportionate divergence at any point in time during the historic period between the backcast aggregates using bottom-up as opposed to top-down splicing depends on two properties of the data: first, the dispersion of the sub-series linking factors and, second, the change in the weights of the old sub-series between that point in time and the linking point. The divergence would be zero were all the sub-series linking factors the same, as previously noted, or were the weights of the old sub-series to remain unchanged from the linking point back through the historic period. It also follows that $D_{t=\ell} = 0$ as $v_{i,t} - \overline{v}_i \Rightarrow v_{i,t=\ell} - \overline{v}_i = 0$.[65]

A further corollary is that the change from one period to the next in the divergence between the spliced aggregations depends on the dispersion of the linking factors and the relative rates of growth of the old sub-series that cause their weights to change. Noting that the linking factors and linking-point weights are constants, the first-difference of equation (A11) can be written:

$$\Delta D_t = \sum_{i \neq j} \left(\frac{\Phi_i - \Phi_j}{\Phi} \right) \Delta v_{i,t} \tag{A12}$$

where Δ denotes the one-period change in the associated variable.

Relevant to the data analysis in the paper is the special case in which $n = 2$. In the two sub-series case, $D_t = \left(\frac{\Phi_1 - \Phi_2}{\Phi} \right) (v_1 - \overline{v}_1)$ and $\Delta D_t = \left(\frac{\Phi_1 - \Phi_2}{\Phi} \right) \Delta v_1$.

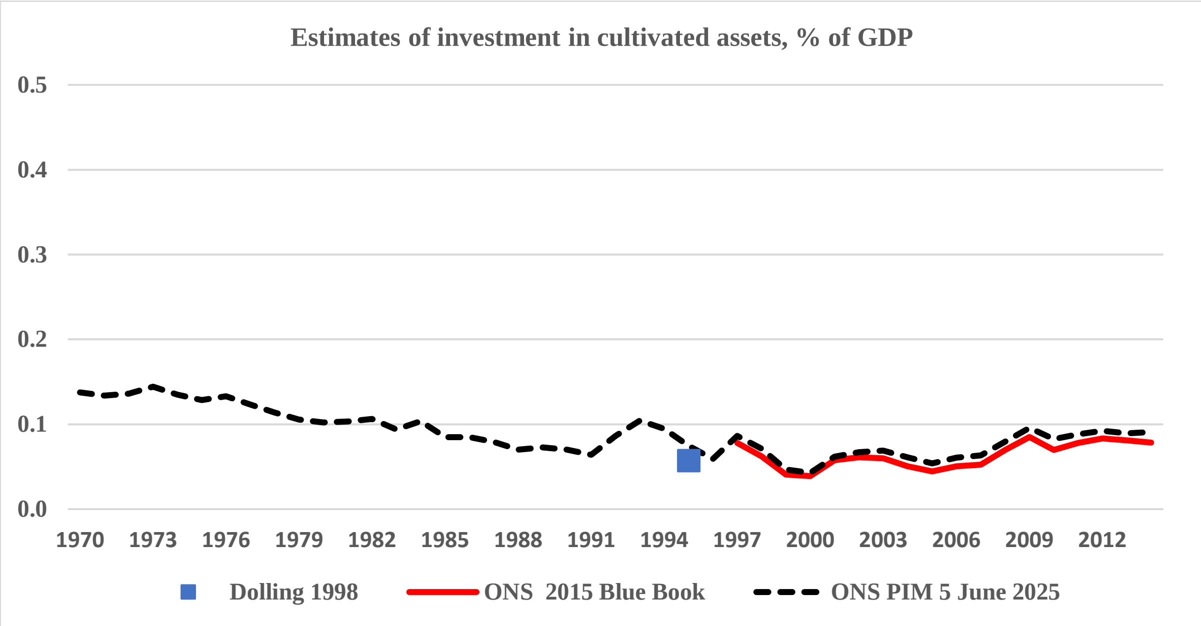
Appendix B: Data construction

The data required for the decomposition of the capital stock team's calculations are partly drawn from its advisory and partly estimated, with the aim of remaining as close as possible to the team's original, but unavailable, dataset.[66] Data can be extracted by digitisation of the relevant 'figures' in the advisory quantifying the impact of the team's splicing on what the advisory refers to as investment in 'plant' between 1970 and 1997 and the corresponding figures, together with a linking factor, for the telecommunications industry up to 1996. A 1997 figure of similar vintage to the team's 'post-linked' telecommunications data can be obtained from the ONS industry-by-asset dataset released in 2016 and consistent with the 2015 Blue Book.[67] By residual, it is possible to infer the impact of the team's splicing on the record of plant investment by all industries apart from the telecommunications industry.

Not provided in the advisory are the team's time series data recovered from the 2010 capital stocks publication, either before or after conversion to standard industrial classification 2007. To infer other data required for the analytical decomposition, reliance is placed on published national accounts investment data under the heading 'other machinery and equipment and cultivated assets', as presented in the 2010 Blue Book. It is acknowledged that the national accounts data may be inconsistent with the PIM data that the capital stock team sought to re-create, but there is no alternative source. The national accounts data provide the closest approximation to the ideal. There are two other matters that arise from a close reading of the capital stock team's advisory: first, whether or not cultivated assets are to be included in its 'plant' category and, second, the accuracy of the capital stock team's re-creation of the investment data in the 2010 capital stocks publication.

The advisory states that 'cultivated assets e.g. livestock, orchards were '[n]ot included' in the 'previous publication'.[68] By 'previous publication, the advisory is referring to '[e]stimates of GFCF [gross fixed capital formation] based on pre-Blue Book 2013 methodology' that were used in the 'estimates of capital stocks and consumption of fixed capital'. The inference is that cultivated assets were not included in what the 2010 capital stocks publication called 'plant and machinery' and so were not part of the 2010 dataset re-created by the capital stock team.[69] The electronic dataset accompanying the 2010 capital stocks publication, however, refers to the publication's 'plant and machinery' series as 'Total Other Machinery & Equipment & Cultivated Assets', with data available from 1996 that are only trivially different in the 1996 to 2005 interval to the 2010 Blue Book data which include cultivated assets and are the same as those data from 2006 to 2009. Similar minor discrepancies in the 1996 to 2005 interval are seen in other investment asset categories save for the cost of transfers. The discrepancies are likely to be the result of revisions that occurred between different ‘2010’ vintages of the PIM and national accounts datasets. The inference drawn, contrary to the statement in the capital stock team’s advisory, is that 2010 capital stocks publication data included cultivated assets within what it labelled ‘plant and machinery’ investment, but that there were no corresponding estimates of the stock of cultivated assets.

Chart B1: Investment in cultivated assets – alternative estimates



Sources: ONS, UK Economic Accounts (28 March 2025); Dolling (1998), p. 103; ONS Industry by Asset (9 September 2016); ONS PIM (5 June 2025). **Notes**: The PIM dataset and current price GDP data are the latest vintages available at the time of calculation.

As a test of this inference, estimates have been made of the magnitude of investment in cultivated assets, and the analytical decomposition re-run with this asset class excluded from the ‘other machinery and equipment and cultivated assets’ national accounts category. Chart B1 shows estimates of cultivated assets, traced as shares of latest estimates of current price GDP. Although the information is limited, the diminutive scale of the asset class provides reassurance that no material error would arise from basing a backcast over the 1970 to 1996 interval on the latest PIM data in combination with the 2015 Blue Book industry-by-asset dataset. The backcast is not inconsistent with Dolling’s estimate of the equivalent figure in 1995.[70] As the scale of this class of asset is so diminutive, it is no surprise that its exclusion made no appreciable difference to the analytical decomposition. The results for ‘plant’ reported in the main text include cultivated assets.

Doubts over the accuracy of the capital stocks' team's re-creation of the 2010 capital stocks publication investment data are raised by Table B1. It records in the first data column the 1997 shares, rounded as presented to the nearest whole number, of each asset class in a total of 'certain assets' – therefore probably not all assets – in what the advisory's 'Figure 22', a pie-chart, describes as the '2010 publication'. These shares are compared in the remaining columns with the 1997 shares of total investment accounted for by different asset classes as seen in other publications that report data (Jones et al. (14 March 2014)) or are data (capital stocks publication; the Blue Book) of 2010 vintage. Two differences stand out: the published shares for dwellings are considerably above, and the published shares for other machinery and equipment are considerable below, those attributed by the advisory to the '2010 publication'. Neither the effects of rounding nor the expression of shares in relation to a total that probably excludes some minor items can account for such large discrepancies. Since the advisory is describing a '2010 publication', which is based on the 2003 standard industrial classification, the conversion of the data to the 2007 standard industrial classification should have played no part in creating the discrepancies shown in Table B1.

The possibility exists that the capital stock team may not have accurately re-created the 2010 capital stocks publication dataset. The team had at its disposal data on investment in constant prices at a very fine level which had to be reflated using 'the same (historic) deflators which created the KP [constant price] series in 2010' and 'at the same level of detail'.[71] Such a mass-reconstruction process involving the matching of constant price and price deflator series may have led to errors which surfaced in the discrepancies from the published and, save for some trivial differences, identical estimates in the 2010 capital stocks publication and the 2010 Blue Book. If errors were made, they would likely have affected the team's estimated linking factors as well as the historic 'old' series used for the backcasting.

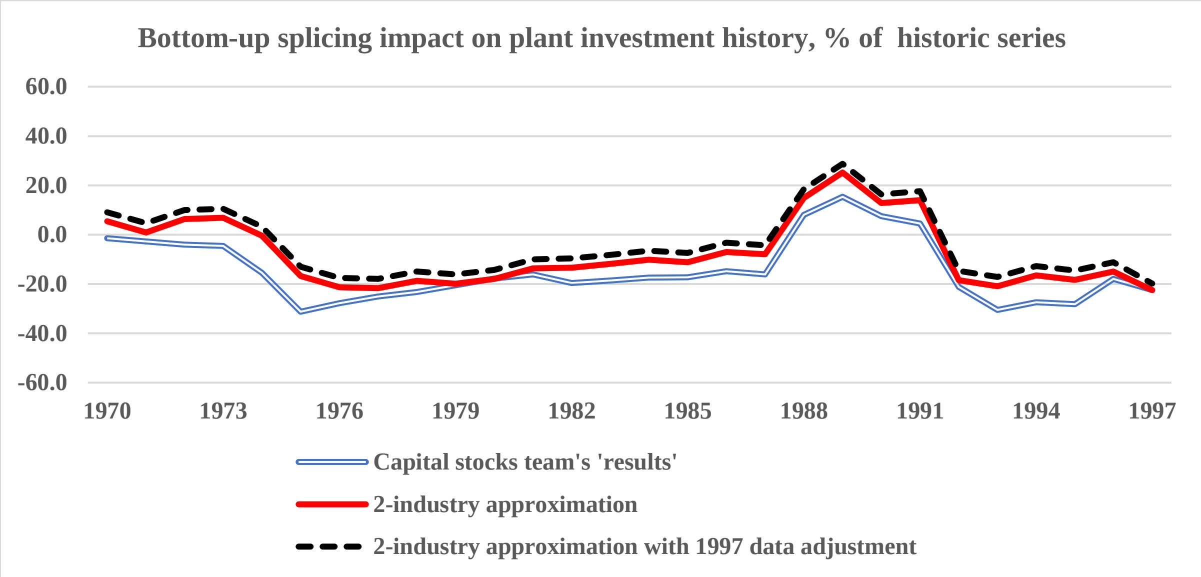
Table B1: Investment asset shares in 1997 – 2010 vintage estimates

Asset class per cent of total	Publication: Mosquera	Jones	Cap. Stocks	Blue Book
Buildings	28	25.9	25.8	25.9
Dwellings	11	17.3	17.3	17.3
Transport	9	9.1	9.1	9.1
Machinery & equipment*	42	37.2	37.3	37.2
Software	4			4.6**
Entertainment	1			
Mineral exploration	1			
Transfer costs	4	4.1	4.1	4.1
Total of 'certain assets'	100			
Intangible		6.5	6.5	6.5
Total all assets		100	100	100***

Sources: Mosquera et al. (2 July 2014), Figure 22, p. 23; Jones et al. (14 March 2014), Annex A, pp. 28-33; 'Cap. Stocks' – Ombundsen et al. (31 July 2010); 'Blue Book' – Dye and Sosimi (eds) (30 July 2010); Oulton and Srinivasan (2003), Table C.2, p. 81; UK Economic Accounts (28 March 2007, 28 March 2008). **Notes:** 'certain assets' refers to the title of Figure 22, in Mosquera et al. (2 July 2014): 'Share of current price GFCF [gross fixed capital formation] attributable to certain assets in 1997 (2010 publication)'; * National accounts category 'Other machinery and equipment including cultivate assets'. It is assumed that this category corresponds to the asset class labelled 'other machinery' in Mosquera et al. (2 July 2014), Figure 22. In the 2010 Blue Book, investment in cultivated assets in 1997 accounted for an estimated 0.5 per cent of total investment; ** This estimate is derived from Oulton and Srinivasan (2003) updated by revisions to software in the 2007 Blue Book (equal to the revisions to the gross operating surplus of private sector corporations), scaled by the very similar level of investment in intangibles in the 2007 and 2010 Blue Books; *** Sum of above excluding the individual software estimate which is included in intangible investment.

As the precise status and role of the data reported in the advisory's 'Figure 22' are unclear, it cannot be decisively concluded that there were basic calculation errors that affected the team's splicing. To examine sensitivities, the analytical decomposition reported in the main text was re-run after aligning – with allowance for rounding error - the 1997 figure for plant investment with the advisory's reported 42 per cent investment share. Chart B2 shows that the impact on the analytical decomposition was trivial, displacing slightly upwards the trace of the analytical results produced without such an adjustment.

Chart B2: Inflation and deflation of history due to bottom-up splicing



Sources: See Chart 1, Mosquera et al. (2 July 20140; Dye and Sosimi (eds) (30 July 2010), ONS Industry by Asset (9 September 2016). **Notes**: '2-industry approximation' is the analytical decomposition based on an attempted reconstruction of the Mosquera et al. (2 July 2014) 'old' (2010 Capital stocks publication) series with industries outside of the telecommunications industry treated as an individually-identical group. The adjusted version aligns the level of the total 'old' series in 1997 with that implied by 'Figure 22' in Mosquera (2 July 2014), p. 23.

Appendix C: ONS Business investment time series notice, 15 May 2025.

Dataset

Business investment time series

Contact: Gross Capital Formation team

Release date: 15 May 2025

Next release: 30 June 2025

Dataset ID: CXNV

About this Dataset

Detailed breakdown of business investment by industry and asset, in current prices, chained volume measures, non-seasonally adjusted and seasonally adjusted.

Edition in this dataset

csv (516.3 KB) | xlsx (397.0 KB) | structured text (519.2 KB)

Previous versions of this data are available.

Important notes and usage information

There are differences between the business investment chained volume measure seasonally adjusted estimates in this dataset and the official estimates published in the Business investment release. These differences are because of the different seasonal adjustment approaches. In this dataset, industry-level series are seasonally adjusted and aggregated. In the Business investment release, the asset-level series are seasonally adjusted and aggregated. This can give different results for the aggregated series. The estimates in the Business investment release should be considered the official statistics.

Please note that not all series will sum to totals because of rounding. These differences are well within the uncertainty of the estimates.

Based on specific user feedback and engagement, we are reviewing our published data for the historic time series for gross fixed capital formation (GFCF) for the years preceding 1997. Until these investigations have been completed, the historical data series for GFCF should be treated as being under review. Current data for GFCF from 1997 onwards are not affected. Our investigations are expected to be completed by 1 June 2025. We would welcome any further comments and user views. Because of the compilation approach for historical data, any change to a component will be closely assessed to see if this results in revisions to aggregate estimates, including gross domestic product (GDP).

The pre-1997 GFCF dataset, when initially created in 2014, was created by linking the historic estimates to the 1997-onwards series. This is because the short overlap period produced more plausible results than a linking factor derived from a longer time period. However, this introduced uncertainty, as it is not possible to test how well the linking factor accounts for methods changes. This is because no information on the real impact was available, which means that the historic GFCF estimates are inherently more uncertain than the 1997-onwards estimates.

We have previously advised users that we have been reviewing our published data for the historic time series for gross fixed capital formation in response to specific user feedback and engagement. As a result of this review we have identified areas for further investigation especially with regard to our dwellings data pre-1997 - transport data has also been reviewed but we propose no further investigation into that asset at this stage. We continue to advise therefore that the GFCF pre-1997 data series should be treated as being under review. GFCF data for 1997 onwards are not affected.

We will provide further updates by 30 September 2025 as we consolidate our investigations. In the meantime, we continue to welcome any further comments and views.

View all data related to gross domestic product (gdp)

Contact details for this dataset

Gross Capital Formation team
gcf@ons.gov.uk
+44 1633 455250

Publications that use this data

Business investment in the UK

Notes

1. The inclusion of R&D and other changes was required under the European national accounting standard known by its acronym ESA10, which replaced another acronym ESA95.

2. Martin (2024), chpt. 3. pp. 37-50.

3. For the latest version of his paper, see Jump (2 September 2024).

4. Ibid., p. 21.

5. Mosquera et al. (2 July 2014). ‘Advisory’ is my term for ONS published content that can range from a one-page note to multi-authored articles, now to be found, sometimes with great effort, on the ONS website.

6. The Quarterly Capital Expenditure Survey which provided about half of the required national accounts investment data complied with the earlier SIC 2003 standard until the first quarter of 2011 and thereafter with SIC 2007 (Nolan and Jones (29 May 2014), p. 3). Hughes et al. (December 2009) and Everett (29 September 2011) provide general background details of the original SIC conversion of national accounts data. The term ‘contemporary period’ to describe the national accounts data beginning in 1997 is mine.

7. Martin (2024), p. 26, ft. nt. 73.

8. Jones et al. (14 March 2014) gives extensive details of the revisions, only some of which have been cited here.

9. Harris (25 June 2014), p. 12.

10. Harris (17 January 2013; 25 June 2014) summarise the capital stocks data suspension and recovery. See also Mosquera et al. (2 July 2014), p. 4. The focus in the narrative here is the current price investment series; the ONS had also to merge the investment price deflators: see Mosquera et al. (2 July 2014), pp 7-8.

11. <https://www.linkedin.com/in/jonathan-f-48532068/> [accessed 31 May 2025].

12. <https://www.linkedin.com/in/andrew-banks-2a7a8180/> [accessed 31 May 2025].

13. Bean (2016), para. 4.106, p. 153. His quoted statements were made in the context of an ONS announcement to ‘increase the number of professional economists’.

14. <(31) Louisa Nolan | LinkedIn> [accessed 7 July 2025].

15. Mosquera et al. (2 July 2014), p. 32.

16. The first capital stocks publication following the 2011 suspension (2 July 2014) reported capital stock data for 37 industries. The calculations are undertaken at a much finer level of detail, however. A recent version of the PIMS input dataset (8 December 2023) involves over 140 entries related to investment expenditure by industries in the standard industrial classification together with investment expenditures by government, the latter classified by function and mapped to the SIC.

17. The earliest vintage published by the ONS that I have so far found is dated 1 December 2020. Various academic studies have also used the ONS PIM or PIM-like data (for example, Oulton and Srinivasan (2003); Oulton and Wallis (2016); Rincon-Aznar, Riley and Young (31 January 2017)).

18. ‘Capital stock estimates user guide’ < GitHub - ONSdigital/Capstocks: R scripts and readme for producing capital stocks estimates> [accessed 1 June 2025]. The admission of possible inconsistency with the published national accounts data may be contrasted with the description of the capital stocks estimates that first appeared in the 1956 Blue Book. Rita Maurice notes, ‘From 1948 the estimates of fixed capital formation used in the perpetual inventory are the latest estimates made for the national accounts’. Maurice (1968), p. 385. It may be inferred that the inconsistencies to which the user guide draws attention (and helpfully confirmed in correspondence with the author (e-mails 21 May to 24 June 2024)) arose from later data management problems.

19. Martin (2024), p. 35-36; ft. nt. 118. The cause of the discontinuity in the dwellings investment data is unknown, but it may be related to the ONS decision in 2014 to divide a previous single-industry entry for dwellings into a four-industry entry that distinguished, amongst other things, central government dwellings from the dwellings investments of other industries (Harris (25 June 2014), pp. 13-14). This re-coding may have collided with an inconsistent categorisation of local government housing to the Housing Revenue Account (HRA). For details of the HRA, see Martin (September 2009), p. 744, ft nt. 16. The same PIM dataset was affected by a drop-out in the record of central government research and development: data cells record zero values before a 1987 figure of £841 million (Martin (2024), p. 62).

20. Baybutt (September 2020). I am grateful to Ryland Thomas, senior economist at the Bank of England, for drawing my attention to this recording of the presentation that I had (remotely) attended.

21. Ibid., mins. 43:19 - 43:30 (approx.).

22. Ibid., mins. 49:25 – 52:15 (approx.).

23. Compared with the 8 December 2023 release, the PIM dataset dated 5 June 2025, the latest as of this writing, contains no revisions to the investment in dwellings data before 1997.

24. Harris (25 June 2014), p. 10.

25. Aigner and Goldfeld (1974).

26. Details in Harrison's e-mail to ONS and a copy list, 14 March 2025.

27. Mosquera (2 July 2014), p. 6.

28. Hughes et al. (December 2009) p. 54 presciently noted that the SIC conversion 'should not result in any overall revisions to the whole economy because it is just a reclassification of components'.

29. Nolan and Jones (29 May 2014), pp. 4-6; Nolan and Jones state (p. 4) 'It should be noted that the pre-1997 data have been converted separately from the 1997 – 2010 data. Work is ongoing to quality assure the pre-1997 converted series, and to ensure that no discontinuities are introduced.' Mosquera et al. (2 July 2014), p. 6 and the longer description provided in Harris (25 June 2014), pp. 11-12 do not refer to corrections to the pre-1997 data after conversion comparable to those applied by the ONS to the national accounts capital asset data from 1997.

30. Reproduction of ONS information is licensed under the 'Open Government Licence v.3.0'.

31. The figure is derived from the price deflator for total capital investment as published in the UK Economic Accounts released 27 June 2014, which was consistent with the European accounting standard, ESA95.

32. To calculate the GDP shares, a case can be made for the use of vintages of the GDP data that applied at the time. The use of the latest vintage GDP data simplifies the exposition by confining differences revealed by these calculations to changes in the numerator of the share expression, and does so without affecting the qualitative conclusions that may be drawn.

33. Mosquera et al. (2 July 2014), p. 9 states 'An examination of this effect [the effect of splicing using a 'single point-in time [linking] factor'] by asset shows that between 1988 and 1991 the difference is primarily from the '"plant"' asset, which moves from subtracting approximately £5-7bn per annum from GFCF to adding between £1-6bn per annum to GFCF [gross fixed capital formation]'.

34. Martin (2024), p. 7 (Chart 2), pp. 43-45.

35. Ibid., pp. 33-34.

36. Bean (2016), p. 28, Box 2.A.

37. For econometric time series approaches, see, for example, Chow and Lin (1971), Reuveny and Kang (1996), Greasley and Oxley (2000), Di Fonzo (2003), Caporin and Sartore (2005); Guerrero and Corona (2018); Contreras-Reyes and Idrovo-Aguirre (2019, 2020); Kapetanios and Papailias (2023). For approaches typically used to convert data from one industrial classification to another, see, for example, Buiten, Kampen and Vergouw (2009); Smith and James (2017).

38. International Monetary Fund (2018), pp. 81- 88; para. 5.25, p. 81 states with regard to backcasting used in the context of national accounting, 'Splicing (or linking) is the simplest and most common backcasting method'. Other examples of the use of the technique include the EU KLEMS datasets (O'Mahoney and Timmer (2009)) and the backcast GDP expenditure data in the Bank of England's data compendium, the 'Millennium of Macroeconomic Data for the UK' (Thomas and Dimsdale (2017)).

39. Tyndall (2014), Section B2.2 (unpaginated); Statistics Canada (2014), slides 12, 22.

40. International Monetary Fund (2018), para. 5.30, p. 84; United Nations (2018), Chapter 4, pp. 38-43; see also Eurostat (2010), p. 15.

41. 'Chapter 4: Analyze and Verify; B. A Priori Expectations', United Nations (2018), p. 38, para. 4.6.

42. Mosquera et al. (2 July 2014), p. 13.

43. There are many references to sense checking in Bean's review, for example: 'A paucity of economic expertise, together with cumbersome systems, have meant inadequate sense checking takes place before data are released.' Bean (2016), para. 5.38, p. 201, See, also, Martin (2024), p. 18.

44. Mosquera et al (2 July 2014), p. 7 states that the choice of a single year, 1997, as the linking point 'produced more plausible results than a linking factor derived from a longer time period'.

45. Ibid., Figures 3 - 8; pp. 10-12.

46. Mosquera et al. (2 July 2014), p. 9.

47. Hill and Fox (1997); De la Fuente (2009) and Pedauga (2009) provide formal analyses of certain aspects of splicing and other simple backcasting techniques. De la Fuente (2009) usefully explores a more flexible splicing methodology which is not within the scope of this note.

48. Suppose the linking factors for industries A and B are 8 and 0.8 respectively. The sum of a misallocated £100 is zero before splicing but minus £720 after splicing (minus £100 times 8 for A plus £100 times 0.8 for B).

49. It may be inferred that the capital stock team chose bottom up rather than top down because the former method 'preserves the detail of the lower level series' (Mosquera et al. (2 July 2014), p. 7).

50. The team's data were developed using the SAS statistical software suite, but could not be found during an ONS initial investigation that began in late-2019 (e-mail correspondence, 3 March 2020).

51. The correlation coefficient between the two series over the 1970 to 1996 interval is 0.96. An ordinary least squares regression with the team's results as the dependent variable returns a coefficient on the 2-industry approximation of 0.91 and a constant term of minus 7.7.

52. See, for example, Dagum and Cholette (2006); Sefton and Weale (1995). Martin (2024), pp. 31-32, Chart 12, provides an example of top-down splicing applied to total investment data.

53. <UK Statistics Assembly 2025 – UK Statistics Authority> [accessed 14 June 2025].

54. Briscoe (January 2025).

55. See, also, Martin (2024), pp. 1-2.

56. E-mail correspondence, 29 October 2018.

57. E-mail correspondence, 30 October 2019.

58. E-mail correspondence, 3 March 2020.

59. Martin (2024), p. 2. The ONS had issued a warning about the inherent uncertainty of the investment data before 1997 as compared with data beginning in 1997 (Mosquera et al. (2 July 2014), p. 7). The warning was at best obscure, and did not impart any sense of what was later discovered; see Martin (2024), p. 2, ft. nt. 7.

60. The Office for Statistics Regulation (OSR) – the regulatory arm of the UK Statistics Authority - prefers the term 'Accredited Official Statistics': <Accredited official statistics – Office for Statistics Regulation (statisticsauthority.gov.uk)> [accessed 24 June 2025]. The OSR's January 2024 'quality focused assessment' did not identify the suspected historic problems with the ONS company profits data. See Office for Statistics Regulation (24 January 2025) for an entry noting its 'casework' enquiry into the historic investment data.

61. Hand (4 March 2025), p. 7.

62. Office for Statistics Regulation (February 2018), p. 11, states 'Producers show they comply with the Code by holding themselves accountable to adherence to clear publication policies – for example, […] by notifying users promptly of errors […]'. This statement remains unchanged in the May 2022 revision of the Code, p. 11.

63. For example, in UK practice, CVM data are additive only from the 'reference' year, the date at which a CVM index is scaled to equate with its current price equivalent.

64. Equation (A8) relies on the fact that $\Phi_i \overline{v}_i = \frac{X_{i,t=\ell}}{Y_{i,t=\ell}} \frac{Y_{i,t=\ell}}{Y_{t=\ell}} = \frac{X_{i,t=\ell}}{Y_{t=\ell}}$ so that

$$\sum_{i=1}^{i=n} \Phi_i \overline{v}_i = \frac{\sum_{i=1}^{i=n} X_{i,t=\ell}}{Y_{t=\ell}} = \frac{X_{t=\ell}}{Y_{t=\ell}}.$$

65. The same property can be derived from equation (A7). Re-written at the linking point $t = \ell$, the equation becomes: $D_{t=\ell} = \sum_{i=1}^{i=n} \frac{\Phi_i \overline{v}_i}{\Phi} - 1$, which, given equation (A8), simplifies to $D_{t=\ell} = \frac{\Phi}{\Phi} - 1 = 0$.

66. As noted in the main text, the ONS reported that it could not find the advisory's dataset (e-mail correspondence, 3 March 2020). As far as can be ascertained, the 2010 equivalent of recent 'PIM inputs' databases has not been published.

67. ONS, Industry by Asset, (9 September 2016).

68. Mosquera et al. (2 July 2014), annex 2, p. 35. The same statement is made in Harris (25 June 2014), p. 2.

69. Omundsen et al. (31 July 2010), Table 4.1.1; the total for 'plant and machinery' has the database code DLXI.

70. Dolling (1998), p. 103.

71. Mosquera et al. (2 July 2014), p. 6.

References

Aigner, D. J. and Goldfeld, S. M. (1974), ‘Estimation and prediction from aggregate data when aggregates are measured more accurately than their components’ *Econometrica*, 42.1, pp. 113–34.

Baybutt, C. (September 2020), ‘Rebuilding Two Centuries of Housing Investment Estimates’, Contributed Session A, Capital Stock Conference on Economic Measurement 2020, Economic Statistics Centre of Excellence, mins. 30:40 – 52:49 (approx.) <EM2020 Contributed Session A - YouTube> [accessed 1 June 2025]

Bean, C. (2016) ‘Independent Review of UK Economic Statistics’, March 2016, <https://www.gov.uk/government/publications/independent-review-of-uk-economic-statistics-final-report)> [accessed 31 May 2025].

Briscoe, S. (January 2025), ‘Simon’s words on user engagement (at the UK Statistics Assembly)’ <https://simonbriscoeblog.wordpress.com/2025/06/24/simons-words-on-user-engagement-at-the-uksa-assembly> [accessed 24 June 2025].

Buiten, G., Kampen, J. and Vergouw, S. (2009), ‘Producing historic time series for STS-statistics in NACE Rev.2’, Discussion paper (09001), Statistics Netherlands.

Caporin, M. and Sartore, D. (2005), ‘Methodological aspects of time series back-calculation’, Working Papers and Studies, European Commission, Luxembourg: Office for Official Publication of the European Communities

Chow, G. and Lin, A. (1971), ‘Best linear unbiased interpolation, distribution, and extrapolation of time series by related series’, *The Review of Economics and Statistics*, 53.4, pp. 372-75.

Contreras-Reyes, J. and Idrovo-Aguirre, B. (2019), ‘Backcasting cement production and characterizing cement’s economic cycles for Chile 1991-2015’, *Empirical Economics*, 57, pp. 1829-52.

—— (2020), 'Backcasting and forecasting time series using detrended cross-correlation analysis', *Physica A: Statistical Mechanics and its Applications*, 560.

Dagum, E. and Cholette, P. (2006) *Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series*, Springer Science+Business Media.

De la Fuente, A. (2009), 'A Mixed Splicing Procedure for Economic Time Series', CESifo Working Paper no. 2876.

Di Fonzo, T. (2003), 'Constrained retropolation of high-frequency data using related series: A simple dynamic model approach', *Statistical Methods & Applications*, 12, pp. 109-119.

Dolling, M. (ed.) (1998), 'Introducing the European System of Accounts 1995 in the UK: The changes in the accounting system and the reasoning behind them', The Stationery Office.

Dye, J. and Sosimi, J. (eds) (30 July 2010), *United Kingdom National Accounts: The Blue Book*, 2010 edn, dataset, ONS.

Eurostat (2010), Backcasting Handbook: Implementation of NACE Rev.1 for EU', Eurostat. <https://circabc.europa.eu/sd/a/be4014a9-0764-422c-a700-0b2a46e5929c/Handbook%20part%204%20-%20Back%20casting%20-%20EN.pdf>, [accessed 7 April 2025].

Everett, G. (29 September 2011), 'Method Changes in the 2011 Blue Book', ONS https://webarchive.nationalarchives.gov.uk/ukgwa/20160105193657mp_/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/articles/methods-changes-in-the-2011-blue-book.pdf > [accessed 21 June 2025].

Greasley, D. and Oxley, L. (2000), 'Measuring New Zealand's GDP 1865-1933: A Cointegration-based approach', *Review of Income and Wealth*, 46.3, pp. 351-368.

Guerrero, V. and Corona, F. (2018), 'Retropolating some relevant series of Mexico's System of National Accounts at constant prices: The case of Mexico City's GDP, *Statistica Neerlandica*, 72, pp. 495-519.

Hand. D. (4 March 2025), 'Independent Report on the 2025 UK Statistics Assembly', updated 11 March 2025 <Independent report on the 2025 UK Statistics Assembly – UK Statistics Authority> [accessed 14 June 2025].

Harris, W. (17 January 2013) ‘Capital Stock, Capital Consumption – Update on progress to reinstate estimates’, ONS. <[ARCHIVED CONTENT] Capital Stock, Capital Consumption - Update on progress to reinstate estimates - ONS> [accessed 31 May 2025].

—— (25 June 2014) ‘Methodological changes to the estimation of capital stocks and consumption of fixed capital, ONS. <https://webarchive.nationalarchives.gov.uk/ukgwa/20160107010200mp_/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/articles/2011-present/methodological-changes-to-the-estimation.pdf> [accessed 31 May 2025].

Hill, R. and Fox, K. (1997), ‘Splicing Index Numbers’, *Journal of Business & Economic Statistics*, 15.3, pp. 387-9.

Hughes, J., James, G., Evans, A. and Prestwood, D. (December 2009), ‘Implementation of Standard Industrial Classification 2007: December 2009 update’, *Economic & Labour Market Review*, 3.12, pp. 51-55.

International Monetary Fund (2018), ‘Backcasting’, in *Quarterly National Accounts Manual, 2017 Edition*, pp. 81-84, Washington DC. <IMF Statistics Quaterly National Accounts Manual and Guide 2017 Edition> [accessed 7 April 2025].

Jones, J., Nolan, L., Banks, A., Wales, P. and Morgan, D. (14 March 2014), ‘Explaining United Kingdom investment estimates: past, present and future’, ONS.

Jump, R. (2 September 2024), ‘Revisiting the Role of Profits in the Collapse of Britain’s post-war consensus’, doi:10.2139/ssrn.4578981 <https://ssrn.com/abstract=4578981> [accessed 15 June 2025].

Kapetanios, G. and Papailias, F. (2023), ‘Nowcasting with Backcalculated Short Time Series & Empirical Evidence’, ESCoE Discussion Paper 2023-08.

Martin, B. (September 2009), Resurrecting the U.K. Historic Sector National Accounts’, *Review of Income and Wealth*, 55.3, pp. 737-51.

—— (2024), ‘You really never had it so good, or why Britain’s post-war national accounts could lead you astray’, Centre for Business Research, University of Cambridge, Working Paper no. 541 <Bill Martin - Cambridge Judge Business School> [accessed 26 June 2025].

Maurice, R. (ed.) (1968), *National Accounts Statistics: Sources and Methods*, Central Statistical Office, Her Majesty's Stationery Office.

Mosquera, J. F., Banks, A., Nolan, L., Harris W. and Carpenter, O. (2 July 2014), 'Impact of the methodological changes to the estimation of capital stocks and consumption of fixed capital', ONS. <https://webarchive.nationalarchives.gov.uk/ukgwa/20160107010200mp_/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/articles/2011-present/methodological-changes-to-the-estimation.pdf> [accessed 31 May 2025].

Nolan, L. and Jones J. (29 May 2014) 'Changes to National Accounts: gross fixed capital formation (investment) – changes for Blue Book 2014 (excluding ESA10), ONS. <https://webarchive.nationalarchives.gov.uk/ukgwa/20160105193657mp_/http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/articles/2011-present/gross-fixed-capital-formation--investment-.pdf> [accessed 31 May 2025].

Office for Statistics Regulation (February 2018), 'Code of Practice for Statistics: Ensuring official statistics serve the public good, edn 2.0 <Code-of-Practice-for-Statistics.pdf> [accessed 15 June 2025].

—— 24 January 2025, 'Historic fixed capital formation data' Casework Issues Log 2024-25 <Casework – Office for Statistics Regulation> [accessed 26 June 2025].

O'Mahony, M. and Timmer, M. (2009), 'Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database', *Economic Journal*, 119, 538, pp. F374-F403, doi:10.1111/j.1468-0297.2009.02280.x.

Omundsen, B., McClaren, C., Saunders, B. and Ware, T. (eds) (31 July 2010) *Capital Stocks, Capital Consumption and Non-Financial Balance Sheets 2010*, dataset, ONS.

ONS, 'Industry by Asset' (9 September 2016): 'Annual gross fixed capital formation by industry and asset, 9 September 2016 < Annual gross fixed capital formation by industry and asset - Office for National Statistics> [accessed 6 June 2025].

—— ‘PIM’ (5 June 2025): ‘Perpetual inventory method inputs, the national balance sheet and capital stocks, preliminary estimates, UK: 2025’, <The national balance sheet and capital stocks, preliminary estimates, UK: 2025 - Office for National Statistics> [accessed 5 June 2025]

—— ‘Business Investment Time Series’ (28 March 2025, 15 May 2025): ‘Important notes and usage information’ <Business investment time series - Office for National Statistics> [accessed 12 June 2025].

Oulton, N., and Srinivasan, S. (2003) ‘Capital stocks, capital services, and depreciation: an integrated framework’, Working Paper no. 192, Bank of England.

Oulton, N., and Wallis, G. (2016) 'Capital stocks and capital services: integrated and consistent estimates for the United Kingdom, 1950–2013', Economic Modelling, 54, pp. 117-125.

Pedauga, Luis, E. (2009) ‘An alternative method for splicing time series: Case Venezuela 1950-2005’, MPRA, paper no. 35409. <https://mpra.ub.uni-muenchen.de/35409/MPRA> [accessed 1 May 2025].

Reuveny, R. and Kang, H. (1996), ‘International Conflict and Cooperation: Splicing COPDAB and WEIS Series’, *International Studies Quarterly*, 40.2, pp. 281–305.

Rincon-Aznar, A; Riley, R. and Young, G. (31 January 2017), ‘Academic Review of Asset Lives in the UK: Report to the Office for National Statistics’, National Institute of Economic and Social Research, Discussion Paper no. 474 <https://www.niesr.ac.uk/wp-content/uploads/2021/10/DP474-4.pdf> [accessed 1 June 2025].

Sefton, J. and Weale, M. (1995), *Reconciliation of National Income and Expenditure*, Cambridge University Press.

Smith P. and James, G. (2017), ‘Changing Industrial Classification to SIC (2007) at the UK Office for National Statistics’, *Journal of Official Statistics*, 33.1, pp. 223-47.

Statistics Canada (2014), ‘Advisory Expert Group Backcasting’, <9th Meeting of the Advisory Expert Group on National Accounts> [accessed 7 April 2025].

Thomas, R. and Dimsdale, N. (2017), ‘A Millennium of UK Data’, Bank of England OBRA dataset, http://www.bankofengland.co.uk/research/Pages/onebank/threecenturies.aspx [accessed 2 April 2025].

Tyndall, J. (2014), ‘Backcasting, Bridging, Parallel Runs, and Seasonal Adjustment, Australian Bureau of Statistics, SNA/M1.14/8.1. <9th Meeting of the Advisory Expert Group on National Accounts> [accessed 7 April 2025].

United Nations (2018), *Handbook on Backcasting (draft),* SNA/M1.18/8.iii.b, New York. <SNA/M1.18/8.iii.b - Handbook on Backcasting> [accessed 7 April 2025].

Abbreviations

ESA10: European system of national and regional accounts, 2010

ESA95: European system of national and regional accounts, 1995

ESCoE: Economic Statistics Centre of Excellence

EU KLEMS: European Union Capital (K) Labour (L) Energy (E) Materials (M) Services (S)

ONS: Office for National Statistics.

OSR: Office for Statistics Regulation

PIM: Perpetual Inventory Method

UK: United Kingdom