Productivity measurement: Reassessing the production function from micro to macro

Authors:

Josh Martin

Bank of England, ESCoE and The Productivity Institute

Rebecca Riley

King's College London and The Productivity Institute

Date:

July 2023

The Productivity Institute

Working Paper No.033

Key words
productivity, economic measurement, production function

Authors' contacts:
rebecca.riley@kcl.ac.uk

Acknowledgements
Financial support from the Economic and Social Research Council (grant: ES/V002740/1) is gratefully acknowledged. Thanks to Diane Coyle, Mary O'Mahony, Bart Van Ark and Tony Venables for helpful comments. Any views expressed are solely those of the authors and so cannot be taken to represent those of the Economic Statistics Centre of Excellence (ESCoE), its partner institutions, the Office for National Statistics (ONS), the Bank of England or to state Bank of England policy. This paper should therefore not be reported as representing the views of the Bank of England or members of the Monetary Policy Committee, Financial Policy Committee or Prudential Regulation Committee.

Suggested citation

J. Martin, R. Riley (2023) *Productivity measurement: Reassessing the production function from micro to macro* Working Paper No. 33, The Productivity Institute.

The Productivity Institute is an organisation that works across academia, business and policy to better understand, measure and enable productivity across the UK. It is funded by the Economic and Social Research Council (grant number ES/V002740/1).

The Productivity Institute is headquartered at Alliance Manchester Business School, The University of Manchester, Booth Street West, Manchester, M15 6PB. More information can be found on The Productivity Institute's website. Contact us at theproductivityinstitute@manchester.ac.uk

Abstract

The productivity growth slowdown in advanced economies during the early decades of the 21st century has led to renewed interest in economic measurement. Measured productivity growth has largely evaporated, yet in many ways the average person is better off than at any time in history and technological advance is ever evident. Are we simply, or at least in part, mismeasuring productivity change? More fundamentally, are we measuring an outdated or otherwise less relevant economic concept? What should and can we measure in the interest of developing evidence-based policy solutions to support productivity growth? This paper reviews some of the recent advances in economic measurement and points to an expanded productivity measurement research agenda arising with these questions.

1. Introduction

At its core, productivity measures the rate at which "inputs", such as our time spent working, are transformed into "outputs", i.e., goods and services. It is no surprise then that productivity underpins living standards. This has been evidenced by significant coincident differences in productivity and living standards across countries and time.

For the last 15 years or so productivity growth in the UK and other advanced economies has slowed significantly. The collapse in productivity growth in the UK has been amongst the most pronounced among advanced economies, but the broad patterns of stagnation are in many respects similar across the developed world (Riley, Rincon-Aznar and Samek, 2018; Fernald and Inklaar, 2022). There has been much research to understand these developments, but little in the way of a comprehensive set of explanations providing an easily identifiable remedy. In part this reflects the myriad of factors that influence productivity, arising with its near interchangeability with the concepts of economic performance and growth.

At the same time, globalisation and digitalisation have changed fundamentally the way in which inputs are transformed into outputs, affecting our ability to measure, analyse and understand productivity growth and therefore our ability to prescribe effective policy for enhancing it. And, despite the centrality of productivity to living standards over centuries, there is increasingly a sense that current discourse around productivity is off topic. This in part reflects the changing nature of production that demands a refresh of analytical approaches, but also the relative absence of critical inputs and outputs such as the environment, health and intangibles from mainstream productivity narratives and statistics.

Against this backdrop, this paper reviews some of the recent and emerging developments in productivity measurement and methods of analysis. The issues considered are mostly universal in nature, but UK evidence and practice is highlighted where relevant. The overarching questions that motivate and frame the scope of this review are: To what extent does the recent slowdown in productivity reflect mismeasurement? How might we best advance productivity statistics to support evidence-based growth policy? And is there an increasing disconnect between measured productivity and welfare? We focus on a few important research topics that emerge in response to these questions, and the conceptual and methodological issues they raise for productivity measurement and analysis. It is not our intention to provide a general and comprehensive review of productivity measurement and methods of analysis, which can be found elsewhere.[1] Nor do we provide an overview of datasets and statistics that might inform productivity analysis.[2] Further, and to be clear, when we use the term "mismeasurement" we are not referring to any measurement "errors" by statistical agencies, researchers or other analysts. Rather, we are referring to the limitations of

[1] For an extensive and detailed overview of the many detailed issues involved in productivity measurement and analysis methods more generally see the volume edited by Grifell-Tatjé, Knox Lovell and Sickles (2018). Sickles and Zelenyuk (2019) provide a comprehensive volume on productivity and efficiency measurement.
[2] Oulton (2021) provides an overview of the theory underlying the measurement of productivity and current practice in the UK, linking to relevant datasets.

current practice and the extent to which we are measuring the concepts that might help inform policy to promote growth and prosperity.

The productivity measures considered in this paper are principally labour productivity and total factor productivity (TFP). What exactly these describe depends on what is captured in the numerator (outputs) and the denominator (inputs) of these metrics. This can vary quite substantially across studies and sometimes with significant implications for the conclusions that can be drawn.[3] Labour productivity measured as gross value added (GVA) or GDP, per head or per hour worked, is commonly used in policy analysis and maps relatively easily to living standards, although as we will see this comparison has its limitations. It is also attractive in its relative simplicity compared to TFP and in its straightforward interpretation in both levels and growth rate terms. Closer to a concept of change in efficiency, TFP growth is commonly measured as growth in gross output or GVA less growth in an aggregate of inputs to production.[4] Changes in inputs are usually aggregated using output elasticities inherent to a production function, equivalent to the ratio of input (factor) costs to sales under perfect competition.

As a starting point for productivity measurement and analysis, the standard production function framework provides a powerful basis for measurement that sits naturally with the dual entry structure of the National Accounts and the distinction there between stock and flow measures. It can, and often is, deployed at more disaggregate levels, for example at the level of an industry, region or firm. To be informative, the framework requires that: inputs and outputs can be measured; for the same and relevant unit of analysis; and that the production technology is well-behaved. So, what can we do when, potentially, none of these apply? This is the substance of research on and challenge to productivity measurement today. This paper reviews some recent advances and highlights where substantive challenges remain. It also sets these within the longstanding and broader "beyond GDP" discussion, reignited with the work of Stiglitz, Sen, Fitoussi (2009).

Prices versus quantity and quality

First, at the heart of reflecting modern production structures in estimates of growth and productivity lie a number of issues relating to the adjudication of growth in nominal values into changes in the quantity and quality of goods and services on the one hand, and changes in price indices on the other hand. These issues are in many respects not new. Writing in the US context, but applicable more generally, Feldstein (2017) reviews the difficulties in estimating the value created by new products and quality change. He concludes that real output is rising more quickly (and the associated price index and inflation less quickly) than official estimates suggest, and that "substantial errors of unknown size remain in our ability to measure both real output and inflation".

[3] Amongst other factors, differences across studies are often dictated by the unit of analysis, the data to hand and the objectives of the analysis.
[4] Following the seminal work of Solow (1956, 1957), and thus sometimes known as the "Solow residual".

These issues have come to the fore in the context of the digitalisation of the economy and waning productivity growth. The extent to which digitalisation has contributed to the stagnation in measured productivity growth since the start of this century remains to be quantified, as discussed in section 2 of this paper. Recent research (e.g., Abdirahman, Coyle, Heys and Stewart, 2020; Nakamura, 2020) suggests it may have been a significant contributing factor. In the very least, research points to the need to adopt new valuation methods to capture the productivity effects of digitalisation (e.g., Brynjolfsson et al., 2019, 2020), including better measurement of time use. In some respects, the issues here mimic established difficulties in measuring the output and productivity of the non-market sector, the outputs of which are usually available at zero price at the point of consumption. Atkinson (2005) recommended a value-added as opposed to cost-based approach to the measurement of public services output, key to which is an assessment of quality change. In the UK the ONS produces a set of public services productivity statistics, implementing these recommendations for much of public services output. New data sources may offer an opportunity to further improve these estimates.

Unit of analysis

The measurement and interpretation of TFP depends on the unit level for which it is measured. For example, we might consider measuring transport infrastructure at the level of a country or region and measuring explicit contributions from improvements or deterioration of infrastructure to changes in labour productivity (Grice, 2016). Measured TFP could then exclude these contributions. At the level of the firm this makes less sense. Changes in infrastructure will undoubtedly affect firm performance, but these are not internal to the firm and are necessarily captured in a measure of a firm's TFP. In section 3 we discuss measurement issues specific to different units of analysis: national economies, subnational jurisdictions, industries and firms.

At the level of national economies, international comparisons of labour productivity and TFP are plagued by differences in measurement practice and feasibility across countries. Whilst statistical agencies typically adhere to guidelines set out in international accounting systems, such as the System of National Accounts (SNA) and European System of Accounts (ESA), there is sufficient wriggle room in practice to muddy comparisons of productivity levels and growth between countries. There is much to be gained from studies that harmonise measurement for the purposes of international comparison (e.g., Ward, Zinni and Marianna, 2018). Further complicating estimates of national productivity is the fact that a lot of production takes place across national boundaries, within multinational corporations. One implication of this is that the outputs of production may be recorded in one country and the associated inputs of production recorded in another. The consequences for measured productivity trends at the national and industry level can be significant (Guvenen, Mataloni, Rassier and Ruhl, 2022) and have in some cases led to the adoption of alternative productivity metrics. The globalisation of production emphasises the importance of exploring productivity through the lens of international supply chains.

At subnational level, productivity measurement is also affected by firms reporting inputs and outputs across jurisdictions, rather than for a particular location, and is conceptually challenged. Further, the issue of separating growth into real output and inflation is exacerbated at subnational level, where the general absence of information on prices complicates the study of productivity. New data sources and advances in small area estimation and nowcasting may enhance subnational productivity metrics (e.g., Koop, McIntyre, Mitchell and Poon 2020; Koop, McIntyre, Mitchell, Poon and Wu, 2023).

Over the last two decades there has been significant advance in the development and use of large-scale firm-level datasets to study the productivity of firms and the interaction between firms' productivity, market structure and aggregate productivity (e.g., Altomonte and di Mauro, 2022; De Loecker and Syverson, 2023). However, significant barriers to the wider use of these data mean that their full potential is yet to be realised, both in the UK and elsewhere. Data sources that measure firm-level prices and product composition are largely yet to be exploited.

Missing capitals

Measured TFP growth captures output growth that cannot be explained by growth in measured production inputs. The measured TFP growth residual reflects growth arising from unmeasured inputs and spillovers. Capital services that are often accounted for in these calculations include tangible capital services such as machinery and equipment, but there is a slew of other capital assets that might also be included. What is included matters for the interpretation of measured TFP. Capital assets that are often "missing" from TFP calculations include natural, social, intangible and human capital. Measured TFP growth then captures the contributions of all of these "missing" assets to changes in output. We discuss two key missing capitals, intangibles and human capital, in section 4, and natural capital in section 5. We also discuss recent work to adjust TFP measures for changes in capital utilisation.

Firms' investments in intangibles (e.g., R&D, software, data, management systems, brand value and workforce training) are known to be associated with better performance, but are typically not observed or taken into account in estimates of firm-level productivity. When intangibles are measured, they appear to be concentrated amongst relatively few firms (see e.g., Riley and Rosazza-Bondibene, 2019) such that failing to account for these may increase estimated TFP dispersion or mark-ups. Official estimates for the UK suggest that in aggregate investments in intangibles are of a similar magnitude to investments in tangible assets such as buildings, machinery and equipment (ONS, 2021). Incorporating these in productivity estimation involves making adjustments to production outputs as well as inputs, and is complicated by difficulties in evaluating appropriate deflators and depreciation rates (Corrado, Haskel, Jona-Lasinio and Iommi, 2022). Growth accounting analysis often takes account of changes in the human capital embodied in the workforce through a labour composition adjustment. Recent research points to a number of potential extensions that might improve this adjustment. The importance of investment in intangible and human capital for technology adoption means that, when these capitals are unaccounted for or "missing", estimated productivity growth may be biased downwards during the adoption phase for a general-purpose technology (Brynjolfsson, Rock and Syverson, 2021).

The properties of intangible capital, described in Haskel and Westlake (2018), and human capital, are such that they challenge the standard production function approach to evaluating productivity. In this approach, production spillovers and economies of scale, arising with the use of intangible and human capital, manifest themselves in measured TFP. The contributions of these capitals to growth are then still in some sense unaccounted for, even when included within the typical growth accounting framework.

The economy and the environment

While the aforementioned issues arguably relate to accurate measurement of a traditional concept of productivity, another issue is whether the concept is the correct one in the first place, or whether it is still relevant in the current climate. A key issue here is the role of the environment in the economy. As we discuss in section 5 of this paper, there are at least three ways in which "the environment" is relevant for the conceptualisation and measurement of productivity.

First, natural capital can be considered a "missing" capital input in most productivity analysis. While some studies have made efforts to include natural capital (Cárdenas Rodríguez, Haščič and Souchier, 2018; Brandt, Schreyer and Zipperer, 2013 and 2014), these usually limit the scope of "natural capital" to only traded natural commodities (such as fossil fuels and minerals) or land. The inclusion of a very broad measure of natural capital in productivity analysis is hampered by the lack of necessary information on all natural capital assets, including stock, flow and price information.

Second, negative environmental externalities are not captured in standard productivity analysis, on the implicit assumption of "free disposal" of bad outputs (such as pollution). Some studies have accounted for the role of some negative environmental externalities, such as greenhouse gas emissions and other pollutants, in national (Cárdenas Rodríguez, Haščič and Souchier, 2018; Brandt et al., 2014) or industrial (Agarwala and Martin, 2022) productivity analysis. But these are arguably still quite narrow in capturing such negative externalities, not including for instance biodiversity loss.

Third, expenditure to reduce environmental impact of economic activity, and on environmental protection, is largely not included in GDP since it is treated as a current expenditure rather than as investment. Spending on various activities to protect the environment (such as changing business models or suppliers, using more expensive but more environmentally friendly inputs, and waste remediation) could be considered investments in the environment, which currently go unmeasured in GDP. Although pertinent to managing the net-zero transition, this issue has received relatively little attention thus far.

GDP and welfare

Discourse on the inadequacy of GDP in capturing outputs that matter for our well-being and longer-term sustainability is not new (see Aitken (2019) for an overview). Yet GDP is central to productivity measurement, and measured productivity has been shown to be central to living standards and welfare, although the linkages are complex (see Sharpe, Sichel, Van Ark, 2022). The question is whether measured productivity growth has become less good at capturing advances in living standards of late. This could happen if the factors not captured in GDP have different effects on welfare than they did before or if, as we discuss throughout this paper, structural change in the economy has rendered GDP less representative of activity in the modern economy. In section 6 we consider this contrast and suggest that overcoming the issues described in this paper, and incorporating the associated solutions in national accounts, would overcome many of the limitations of GDP per capita and measured labour productivity in describing trends in welfare and prosperity.

2. Prices versus Quantity and Quality

Distinguishing between inflation and growth in real output is fundamental to productivity measurement. It is also acutely related to the accurate measurement of long-term trends in living standards and welfare. There are many long-standing issues that complicate the distinction between changes in price and changes in quality and quantity, including goods and services that are free at the point of consumption, technical change and product innovation, the heterogeneity of services products, and conceptualisation of value in the non-profit and financial services sectors (see e.g., Bean, 2016). We discuss some of the issues and related advances in measurement research in the context of the digitalisation of the economy and the measurement of non-market services.

Digitalisation and the value of free

Feldstein (2017) points to potentially significant uncertainty in our ability to measure economic growth versus inflation. Nakamura (2020) also notes this uncertainty, and, in large part due to the digitalisation of the economy, suggests these uncertainties have increased in the last couple of decades and may have contributed to the appearance of secular stagnation in the US. Others have challenged this view. Syverson (2017) suggests that mismeasurement of growth is unlikely to account for a substantial part of the productivity slowdown. Amongst the reasons for this conclusion, he notes that the size of the slowdown across countries is uncorrelated with information and communication technology intensity and that estimates of missing surplus associated with online products are likely small compared with the magnitude of the slowdown.

The extent to which mismeasurement of growth may contribute to productivity stagnation in the 21st century remains an open question, although mismeasurement is clearly only one of a number of explanations. At the very least, the implications of digitalisation pose challenges to productivity measurement and analysis related to the pricing structures of digital products and to evaluating improvements in digital products.

One issue in the context of digitalisation is that many digital products are consumed for free. Thus, their welfare contributions are not well-captured in GDP. Brynjolfsson, Collis, Diewert, Eggers and Fox (BCDEF) (2019) propose a new metric, GDP-B, designed to capture the change in welfare associated with the consumption of new and free products and, crucially, provide suggestions for how this might be estimated (see also BCDEF, 2020). Following Brynjolfsson, Collis and Eggers (2019), they implement incentive compatible discrete choice experiments to assess the effect of Facebook on consumer welfare in the US and suggest that the impact of Facebook has been to add 0.05 to 0.11 percentage points to annual growth in GDP-B since 2004. Were labour productivity and TFP calculated using GDP-B rather than GDP, these would increase by a commensurate amount. As the authors point out, these are very significant changes considering that Facebook is but one product, and considering the slowdown in annual labour productivity growth is around 1-1.5 percentage points per year.

Taking a steer from BCDEF, Coyle and Nguyen (2020b) use contingent valuation methods to assess the value of free digital goods, as well as other public goods (public parks) and paid-for substitutes in the UK using a survey tool. The paid-for substitutes can be used for benchmarking unpaid equivalents. In contrast to BCDEF (2019) their methodology is not incentive compatible. Three waves of the survey were implemented. A February 2020 wave and another in May 2020 spanned the start of the COVID-19 epidemic in the UK and the first national lockdown that ensued. Their study provides useful insights into some of the issues that will need to be explored in developing wider applicability and use of these methods for economic measurement. The changes over this period in the use and valuations of different products (increasing for parks and online learning; decreasing for Google maps and cinemas) provide some confidence in the ability of the methodology to capture real change. However, valuations for online goods are higher than for comparable physical substitutes, are not consistently estimated, and there are issues of aggregation. Using an alternative approach, Poquiz (2023) estimates the value from the consumption of free digital goods to UK households using the prices of paid for digital equivalents as a proxy within a hedonic regression. This analysis suggests that the value of free videoconferencing, personal email and online news, derived by UK households in 2020, was of the order of 1.1 to 2 percent of household final consumption expenditures.

Schreyer (2021) embeds the value associated with free digital products, over and above revenue associated with marketing and data sales, within a consistent accounting framework where time spent by households constitute inputs into the production of this value. These inputs could in principle be measured via time-use surveys. In a separate approach, Nakamura and Hulten (2021) propose an expanded GDP framework, EGDP, which captures the welfare effects of the increased availability of free information to consumers associated with digital technologies. Consumers convert GDP at resource cost to utility via a technology for consumer decision making. This is embedded within a standard growth accounting framework. Via case studies, they demonstrate the potentially significant divergence between consumer welfare and GDP associated with improvements in the availability of free information. Nakamura, Samuels and Soloveichik (2017) propose a production-based approach to valuing free digital content, consistent with the framework of the National Accounts and not capturing consumer surplus in full. Their estimates suggest free digital (and newspaper) content increases real US GDP growth by 0.01 percentage points per year 1929-1995, 0.07 percentage points per year 1995-2005 and 0.11 percentage points per year 2005-2015. Their estimates do not suggest that failing to value free digital content contributes to an explanation of the slowdown in US TFP growth after 2005. According to these, free digital content increases private business TFP growth by 0.07 percentage points per year 1995-2005 and 2005-2014.

Digitalisation and innovation

Another source of mismeasurement related to digitalisation concerns quality adjustment associated with rapid innovation in digital goods and services. As alluded to in Feldstein (2017), the difficulty of estimating quality change due to innovation has been with us for decades. Indeed, some studies suggest there is little evidence that the slowdown in US productivity growth after 2004 arises from mismeasurement of the benefits of innovation in digital goods and services because similar mismeasurement existed before the slowdown (Byrne, Fernald and Reinsdorf, 2016). Nonetheless, mismeasurement associated with digital innovation may be growing over time and become more material in future (Ahmad, Ribarsky and Reinsdorf, 2017).

Indeed, Coyle and Nguyen (2018) construct price indices for cloud computing and storage products and estimate that the price of these services has fallen significantly since 2010. The market for these cloud services is large and growing rapidly and is not reflected in official statistics. Abdirahman, Coyle, Heys and Stewart (2020) suggest that quality improvements in telecommunications associated with substantial advances in, and the spread of, mobile technology since 2010 are significantly higher than allowed for in conventional statistics. Assessing developments in the UK they estimate that between 2010 and 2015 telecommunications services prices could have fallen by between 35 per cent and 90 per cent.

What might be the impacts on measured productivity of better adjusting deflators for new and improved digital products? The UK Office for National Statistics (ONS) adopted a new deflator for telecommunications services in the 2021 Blue Book (annual national accounts update), based on a unit-value approach, which falls very rapidly, in line with the findings of Abdirahman et al. (2020) and very much in contrast to the deflator used previously. This led to large upward revisions to labour productivity growth in the telecommunications sector, from 12.6 to 27.4 per cent per year 1997-2007 on average, and from -0.6 to 24.3 per cent on average per year 2010-2018. With the implementation of double deflation in the 2021 Blue Book, upward revisions to growth in the telecommunications sector led to downward revisions to growth in sectors that were heavy users of telecommunications services (Martin, 2021). The overall effect of the methodological changes implemented in the 2021 Blue Book was to reduce measured UK labour productivity growth by an average of 0.2 percentage points per year in the decade leading up to the Great Financial Crisis, and to increase measured UK labour productivity growth by an average of 0.1 percentage points per year 2008-2018, thus reducing the size of the slowdown in UK productivity growth or, rather, lessening the UK productivity puzzle.

The sectors that account for most of the UK productivity puzzle are manufacturing, information and communication services, and financial services (Martin and Mackenzie, 2021). There are known measurement issues in all of these sectors. Moreover, cross-country sector comparisons of productivity growth show, despite some differences, significant similarities between the UK and some of the other major economies in the sectoral patterns of productivity developments (Riley, Rincon-Aznar and Samek, 2018). These similarities point to common explanations for productivity weakness across countries. This is consistent with a review of service sector deflators by O'Mahony and Samek (2021a), who find that an underestimate of services output growth is unlikely to explain why productivity growth slowed more in the UK than in many other countries in the last decade. That said, universal mismeasurement of services output growth could still affect productivity measurement. O'Mahony and Samek (2021a) detect common biases across countries due to inadequate adjustments for quality in services that have led to small but significant upward biases in inflation and downward biases in real output growth.

Non-market services

Productivity in public services is another area of productivity measurement plagued by a lack of observable prices. Common practice to measuring growth in real output in the public sector is to equate growth in output with growth in inputs. Perhaps in recognition of this, much productivity analysis eschews public sector industries such as public administration and defence, education, human health and social care. In an influential review, Atkinson (2005) recommended a departure from this approach, suggesting that the objective should be to measure the value added of public services rather than the cost of production. Key to this endeavour is to quality adjust output quantity measures.

In the UK the ONS produces a set of public service productivity statistics, which incorporate quality adjustment for some public services. For example, measured as a share of outputs, 80 per cent of health care services and 74 per cent of education services are quality adjusted (Harris and Martin, 2019). The output estimates used in the public service productivity statistics are different to those used in the official labour productivity statistics, which are based on the UK National Accounts. In accordance with the European System of Accounts (ESA) 2010, estimates of non-market output in the UK National Accounts do not currently incorporate explicit quality adjustments. In the decade to 2018, both healthcare and education output, as measured in the public service productivity statistics, rose by 4¾ percentage points more than in the National Accounts (Zella, 2021).

Education services in ONS public service productivity statistics are quality adjusted using measures of attainment in exams, post-school destinations, and student well-being (proxied by measures of bullying). But the use of standardised test scores may not capture key aspects of underlying skills that should be incorporated in quality adjustments. Corrado, O'Mahony and Samek (2021) suggest that nominal and real output measures based on increments to lifetime earnings that can be attributed to formal education are a useful alternative method to current practice. They point to the need to consider the education sector as a whole, rather than separating schools from further education. Their measures can be incorporated into a national accounting framework by treating education as investment in intangible social infrastructure, with education of foreign students allocated to exports.

The approach to quality adjusting healthcare output in ONS public service productivity statistics is based on Dawson et al. (2005). A key component of this is to account for the change in health-related quality of life associated with hospital treatment; an approximation to the change in quality adjusted life years (QALYs). Bojke, Castelli, Grasic, Mason and Street (2018) suggest current practice may overlook important characteristics of the quality of healthcare and consider additional indicators from the National Health Service (NHS) Outcomes Framework. Davies (2020) provides a range of indicators that might be used to evaluate the performance of healthcare services and draw international comparisons.

In practice, there are significant questions to be explored around how different aspects of quality, as well as different quality adjusted services, should be weighed against one another in estimating public services output and hence productivity (Foxton, Grice, Heys and Lewis, 2019). The COVID-19 pandemic brought to the fore some of the difficulties in adequately capturing the real value added of public services in GDP and highlights the inevitable question of coherence (or not) between production and welfare.

In the UK National Accounts, the contributions of education and healthcare services to growth in real GDP are calculated using cost weighted activity indexes (CWAI), where growth in the counts of different publicly delivered activities are grossed by their estimated unit costs. This methodology resulted in an estimated drop in public services volumes in 2020 that was much larger in magnitude in the UK than elsewhere, where volume estimates are essentially calculated as deflated expenditures, which either increased or did not change (Dey-Chowdhury, McAuley and Walton, 2021). This illustrates that despite common international guidelines for collating key National Accounts statistics there is significant scope for differences in measurement in practice. This makes international comparisons of productivity fraught with difficulty; in health and education these are already hampered by the extent to which these services are delivered by the public or market sector. We return to the issue of international comparability of productivity measures in a later section.

International comparability aside, additional adjustments were made to the statistics to better reflect significant changes in some public services during the pandemic. Schools were shut on several occasions and the amount of education provided by teachers in a remote setting was found to be less than provided in the usual classroom setting. ONS adjusted output estimates of the education sector to reflect this. The estimated fall in education services echoes widespread concern about reductions in attainment and lost education years, particularly amongst groups where home production of education was unable to compensate for the loss of public provision.

In healthcare, the fall in output in 2020 chimes less well with intuition (see Coyle, 2022). The reduction in measured output reflects a shift from more costly non-urgent treatments to less costly treatments and highlights potential weaknesses in the weighting methodology for public services activities. Whether quality adjusted estimates provide a different picture depends on the change in QALYs associated with those treatments that did not take place and those new treatments that did and the extent to which these differences are accounted for. Estimating the value added of test and trace and the vaccination programme is associated with some known measurement issues and has similarities to the discussion around the introduction of new and free digital goods. An additional complication is that the value of vaccinations is state dependent: the value of vaccinations during a pandemic is likely much higher than when the virus is under control. These issues all stem from the lack of market prices for non-market output. Where activity is valued based on cost rather than value (particularly social value), activities which are cheap to deliver but socially valuable will be underweighted.

3. Unit of Analysis

Methods for measuring and analysing productivity depend on the unit of analysis considered, e.g., whether we consider the economy as a whole, particular geographical areas, industries or businesses. There is lot to learn from analysing productivity at all of these levels, and of course they are interdependent. Here we discuss some of the issues arising for productivity analysis related specifically to the unit of measurement.

International comparisons of productivity and globalisation

Much productivity analysis relies on cross-country comparison of macroeconomic aggregates and one of the issues that arises in this context is simply that of international comparability. International guidelines for measurement are extensive. They set out guidelines for best practice and the direction of travel. But, in compiling productivity statistics, and the many statistics that underlie these, there is much room for interpretation. This makes sense, because what is best practice for any one country will depend on feasibility, available data sources, resources for development and the structure of the economy. Nevertheless, differences in practice can have significant consequences for economic interpretation. We have already noted this above in the context of public sector output measurement in 2020 during the COVID-19 pandemic. These issues are also evident in the measurement of labour inputs from the analysis of Ward, Zinni and Marianna (2018). They examine in detail the way that hours worked are estimated across countries. They find scope for harmonisation across countries. A harmonised approach to measurement reduces the significant gap between the level of UK labour productivity and that of its peers by around a third. That said, their harmonised measure for the UK is a simplified implementation of the method, and a more thorough analysis might alter the result somewhat. As a result, the ONS have recently presented a range of comparisons (Mackenzie and Baybutt, 2022), supported by research on presentation of such results (Galvao, 2022). There is scope for further exploration of the importance of differences in measurement approaches in influencing perceived wisdom about relative labour productivities across countries.

Issues of international comparability are also important in the context of incorporating new products into price indices and adjusting for quality improvements associated with technological advance. As noted above, current UK practice in measuring service producer price indices does not appear to generate significant differences in productivity trends in the UK compared to elsewhere (O'Mahony and Samek, 2021a). As statistical agencies increasingly move towards incorporating digital technologies and associated quality improvements into national statistics, questions of international comparability will continue to arise. For example, examining trends in mobile phone inflation as recorded in official consumer price indices, Byrne (2019) finds significant unexplained variation across countries. In the decade after the financial crisis, average annual rates of mobile phone inflation vary between no change and declines of 25 per cent. Given that the types of mobile phones consumed in developed countries are largely similar, he suggests mismeasurement may have exacerbated the apparent variation in price trends.

Another key challenge to consider in productivity measurement is the interpretation of national statistics and national economies in an increasingly globalised economy. The classic example is that of Ireland, where GDP increased by 26 per cent in 2015 in the main due to transfers of intellectual property assets within multi-national companies. This led to the increased use of Gross National Income (GNI) rather than GDP in Irish productivity analysis. The Irish Statistical Agency also introduced a supplementary indicator of the size of the Irish economy, GNI*, intended to exclude the effects of globalisation that render GDP a misleading indicator of the size of the economy. Specifically, GNI* removes from GNI the effects of globalisation associated with highly mobile economic activities. In the wake of "the Irish case", statistical agencies have established "large cases units" like the ONS International Business Unit to monitor and understand the implications of multinational enterprise activities for GDP and other national statistics (see e.g., Eaton, 2019).

Distortions to economic measurement associated with movements in intangible assets across national borders within multinational firms are often related to profit shifting and tax planning (Neubig and Wunsch-Vincent, 2017). Profit shifting breaks the relationship between the location where profits are recorded and the location where economic activity actually takes place, with clear implications for measured productivity. Guvenen et al. (2022) illustrate how offshore profit shifting by US multinational enterprises affects US GDP and productivity measures. To do this they reconstruct US GVA by reapportioning the worldwide profits of US multinational enterprises to the US parent and non-US affiliates, based on the location of labour compensation and sales. Their revised GVA estimates suggest that US productivity growth was higher than recorded in official statistics on average between 1994 and 2004, and little different to that recorded in official statistics on average between 2004 and 2016, making the measured productivity slowdown in the US more significant than official estimates indicate. Discrepancies were particularly large in R&D intensive industries.

The study by Guvenen et al. (2022) relied on a confidential survey of US multinational enterprises and their worldwide operations conducted by the US Bureau of Economic Analysis. The study is therefore not easily replicable for other countries. However, global business accounting data available in ORBIS can be used to illustrate the issues for other countries. Mion and Tong (2021) carry out a reapportioning exercise for UK and non-UK multinational enterprises operating in the UK. This exercise suggests that UK activity adjusted for profit shifting might be lower (higher) than recorded in official statistics in 2017 (2007), directionally consistent with the estimates of Guvenen et al. (2022) for the US. Discrepancies in 2017 were particularly high in the petroleum and gas, mining, telecommunications and pharmaceuticals industries, where adjusted activity was less than recorded. There are many assumptions involved in constructing these estimates, and further exploration is necessary if we are to understand the implications of profit shifting for measured productivity growth in the UK.

In a global economy, the production function for final products extends across borders and highlights the need to consider global value chains as a unit of analysis to understand productivity. This has been facilitated to some extent through measurement initiatives such as WIOD (World Input Output Database) and the OECD-WTO TiVA (Trade in Value Added) statistics. Timmer, Miroudot and de Vries (2019) suggest supply chains may be better described in terms of trade in tasks or functions (fabrication, R&D, management and marketing) rather than products, with advanced economies specialising in R&D and management tasks, while the physical manufacturing process is carried out elsewhere.

Subnational productivity

The issues surrounding globalisation and the measurement of national economic activity and productivity are not unrelated to the issues of measuring these at subnational levels within countries. In particular, the potential for measures of production and income (and hence welfare) to diverge because of financial flows across borders is familiar. Subnational measurement is further complicated because flows across borders within countries are usually not counted, unlike many flows across international borders. There is significant scope to improve upon subnational measurement and to understand the position of regional economies in supply chains through the use of detailed consumer and business transaction data and through changes in accounting and survey practices (Davidson, Black, Connolly and Spowage, 2022). The increased trend towards homeworking since the pandemic brings new considerations into play around the concept of the location of economic activity.

Recent policy focus on addressing geographical disparities across the UK has led to a series of innovations in subnational statistics collated by the ONS and released on an experimental basis. These include estimates of GVA for very small areas like the Census Middle layer Super Output Areas.[5] To support subnational productivity analysis the ONS has also released estimates of gross fixed capital formation by asset type and industry for the 179

[5] Middle layer Super Output Areas (MSOAs) comprise between 2,000 and 6,000 households. There are 6,856 MSOAs in England and 408 in Wales.

administrative ITL3 areas that constitute the UK. In developing these estimates ONS apportion company turnover to local areas in line with the distribution of company employment, implicitly assuming labour productivity is equal across business sites. Investment is partially apportioned in line with employment. It seems likely that estimated productivity and incomes may differ substantially for some smaller areas, not only because of necessary assumptions in compiling these statistics, but also because locations of residence may differ from locations of work.

One of the missing building blocks in developing an accurate picture of spatial differences in economic fortunes is an accurate measure of subnational inflation and differences in price levels across the UK. Hearne (2021) suggests spatial disparities in incomes across the UK are reduced when measured on a real rather than nominal basis. Connolly and Spowage (2021) suggest that small sample sizes lead to instability in consumer expenditure weights and are problematic for the calculation of regional price statistics. ONS chain volume measures of regional productivity take account of the industrial structure of the region but use UK (national) industry deflators.

Industry sectors and the activities of firms

Industry sector is often used as a taxonomy for analysing productivity and is the focus of many economic policies. The idea is that an industry defines a set of relatively homogeneous production activities, distinct from technology outside the industry. Industries are described using Standard Industrial Classifications (SIC). Continuity of these classifications is often regarded as helpful to those analysing longer term productivity trends, facilitating efforts like the EU KLEMS and World KLEMS databases (O'Mahony and Timmer, 2009; Jorgenson, 2012). But continuity also comes at a price and misses important economic change. For example, Coyle and Nguyen (2020a) describe the phenomenon of factoryless manufacturing in the UK and the US and suggest this may lead to misattribution of activity across sectors and countries. They suggest that the decline in manufacturing may be overstated due to this phenomenon and recommend consistent measurement through official surveys.

Others have used information scraped from company websites to improve understanding of industry classifications and networks. These methodologies are in their infancy. Bishop, Mateos-Garcia and Richardson (2022) analyse whether SIC codes can be predicted from text and suggest that their methodologies can be used to measure activities related to sustainability and the green economy, as well as the development and adoption of new technologies. Nathan and Rosso (2015) use text mining to identify digital businesses and find that information and communication services activities are 42 per cent larger than SIC-based estimates suggest. Bernini et al. (2017) use text mining of company websites to identify industrial clusters and networks in the UK. They highlight the importance of complementing data analytic methodologies with qualitative case study research.

Firm-level and micro to macro studies of productivity

The growth accounting framework is widely used for measuring the contributions of different inputs to labour productivity growth and to measuring TFP at the national and industry levels. Firm level studies more typically use econometric methods to recover output elasticities and TFP. Econometric analysis of the production function often considers within the same set up the effects of production function inputs and external factors that influence TFP (e.g., the competition environment, supply chain arrangements, etc). A key concern in these studies is the endogeneity of input choices to unobservable factors that influence TFP. A number of methods have been devised to handle this (see e.g., Levinsohn and Petrin, 2003; Wooldridge, 2009; Ackerberg, Caves and Frazer, 2015).

Businesses or producer level data, collected through surveys or via administrative processes, have been used extensively in productivity analysis and to great effect, facilitated by enhanced data access for research purposes and increased computing ability. Syverson (2011) notes the significant and unexplained variation in TFP across businesses within narrowly defined industries that appears in these data sources. Surveys may consider specific aspects of firms' behaviour and determinants of firms' productivity. Large scale firm surveys and administrative data facilitate analysis of linkages between firm behaviour, interactions between firms, and macroeconomic developments (see e.g., Altomonte and di Mauro, 2022, for an overview of this literature). Some recent applications include: the documentation of productivity dispersion, characterisation of frontier firms and catch up (Andrews, Criscuolo and Gal, 2016); shift-share decompositions of aggregate productivity growth (see Riley, Rosazza-Bondibene and Young, 2015, for an example of implementation in the UK data); determinants of input misallocation across businesses (e.g. Anderson, Riley and Young, 2019; Besley, Roland and Van Reenen, 2020); and documentation of patterns of business dynamism and their implications for aggregate productivity (Decker, Haltiwanger, Jarmin and Miranda, 2020; see Lui et al., 2020, for patterns of business dynamism in the UK).

A major issue for these studies is that the data used is often not designed for the purpose to which it is being applied, with important implications for the use of and conclusions that can be drawn from analysis of these data. This is increasingly recognised in the literature, prompting advances in estimation and identification procedures in recovering firms' productivity and calling for greater attention to underlying imputation and weighting structures (see e.g., De Loecker and Syverson, 2023; Kim, Petrin and Song, 2016; White, Reiter and Petrin, 2018; Bartelsman and Wolf, 2018; Kalemli-Ozcan, Sorensen, Villegas-Sanchez, Volosovych, and Yesiltas, 2015).

One issue that arises is the typical absence of firm specific prices in the analysis of productivity at the level of the firm. Nominal values are usually deflated with industry level indices, ignoring heterogeneity in firms' prices within industries. If heterogeneity in firms' output prices reflects differences in product quality, we may be content for these differences to be reflected in measured TFP, as per the discussion in the previous section. The contention is to what extent measured TFP captures technical efficiency versus other factors that influence prices, such as shifts in demand and imperfect competition. There have been several approaches to disentangling these influences. Where it is possible to observe firms' output prices, it is possible to distinguish between a measure of quantity TFP (TFPQ) and revenue TFP (TFPR), as in Foster, Haltiwanger and Syverson (2008). Others have used observed shifts in demand to distinguish between physical and revenue TFP when firms' output prices are unobserved, combining the production framework with a demand system as in De Loecker (2011). Forlani, Martin, Mion and Muûls (2022) decompose firm heterogeneity in revenue TFP into measures of firm heterogeneity in demand, quantity TFP and markups. These studies point to potentially significant differences between measures of technical efficiency and revenue TFP, the shape and importance of which depend on identification and functional form assumptions and the particular research questions being addressed.

Advances, such as these, in methods for distinguishing between productivity and other dimensions of firm heterogeneity have led to new insights into macroeconomic trends. For example, in a study of UK firms around the Great Financial Crisis, Jacob and Mion (2022) illustrate that the sharp slowdown in measured TFP growth after the crisis was more related to demand weakness pushing down markups rather than productivity. Exploiting harmonised cross-country financial accounts data, Andrews, Criscuolo and Gal (2016) suggest that the labour productivity slowdown since the early 2000s has been characterised by increasing labour productivity divergence between firms at the global frontier, who continue to grow, and other firms, for whom labour productivity growth has stagnated. These patterns remain in place once they adjust for differences in capital deepening and markups across firms, leading them to conclude that documented labour productivity divergence reflects divergence in technological efficiency. De Loecker, Eeckhout and Unger (2020) document the emergence in the US since the 1980s of an upper tail of firms with high markups and profitability, increasing aggregate markups and suggesting a rise in market power.[6] Autor, Dorn, Katz, Patterson and Van Reenen (2020) suggest that globalisation and technical change have led to the rise of superstar firms with very high markups, contributing to a decline in workers' share in growth.[7] ONS analysis of the Annual Business Survey and Annual Business Inquiry suggests patterns of increasing dispersion in markups (Black, 2022) and labour productivity (ONS, 2020) are also evident amongst firms in the UK since the early 2000s.

Another issue that needs to be addressed in measuring productivity at the level of the firm is the absence of information on firms' use of capital services. Company accounts data often includes book values for fixed assets that can be used as proxies for tangible capital stocks. Firm surveys like the UK Annual Business Survey include information on investment in tangible and some intangible assets. These can be used to construct estimates of firms' capital

[6] Rising market power signals imperfect competition. Comin, Gonzalez, Schmitz and Trigari (2022) explore the implications of the zero profit or perfect competition assumption for TFP measurement at the aggregate level. Noting that imperfect competition drives a wedge between output elasticities and factor shares, they develop an estimation method that accounts for this wedge. Their results suggest that standard growth accounting methods underestimate recent TFP growth in the US, Italy and the UK, and overestimate recent TFP growth in Germany, Spain and France.

[7] The extent to which the labour share has fallen outside the US is the subject of some discussion and depends on details of how this is measured (Gutiérrez and Piton, 2020; Teichgräber and Van Reenen, 2021).

stocks (and corresponding capital services) using the perpetual inventory method (PIM). Measurement error is exacerbated by missing longitudinal information on firms' investments and missing information on the appropriate starting stock to use in the PIM, leading to biases in estimated output elasticities and TFP. In the UK there is scope to reduce this source of measurement error through linking company accounts and ONS survey records.

There is clearly much to be learned about productivity from the analysis of business microdata, and there has been an upsurge in use and methods of analysis over the last twenty years. Nonetheless, as a resource, these datasets remain relatively untapped. In the UK this reflects lack of consistent investment in the maintenance and development of these data sources and their documentation, significant non-pecuniary costs associated with data access for research, and difficulties in sharing and linking data between the different institutions that collect these. Together this can present significant disincentives to the exploration of business data for UK productivity research. The ONS has recently made improvements to the usability of some key large scale firm datasets for productivity analysis, including development of the Longitudinal Business Database and preparing contributions to international projects such as CompNet at Halle Institute for Economic Research, and DynEmp and MultiProd at OECD, based on analysis of the Annual Respondents Database. Business data collected by HM Revenue and Customs through administrative processes (e.g., VAT and international trade), and linked with employee data through PAYE records, provide another promising resource for UK productivity research.

4. Missing Capitals

We focus in this section on some of the capitals - intangibles and human - that are often described as missing from productivity calculations. We note recent methods for adjusting for capital utilisation in deriving measures of TFP.

Intangible capital

Intangibles have come to denote a group of purchases that contribute to production over the longer term, and that should therefore be classified as investment, but do not manifest themselves in a tangible physical form. These include spending on R&D, software and databases, as well as management practices, branding and other organisational capital. Intangible investments are increasingly acknowledged as crucial to firms' performance (see e.g., Haskel and Westlake, 2018; Bloom, Sadun and Van Reenen, 2017). And although methods for measuring these investments is still evolving (Martin, 2019), official estimates suggest firms' investments in intangibles are of a similar magnitude to firms' investments in tangible assets such as buildings, machinery and equipment (ONS, 2021). Less than half of these are capitalised in the National Accounts (ONS, 2019).

Much has been learnt about the patterns of intangible investment following the seminal work of Corrado, Hulten and Sichel (2005), which proposed a framework for measuring intangible investment and capital consistent with the principles of national accounting frameworks. On the output side, accounting for intangible investment within this framework increases gross value added and labour productivity levels (not necessarily growth), either because intermediate expenditures are reclassified as investment or because own-account investment is recorded. TFP estimates are further affected insofar as intangible capital services are included in growth accounting frameworks on the input side. In practice, measurement is complicated by a lack of appropriate deflators and depreciation rates (Corrado et al., 2022, provide an overview of the issues), as well as difficulty in collating information on these investments within standard company accounts reporting practices and business surveys (Martin and Baybutt, 2022). Investments in organisational capital, which are not capitalised within the National Accounts, are often proxied by expenditures on employees in management and advertising related occupations. Other intangible investments are usually collated from nationally representative business surveys, although investment in own-account production is usually gleaned from the occupational distribution of labour costs.

Measurement of intangibles has benefitted from several rounds of collaborative international work, largely based in Europe. Previous projects include INNODRIVE[8] (2008-2011), INTAN-Invest[9] (2013-2017), SPINTAN[10] (2013-2017), and GLOBALINTO[11] (2019-2022). These projects have often led to the creation of a cross-country dataset of investment in the full set of intangible assets proposed in Corrado, Hulten and Sichel (2005), estimated in a consistent manner as far as possible. The latest development is the combination of the INTAN-Invest dataset with EU KLEMS (EUKLEMS & INTANProd). This enables cross-country growth accounting with a wide set of intangibles estimated in a consistent manner.

Measures of intangibles at the level of the firm are generally less well-developed compared to country and broad industry aggregates. The ONS and IPO conducted an intangible assets survey in 2010 and 2011. These revealed significant discrepancies between macroeconomic estimates of intangible investment and that which could be derived from firm responses to direct survey questions. The ONS and ESCoE have developed two waves (conducted in 2017 and 2020) of a UK Management and Expectations Survey (MES). These surveys are similar to management surveys developed in other countries, building on the US MOPS and the World Management Survey, and yield longitudinal data on UK businesses' management practices and other production information. The methodology for measuring management practices in these surveys is different to that used for aggregate estimates of management capital, but captures salient features of organisational capital at the level of the firm. The Annual Business Survey, Annual Purchases Survey and Business Enterprise R&D (BERD) survey have additional information on intermediate expenditures and investments that can be employed to estimate firms' use of intangibles (see e.g., Riley and Rosazza-Bondibene,

[8] Project website now closed. A write-up is available: https://cordis.europa.eu/project/id/214576/reporting
[9] http://www.intaninvest.net/
[10] http://www.spintan.net/
[11] https://globalinto.eu/

2019). The development of accounting practices is another avenue to explore in the pursuit of better information on firms' investments in intangible assets.

Survey data on intangibles offer opportunities but are not without their problems. The ONS for many years ran the BERD survey by sampling from a list of known or suspected R&D performers, updated annually from a range of data sources. The resultant estimates of business R&D expenditure were lower than claims for business R&D tax credits, and low internationally, prompting ONS to review the methods. They identified that the survey approach was not accounting for the potentially large number of small businesses carrying out R&D that had not been identified and added to the sampling frame and suggested that this led the previous estimates to understate R&D investment substantially. Changes to survey design are now underway. Another key business survey, the Annual Purchases Survey, was suspended after 2004 to reduce cost and minimise respondent burden. It was reinstated in 2015, but data on the intervening years was never collected. This hampers understanding of changing business spending patterns over a period of significant economic change. More consistency and forward-thinking in business surveys would support analysis of intangibles and productivity.

The issues associated with measuring intangible capital investment and services are plenty. Perhaps even more important, are issues associated with measuring the value contribution of intangible investments. Intangible capital is likely to behave differently in the production process to tangible capital. For example, Haskel and Westlake (2018) characterize intangible capital by its scalability, sunkenness, spillovers, and synergy in production. Furthermore, the returns to intangibles such as R&D may exhibit fat tails, with a small proportion of firms having relatively good returns and most having relatively poor returns. For these reasons the contributions of intangibles to productivity and growth may be difficult to capture in standard production functions or using growth accounting techniques. Indeed, standard methodologies may underestimate the contributions to growth of intangible investments irrespective of whether these inputs are correctly measured. Given these complexities, case studies and data driven approaches may yield new insights about the value contributions of intangible investments.

Goodridge, Haskel and Wallis (2017) explore spillovers from R&D and other intangible assets using industry data. They find evidence consistent with spillovers from both R&D and non-R&D intangibles. Analysing the MES 2017, Bloom, Kawakubo, Meng, Mizen, Riley, Senga and Van Reenen (2021) find that firms with more structured management practices are better at predicting future own firm outcomes and face less subjective uncertainty, yielding new insights into how organisational capital may influence productivity. Li, Mizen, Riley and Schneebacher (2023), find that during the COVID-19 pandemic, better managed firms were more likely to switch to homeworking and online sales, and suffered a smaller decline in turnover, suggesting organisational capital may support resilience and firms' ability to adjust to adversity.

It will likely be increasingly important to capture the contributions of data assets and artificial intelligence to productivity growth. Recent papers consider these questions within an intangible asset framework (Goodridge, Haskel and Edquist, 2022; Corrado, Haskel, Iommi, Jona-Lasinio and Bontadini, 2023; Corrado, Haskel and Jona-Lasinio, 2021). Measurement and valuation issues are amplified by the opensource nature of these assets.

Human capital

Most growth accounting analysis, such as in EU KLEMS, takes account of changes in the human capital embodied in the workforce through a labour composition adjustment. In this case the only missing capital is that being generated by the education system for students enrolled (this can be taken into account using the methodology of Corrado, O'Mahony and Samek (2021), mentioned above). The measures of labour composition can, however, be refined in many directions. For example, they do not typically include an adjustment for health status (O'Mahony and Samek, 2021b). Labour composition measures are generally based on formal qualifications but preferably would be based on the notion of skill. Current estimates do not explicitly account for on-the-job training, although usually do stratify by age as a proxy for general work experience. Firm-specific training is included in the Corrado, Hulten and Sichel (2005) classification of intangibles as part of a broader set of economic competencies.

Surveys such as the Skills and Employment Surveys are useful in generating information on the skills that people use at work. Online job vacancy data might complement surveys. The text contained in job adverts can be used to explore the links between formal qualifications, occupations and skills. For example, Djumalieva and Sleeman (2018) develop a skills taxonomy from Burning Glass job adverts using machine learning methods. Djumalieva, Lima and Sleeman (2018) create a detailed mapping between these detailed skills groups and standard occupations. Posted salaries contained in some job adverts provide an opportunity to attribute value to particular skills; these may find further use in developing labour composition metrics.

Online job advert data also benefit from their granularity and can be used to create a picture of skills demand for sub-national geographies; see Brüning and Mangeol (2020) for US states and Garasto, Djumalieva, Kanders, Wilcock and Sleeman (2021) for UK Travel-to-Work areas. Combined with survey data these might be used to develop local area estimates of skills mismatch. In addition to complementing indicator approaches to measuring human capital (see e.g., the approaches discussed in Vassilev, Sidhu, Payne and King, 2020), mismatch indicators might be used in standard frameworks to adjust labour composition or human capital services utilisation.

Human capital is typically associated with externalities or spillovers in production (see e.g., Kirby and Riley, 2008), which in standard growth accounting exercises would manifest themselves in TFP. But human capital may have broader social impacts, such as reducing crime rates or increasing trust and social participation, affecting welfare.

Capital utilisation

Because of adjustment costs, measured TFP is pro-cyclical. Firms hoard labour and do not dismount their capital stocks when demand is temporarily weak; neither do they immediately adjust factor inputs in upturns. While the issue has long been known, no consensus on the best method to adjust for variations in capacity utilisation has emerged. The most popular approach is that proposed by Basu, Fernald and Kimball (2006), which assumes that firms operate along both observable and unobservable margins simultaneously, such that changes in (observed) hours worked can proxy for changes in (unobserved) capacity utilisation, capturing both labour effort and capital utilisation. Huo, Levchenko and Pandalai-Nayar (2020) suggest that changes in TFP adjusted for factor utilisation in this way are uncorrelated across countries; any co-movement across countries in unadjusted TFP is associated with shocks to factor utilisation. Martin and Jones (2022) explore an extension of these methods to account for differential utilisation rates across different types of capital inputs in estimating UK TFP. Their analysis is based on an assessment of the complementarities between different occupations and capital assets. Amongst others, Comin, Gonzalez, Schmitz and Trigari (2022) adjust TFP for factor utilisation using proxies from business surveys. In a different vein, Diewert and Fox (2018) posit that in so far as TFP measures technical progress it can never fall, because technologies cannot be un-invented. They devise a method for splitting standard TFP indices into an index of technical progress and a measure of cost efficiency. During recessions, cost efficiency decreases and TFP remains flat. More recently, the trend towards increased working from home, brought about by the pandemic, points to the need to adjust productivity measurement for increased utilisation of capital in the home (Eberly, Haskel and Mizen, 2021).

5. The Economy and the Environment

There are at least three ways in which "the environment" is relevant for the conceptualisation and measurement of productivity: 1) natural capital as an input, 2) environmental damage as a "bad" output, 3) investment in the environment as an unmeasured (good) output. We take these in turn.

Natural capital

Amongst the "missing capitals", perhaps the most egregious omission is natural capital. Natural capital covers all manner of natural resources, such as land, soil and subsoil assets, forests, woods and grasslands, water resources, naturally occurring animal and plant life, and more besides. Some of these are treated as "non-produced assets" in national income accounting, which means that they are recognised as capital, but no investment goes into their creation. In practice, many NSIs do not make estimates of the value of these capital stocks, or the capital services which they provide to production. As a result, natural capital is rarely included as a capital input in productivity analysis.

The UK ONS, and some other NSIs, make estimates of the value of natural capital separately to their national income accounting. The measurement of natural capital follows the UN's System of Environmental-Economic Accounting (SEEA), a complement to the System of National Accounts (SNA) used for national income accounting.

Following SEEA, stocks of natural capital are measured either through non-monetary measures, or monetary measures. Non-monetary measures vary according to the natural capital asset in question, and include such measures as their area, quality and condition. Given the range of measurement approaches and units, these measurements are not well suited for inclusion in productivity analysis. Monetary values of the stock of natural capital are based on the net present value of future benefit flows expected to be delivered by the natural capital over the course of its life. This relies on detailed modelling, and a range of uncertain data inputs. These approaches contrast with the measurement of stocks of produced assets, which usually use the Perpetual Inventory Method (PIM), summing long series of past investment data and applying depreciation rates. Since natural capital is not produced, and so there is no recorded investment, a PIM is not feasible.

As a result of these contrasting measurement approaches, there has been limited progress in incorporating natural capital as an input into productivity analysis alongside other capitals in conventional growth accounting analyses. In growth accounting exercises, tangible and intangible (produced) capital are treated symmetrically, with weights determined by the user cost of capital, following Hall and Jorgenson (1967). It is difficult to conceptualise natural capital in the same way, and terms in the user cost equation are less compatible with natural capital. Freeman, Inklaar and Diewert (2021) discuss how natural resources should be treated in international productivity comparisons when not all countries use the same natural resource inputs and propose to treat it as a "missing goods" problem. Diewert and Fox (2016) discuss approaches to the measurement of user costs for non-renewable resources in growth accounting.

Some studies have accounted for natural capital as an input in productivity analysis, either by limiting the scope of natural capital assets considered, or by limiting the scope to particular industries where measurement is more tractable. For instance, Cárdenas Rodríguez, Haščič and Souchier (2018), building on Brandt, Schreyer and Zipperer (2013, 2014), estimate environmentally adjusted multi-factor productivity (EAMFP) including natural capital as an input and accounting for pollution as a 'bad' output across OECD countries. To incorporate natural capital, they estimate an input cost share based on the "market price net of extraction costs", sourced from the World Bank's Wealth Accounting and the Valuation of Ecosystem Services (WAVES) database. In other words, they estimate the cost of a 'raw' unit of natural capital (i.e., net of the cost of its extraction), and use this to estimate its cost share. This relies on having observable market prices, and thus the scope of natural capital is limited to natural commodities, specifically four types of fossil fuels (hard coal, soft coal, gas, oil) and ten minerals (bauxite, copper, gold, iron ore, lead, nickel, phosphate, silver, tin and zinc). This is clearly not exhaustive of all natural capital as broadly conceptualised, omitting at least land, aquatic resources and freshwater. Nonetheless, this is one of few studies to make any adjustment for natural capital inputs, which is commendable. Dang and Mourougane (2014a, b) follow a similar approach for selected Asian countries.

One major drawback of this approach is that these natural capital inputs are valued at market prices, which are volatile. While the authors smooth the values over five years to reduce the effect of volatility, the cost shares are nonetheless highly influenced by the prevailing market price for the commodities. As such, the contribution of natural capital to output growth appears to vary markedly over the economic cycle, in line with the market price of the commodities. Further, the market prices may not reflect the true social or environmental 'value' or cost of the commodities. The valuation, and thus contribution to output growth, of fossil fuels such as oil and gas are very large and dominate the results.

Land, which is arguably one major type of natural capital, is occasionally included in growth accounting analysis of productivity. Like broader measures of natural capital, the omission of land from most research is likely due to the absence of high quality and timely data on land volumes, as well as conceptual issues. Shimizu, Diewert and Nomura (2023) discuss productivity measurement with the inclusion of land types and propose changes to the treatment of land in the System of National Accounts. The OECD Measuring Capital Manual (OECD, 2009) discusses the measurement of land as an input into production but notes numerous conceptual and practice challenges. Studies of agricultural productivity often include some measure of natural capital input, most notably land. For instance, estimates of productivity of the agriculture sector in the US produced by the US Department of Agriculture[12] include land as an input.

The impact of including natural capital in productivity measures depends on the relative growth rates of the included natural capital and the other measured inputs (labour and produced capital). Usually natural capital (especially land) grows relatively slowly, which drags down the aggregate inputs measure once it is included, and thus pushes up estimates of productivity growth.

Accounting for negative environmental externalities

Another way in which environmental factors are considered in productivity analysis is by accounting for negative environmental externalities from production. There are many such negative externalities, but the most commonly measured are greenhouse gas emissions and other pollutants. These are omitted from standard productivity analysis since they are not tracked as output in the National Accounts or other business microdata, and the price indices of outputs do not account for externalities (positive or negative).

To account for these negative externalities, two broad approaches have been taken. First, if a unit price can be ascribed to the negative externalities (e.g., a price per ton of carbon emissions), then they can be accounted for in output measures by combining them with 'traditional' output measures on a monetary basis. This approach can be used to adjust both

[12] https://www.ers.usda.gov/data-products/agricultural-productivity-in-the-u-s/

the level and real growth of output, both of which are relevant for productivity analysis. This is the procedure used by Agarwala and Martin (2022) in exploring environmentally adjusted productivity measures for over 40 industries of the UK, between 1997 and 2019.

This approach relies on the choice of ‘price’ for negative externalities, which are usually not observed in the market. Agarwala and Martin (2022) use prices for carbon emissions and other pollutants assigned by UK government. However, the price of carbon is hotly contested, and there are many estimates based on different methods and assumptions (see Tol, 2005, for a review). For some types of negative externality (for instance, biodiversity loss) prices are less clearly defined. Shadow prices, estimated through stated preference experiments or estimated econometrically (see more below), can enable more negative externalities to be priced and factored in.

This approach can be thought of as internalising (some of) the negative environmental externalities generated through the production process, although this is not done directly at the level of each product or activity. If an activity generates positive value added in the traditional sense, but that is offset by negative environmental externalities, then the net result may be zero on the level of (adjusted) GDP. This is akin to “accounting prices” proposed in the Independent Review on the Economics of Biodiversity: The Dasgupta Review (Dasgupta, 2021), which proposed to modify the ‘price’ of a unit of output to reflect its social value.

A second broad approach is the use of non-parametric techniques such as Data Envelopment Analysis (DEA). This avoids the need to assign prices or weights to different types of output (good and bad), and instead a set of equations is solved to determine the frontier of feasible outputs given available inputs. Changes in the distance to that frontier determine productivity growth. The methodology is often used in environmental and agricultural economics.

Brandt et al. (2014) attempt to apply the DEA approach, but were unable due to data limitations, and so instead estimate elasticities of various bad outputs (emissions and pollutants) with respect to the good output (GDP), solving for shadow prices of bad outputs. Gu, Hussain and Willox (2019) do similarly for Canadian manufacturing firms, exploiting firm-level data to estimate elasticities for types of pollutants. Rodríguez, Haščič and Souchier (2018) estimate country-specific elasticities, and in doing so omit some pollutants from their estimates, the growth of which they find not have significant relationships with GDP growth.

Negative environmental externalities are not limited to greenhouse gas emissions and pollutants. Frontier Economics (2017) account for the quality drinking water and river water in analysis of the productivity of the water and sewerage industry in the UK. This approach shares some similarities with public service productivity measures produced by the ONS, adjusting quantity output measures with quality measures. This may be fruitful for sector-specific studies but is likely less relevant to economy-wide analysis.

Environmental investment

A final, underexplored issue is how environmental protection expenditures are treated. As economies transition towards Net Zero, businesses and governments are making significant expenditures to reducing the environmental impact of production processes. Except where this qualifies as investment in produced assets (e.g., new machinery), this expenditure will largely be treated as intermediate consumption and thus netted out in the calculation of value added. But these expenditures could be seen as investments in "the environment", and thus would add to GDP rather than reduce it.

Agarwala and Martin (2022) estimate unmeasured environmental investment in the UK using a sum of costs approach, with adjustments to avoid double counting with activity already treated as final output (e.g., investment, exports) in the National Accounts. They estimate this adds about 5% to the level of UK GDP, although does little to change the growth rate of GDP in real terms between 1997 and 2019. Just as for currently unmeasured intangible investment (e.g., Goodridge, Haskel and Wallis, 2018), tracking additional environmental investment would have an ambiguous effect on TFP since it raises both measured output and measured inputs (through additional capital inputs).

6. GDP and Welfare

As described above, the point of departure for productivity measurement is usually a measure of gross value added or GDP. There is a long-standing debate about the inadequacy of GDP in capturing outputs that matter for our well-being and longer-term sustainability (Stiglitz, Sen, Fitoussi (2009); see Aitken (2019) for a recent discussion). As a measure of output or production, GDP captures goods and services that are produced and sold in the market at a price. It also captures aspects of goods and services produced without an explicit price, such as the output of the public and non-profit sectors, and owner-occupier housing imputed rental.[13] But, it is relatively silent about salient features of the economy and society that also impact well-being, such as household production, distribution, health outcomes and uses of the natural environment. This discussion has gained traction in light of the measured slowdown in productivity growth in advanced economies and a sharpened focus on managing climate change.

GDP, which is principally a production metric, or GDP per capita (labour productivity adjusted by hours worked per capita), may never have been a particularly good measure of welfare or broader well-being. However, its prominence as a key economic statistic has often meant that it has been used as such. This is not entirely without good reason. In so far as it measures income (through the consistent framework of the National Accounts) that can be

[13] Imputed rental is the value associated with the benefits that owner-occupiers get from living in a house that they own, instead of renting it out. It is included primarily for cross-country comparability given differences in home ownership rates across countries. It is a major reason why GDP measures more than just market production.

used for consumption, it is surely highly correlated with economic welfare over time and across countries. But, these correlations are imperfect and may break down. This could happen if the welfare effects of factors that are captured in GDP (e.g., average incomes) and the welfare effects of factors that are not well-captured in GDP (e.g., health outcomes and the environment) begin to diverge. And, as we have discussed in previous sections, this could happen because of structural change in the economy that affects the ability of GDP to fully reflect output (and associated economic welfare) in modern production structures, in particular that associated with digital outputs, product innovations, the internationalisation of production and investments in intangibles. The magnitudes and relative importance of these issues for the coherence between measured productivity growth and developments in living standards and welfare are uncertain, although the evidence reviewed in this paper does provide some clues.

Several studies consider the divergence between real GDP growth and growth in real median household disposable income (a measure closer to welfare). For example, Nolan, Roser and Thewissen (2019) consider differences between growth in GDP per capita and median household incomes in 27 OECD countries. They find that GDP per capita rose faster than median income in 23 out of these 27 countries in the 20-30 years leading up to the early 2010s. They decompose these differences into that arising from relative price movements, developments in GDP versus GNI, discrepancies between National Accounts measures and household surveys, changes in household size and last, inequality, as captured by differences in growth in mean and median household incomes. In the US, between 1979 and 2013, the rate at which growth in GDP per capita outstripped growth in median household incomes was particularly high, at 1.27 percentage points per annum; median household incomes rose by a mere 0.32 percentage points per annum. Faster measured growth in consumer prices than in the GDP deflator accounts for almost half of the shortfall in growth in median household incomes in the US. The shortfall in growth in household incomes is much larger in the US than in most other OECD countries. In the UK, between 1979 and 2010, median household income growth was much stronger than in the US and lagged growth in GDP per capita by just 0.21 percentage points per annum. The difference to the US is primarily due to differences in the development of relative price indices.

In a similar vein, van Reenen and Teichgräber (2021) decompose the apparent decoupling of median wages from productivity growth in the UK between 1981 and 2019 into the effects of non-wage labour income, different price measures, changes in inequality, and differences in data sources. Changes in inequality account for much of any divergence. Oulton (2022) demonstrates that productivity growth accounted for the vast majority of growth in median equivalised household disposable income in the UK between 1977 and 2019. Fundamentally, measured productivity growth is still important for living standards in the UK.

Other studies, building on Nordhaus and Tobin (1972), develop broader measures of welfare. The measure developed by Jones and Klenow (2016) suggests that the US, widely regarded as representing the global productivity frontier, may compare less favourably to other advanced economies on welfare measures that go beyond GDP per capita. Their measure takes into account the lifetime utility from consumption and leisure, as well as inequality in consumption and leisure, and suggests that European countries (the UK, France, Spain and Italy) are less behind the US when compared on a welfare basis than on a GDP per capita basis. This is because the European countries outperform the US on leisure time, life expectancy and equality. Aitken and Weale (2021) explore the development of this measure of welfare for lower tier local authorities in England and Wales. The welfare index of Jones and Klenow (2016) is rooted in economic theory and yields a single welfare measure, derived within an expected utility framework. Well-being is often measured using a dashboard approach. This is in part because it can be unclear how one should weight together different aspects of well-being into a single measure in a theoretically coherent manner.

Others have suggested developments of GDP and existing measurement frameworks, as well as supplementary indicators, that improve the measurement of output in the modern economy and, by virtue of this, the coherence of measured output or incomes with economic welfare. Some attempt to embed production within a broader notion of welfare. Many of these developments are reviewed in Jorgenson (2018) and some sit naturally within frameworks for productivity analysis. For example, Corrado, Fox, Goodridge, Haskel, Jona-Lasinio, Sichel and Westlake (2017) suggest developing GDP to better reflect modern economic production structures. This involves doing a better job of quality adjusting prices, measuring the value of free goods and services, and better capturing intangible and environmental capital.[14] Building on the work of Sen (1985), Coyle and Mitra-Kahn (2017) recommend an asset-based sustainability framework or balance sheet approach to measurement. To some extent this approach is already inherent to the National Accounts framework. These proposals closely align with the notion of 'missing capitals' in productivity analysis and require significant advances in measurement.

Heys, Martin and Mkandawire (2019) put forward a "spectrum" of metrics moving at one end from a market sector GVA metric, commonly used in productivity analysis, and gradually building this up to a measure of well-being. Their approach is a practical one, rooted in the National Accounts framework and is attractive for its methodological consistency. In particular, their starting point is that market prices remain the most objective way to aggregate goods and services when creating a single measure of production-based welfare, and they expand the concept of production in a way that maintains consistency of the production and asset boundaries. The latter feature is particularly important when thinking about broader output metrics in a productivity framework. Bucknall, Christie, Heys and Taylor (2021) employ this spectrum concept to develop a measure of *augmented* Net

[14] Deaton and Schreyer (2020) highlight the role of changing foreign ownership structures in decoupling GDP and welfare. For example, as in the case of Ireland highlighted above, where GDP increased by 26 per cent in 2015 and real household disposable incomes rose by less than 5 per cent. This has led to more widespread use of supplementary metrics, including GNI, in the measurement and analysis of Irish labour productivity. GNI measures income available to domestic residents, rather than income generated by the domestic economy. (Note that "inflated" output due to the inclusion of capital services associated with foreign owned intellectual property assets in GDP should not materially affect estimates of TFP).

National Disposable Income (NNDI+) for the UK using readily available data sources. In comparison to the standard Net National Disposable Income measure, this includes flows of benefits and depreciation from some of the missing capitals (some aspects of natural and intangible assets) and the household accounts (home production), as well as quality adjustment of some public services.

These estimates have since been updated by ONS in Taylor (2023), and they plan to continue routine publication under the heading of “inclusive income”. The latest estimates show that net inclusive income (equivalent to NNDI+ from Bucknall et al. (2021)), grew 17.5% in total between 2005 and 2019, compared to 20.1% for net national disposable income (the closest national accounts measure), and 22.1% for real GDP. Thus, the welfare measures grew less in total than the national accounts measures, although were in general smoother and saw less of a contraction in 2009. This is the reverse of the results in Bucknall et al. (2021) on account of methodological improvements. A large part of these new estimates are household production activities, such as caring, cleaning, and personal transportation. From a productivity perspective, the inclusion of household production will further require estimates of the inputs used in producing this output, including the time use on these activities. It is unclear how a revised or supplementary labour productivity growth series that includes domestic production would compare to existing labour productivity growth estimates.

In previous sections we considered many further issues in developing GDP to better reflect output and hence economic welfare within current production structures. The spectrum framework introduced in Heys et al. (2019), and underlying the ONS inclusive income measures (Taylor, 2023), reflect many of the developments to GDP described in this paper. This bridges the gap between GDP as it currently is, and the broader measures of welfare that are preferred by some. Overcoming the issues described in this paper, and incorporating the associated solutions in national accounts, would likely overcome many of the limitations of GDP per capita and measured labour productivity in describing trends in welfare.

7. Conclusions

In conclusion, the standard production function framework serves as a very useful tool for measurement and analysis of growth and living standards. But, it is also evident that much has happened to the production function in recent decades. To be informative, the framework requires that: inputs and outputs can be measured; for the same and relevant unit of analysis; and that the production technology is well-behaved. In this paper we have provided many examples of where one or several of these conditions do not apply, due to structural economic change arising with digitalisation and globalisation, and long-standing issues in measuring productivity associated with the distinction between real and nominal growth, the definition of value added and our ability to measure production inputs adequately. In this context, we might consider that the production function has, at least in some sense, broken down and that despite many advances in productivity measurement over the last thirty years, it remains the case that “our understanding of what is happening in our economy (and in the world economy) is constrained by the extent and quality of the available data" (Griliches, 1994).

At the same time, there are opportunities to improve understanding of productivity, through new data sources and investment in data infrastructure and analytical frameworks. We have highlighted in this paper the gains in understanding that might be made through economic measurement research, including: the further development of frameworks for valuing digital and publicly provided goods and services; the development of high-quality time-use data; supplementary metrics that better reflect productivity in national economies; the development of business microdata infrastructure; as well as measuring and incorporating "missing" capitals, intangible and natural, in analysis frameworks. Advances on these measurement issues will shed light on the causes of current productivity stagnation. Going forward, they will support the development of evidence-based policy to promote sustainable growth and prosperity.

References

Abdirahman, M., Coyle, D., Heys, R., and Stewart, W., 2020. "A Comparison of Deflators for Telecommunications Services Output", *Economie et Statistique / Economics and Statistics*, Institut National de la Statistique et des Études Économiques (INSEE), issue 517-518-5, p. 103-122.

Ackerberg, D., Caves, K., and Frazer, G., 2015. "Identification properties of recent production function estimators", *Econometrica*, Vol. 83, No. 6, p. 2411-2451.

Agarwala, M. and Martin, J., 2022. "Environmentally-adjusted productivity measures for the UK". The Productivity Institute Working Paper No.028.

Ahmad, N., Ribarsky, J., and Reinsdorf, M., 2017. "Can potential mismeasurement of the digital economy explain the post-crisis slowdown in GDP and productivity growth?", OECD Statistics Working Papers, 2017-09, OECD Publishing, Paris.

Aitken, A., 2019. "Measuring welfare beyond GDP", *National Institute Economic Review*, No. 249, August, R3-R16.

Aitken, A., and Weale, M., 2021. "Welfare Indicators for Lower Tier Local Authorities in England and Wales in 2016", presentation at ESCoE Workshop, *Ever-increasing demand for regional statistics: COVID-19 era and beyond*, June.

Altomonte, C., and di Mauro, F., 2022. *The Economics of Firm Productivity: Concepts, Tools and Evidence*, Cambridge University Press.

Anderson, G., Riley, R., and Young, G., 2019. "Distressed Banks: Distorted Decisions?", Centre for Macroeconomics DP 2019-08.

Andrews, D., Criscuolo, C. and Gal, P., 2016 "The best versus the rest: the global productivity slowdown, divergence across firms and the role of public policy", OECD Productivity Working Papers, 2016-05, OECD Publishing, Paris.

Atkinson, A., 2005. "Atkinson Review: Final report. Measurement of Government Output and Productivity for the National Accounts", Palgrave Macmillan.

Autor, D., Dorn, D., Katz, L., Patterson, C. and Van Reenen, J., 2020. "The fall of the labor share and the rise of superstar firms", *Quarterly Journal of Economics*, 135(2), p. 645–709.

Bartelsman, E. and Wolf, Z., 2018. "Measuring Productivity Dispersion", in Grifell-Tatjé, E., Knox Lovell, C. A., and Sickles, R., 2018. (Eds.) *The Oxford Handbook of Productivity Analysis*, Oxford University Press.

Basu, S., Fernald, J. and Kimball, M., 2006. "Are Technology Improvements Contractionary?", *American Economic Review*, 96 (5), p. 1418–1448.

Bean, C.R., 2016. "Independent review of UK economic statistics." *London, UK: HM Treasury.*

Bernini, M., Riley, R. & Rincon-Aznar, A., Barbera, M., Bratus, A., Sambin, N., Adderley, S., Mulhall, R., and Ramirez, P., 2017. "Industrial Clusters in England", BEIS Research Paper No. 4, September.

Besley, T., Roland, I., and Van Reenen, J., 2020. "The Aggregate Consequences of Default Risk: Evidence from Firm-level Data", NBER Working Paper No. w26686.

Bishop, A., Mateos-Garcia, J. and Richardson, G., 2022. "Using Text Data to Improve Industrial Statistics in the UK", ESCoE Discussion Paper 2022-01, London: ESCoE.

Black, R., 2022. “Estimates of markups, market power and business dynamism from the Annual Business Survey, Great Britain: 1997 to 2019”, *ONS website article.*

Bloom, N., Kawakubo, T., Meng, C., Mizen, P., Riley, R., Senga, T., and Van Reenen, J., 2021. “Do well managed firms make better forecasts?”, NBER Working Paper No. w29591.

Bloom, N., Sadun, R., and Van Reenen, J., 2017. “Management as a Technology?”, NBER Working Paper No. w22327.

Bojke, C., Castelli, A., Grasic, K., Mason, A. and Street, A., 2018. “Accounting for the Quality of NHS Output”, CHE Research Paper 153, Centre for Health Economics, University of York: York.

Brandt N., P. Schreyer and V. Zipperer, 2014. “Productivity Measurement with Natural Capital and Bad Outputs”, OECD Economics Department Working Papers, No. 1154, OECD Publishing, Paris.

Brandt, N., P. Schreyer and V. Zipperer, 2013. “Productivity measurement with natural capital”, OECD Economics Department Working Papers, No. 1092, OECD Publishing, Paris.

Brüning, N., and Mangeol, P., 2020. “What skills do employers seek in graduates?: Using online job posting data to support policy and practice in higher education”, OECD Education Working Papers, No. 231, OECD Publishing, Paris.

Brynjolfsson, E., Collis, A., Diewert, E., Eggers, F., and Fox, K., 2019. “GDP-B: accounting for the value of new and free goods in the digital economy”, NBER Working Paper No. w25695.

Brynjolfsson, E., Collis, A., Diewert, E., Eggers, F., and Fox, K., 2020. “Measuring the Impact of Free Goods on Real Household Consumption”, *AEA Papers and Proceedings*, Vol. 110, p. 25-30.

Brynjolfsson, E., Collis, A., Eggers, F., 2019. “Using massive online choice experiments to measure changes in well-being”, *Proceedings of the National Academy of Sciences of the United States of America*, 116, p. 7250–5.

Brynjolfsson, E., Rock, D., and Syverson, C., 2021. “The Productivity J-Curve: How Intangibles Complement General Purpose Technologies.” *American Economic Journal: Macroeconomics*, 13 (1), p. 333-72.

Byrne, D., 2019. “The Mysterious Cross-Country Dispersion in Mobile Phone Price Trends”, *National Institute Economic Review*, Vol. 249 (1), R39-R46.

Byrne, D., Fernald, J., and Reinsdorf, M., 2016. “Does the United States have a Productivity Slowdown or a Measurement Problem,” Finance and Economics Discussion Series 2016-17, Board of Governors of the Federal Reserve System (U.S.).

Bucknall, R., Christie, S., Heys, R., and Taylor, C., 2021. “GDP and Welfare: Empirical Estimates of a spectrum of opportunity”, ESCoE Discussion Paper 2021-08, London: ESCoE.

Cárdenas Rodríguez M., Haščič I. and Souchier M., 2018. “Environmentally Adjusted Multifactor Productivity: Methodology and Empirical Results for OECD and G20 Countries”, OECD Green Growth Papers, No. 2018/02, OECD Publishing, Paris.

Comin, D., Gonzalez, J., Schmitz, T., and Trigari, A., 2022. “A new measure of utilization-adjusted TFP growth for Europe and the United States”, NBER Working Paper No. w28008.

Connolly, M., and Spowage, M., 2021. “Improving the quality of regional economic indicators: Regional Consumer Prices”, ESCoE Discussion Paper 2021-13, London: ESCoE.

Corrado, C., Fox, K., Goodridge, P., Haskel, J., Jona-Lasinio, C., Sichel, D., and Westlake, S., 2017. "Improving GDP: Demolishing, Repointing or Extending?", Indigo Prize Entry (winning entry).

Corrado, C., Haskel, J., Iommi, M., Jona-Lasinio, C., and Bontadini, F., 2023. "Data, Intangible Capital, and Productivity", Paper originally prepared for the NBER/CRIW Conference on Technology, Productivity, and Economic Growth, March 17-18, 2022.

Corrado, C., Haskel, J., Jona-Lasinio, C., 2021. "Artificial intelligence and productivity: an intangible assets approach", *Oxford Review of Economic Policy*, Vol. 37, No. 3, p. 435-458.

Corrado, C., Haskel, J., Jona-Lasinio, C., and Iommi, M., 2022. "Intangible Capital and Modern Economies", *Journal of Economic Perspectives*, Vol 36, No 3, p. 3-28.

Corrado, C., Hulten, C. and Sichel, D., 2005. "Measuring capital and technology: an expanded framework." In *Measuring capital in the new economy* (p. 11-46). University of Chicago Press.

Corrado, C., O'Mahony, M., and Samek, L., 2021. "How does education contribute to productivity? An intangible infrastructure approach applied to the UK and the US", *36th Annual Virtual General Conference*, IARIW Plenary Session: New Developments in Economic Measurement, IARIW, 23-27 August.

Coyle, D., 2022. "Healthcare as social infrastructure: productivity and the UK NHS during and after COVID-19", The Productivity Institute Working Paper No.017.

Coyle, D., and Mitra-Kahn, B., 2017. "Making the future count", Indigo Prize Entry (winning entry).

Coyle, D. and Nguyen, D., 2018. "Cloud Computing and National Accounting", ESCoE Discussion Paper 2018-19, London: ESCoE.

Coyle, D. and Nguyen, D., 2020a. "No plant, no problem? Factoryless manufacturing, economic measurement and national manufacturing policies", *Review of International Political Economy*. Also available as ESCoE Discussion Paper 2019-15, London: ESCoE.

Coyle, D., and Nguyen, D., 2020b. "Free goods and economic welfare", ESCoE Discussion Paper 2020-18, London: ESCoE.

Dang, T. and A. Mourougane, 2014a. "Adjusting Productivity for Pollution in Selected Asian Economies", OECD Green Growth Papers, No. 2014/01, OECD Publishing, Paris.

Dang, T. and A. Mourougane, 2014b. "Estimating Shadow Prices of Pollution in Selected OECD Countries", OECD Green Growth Papers, No. 2014/02, OECD Publishing, Paris.

Dasgupta, P., 2021. The Economics of Biodiversity: The Dasgupta Review. London: HM Treasury.

Davidson, S., Black, J., Connolly, M., and Spowage, M., 2022. "Improving the Quality of Regional Economic Indicators in the UK: A Framework for the Production of Supply and Use and Input Output Tables for the Four Nations", ESCoE Discussion Paper 2022-14, London: ESCoE.

Davies, J., 2020. "International comparisons of healthcare quality", Nuffield Trust, Indicator Update, 25 June 2020: https://www.nuffieldtrust.org.uk/news-item/international-comparisons-of-healthcare-quality-1

Dawson, D., Gravelle, H., O'Mahony, M., Street, A., Weale, M., Castelli, A., Jacobs, R., Kind, P., Loveridge, P., Martin, S., Stevens, P., and Stokes, L., 2005. "Developing New

Approaches to Measuring NHS Outputs and Activity", CHE Research Paper 6, Centre for Health Economics, University of York: York.

Deaton, A., and Schreyer, P., 2020. "GDP, wellbeing, and health: thoughts on the 2017 round of the International Comparison Program", NBER Working Paper No. w28177.

De Loecker, J., 2011. "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity", *Econometrica*, Vol. 79, No. 5, p. 1407–1451.

De Loecker, J., Eeckhout, J., and Unger, G., 2020. "The rise of market power and the macroeconomic implications", *Quarterly Journal of Economics*, 135(2), p. 561–644.

De Loecker, J., and Syverson, C., 2023. "An Industrial Organization Perspective on Productivity", Chapter 3 in *Handbook of Industrial Organization*, Vol 4, Eds. Ho, K., Hortaçsu, A., and Lizzeri, A., Elsevier.

Decker, R. A., Haltiwanger, J., Jarmin, R. S., and Miranda, J., 2020. "Changing Business Dynamism and Productivity: Shocks versus Responsiveness", *American Economic Review*, 110 (12), p. 3952-90.

Dey-Chowdhury, S., McAuley, N., and Walton, A., 2021. "International comparisons of GDP during the coronavirus (COVID-19) pandemic", Office for National Statistics, February.

Diewert, E. and Fox, K., 2016. "The User Cost of Non-renewable Resources and Green Accounting". Mimeo.

Diewert, W. E., and Fox, K., 2018. "A decomposition of US business sector TFP growth into technical progress and cost efficiency components," *Journal of Productivity Analysis*, Vol. 50(1), p. 71-84.

Freeman, D., Inklaar, R., and Diewert, E., 2021. "Natural resources and missing inputs in international productivity comparisons", *Review of Income and Wealth*, 67, No.1, p. 1-17.

Djumalieva, J., and Sleeman, C., 2018. "An Open and Data-driven Taxonomy of Skills Extracted from Online Job Adverts", ESCoE Discussion Paper 2018-13, London: ESCoE.

Djumalieva, J., Lima, A., and Sleeman, C., 2018. "Classifying Occupations According to Their Skill Requirements in Job Advertisements", ESCoE Discussion Paper 2018-04, London: ESCoE.

Eaton, S. 2019. "Understanding multinational enterprises: insights from the International Business Unit and foreign direct investment statistics", ONS website article.

Eberly, J.C., Haskel, J. and Mizen, P., 2021. "'Potential Capital', Working From Home, and Economic Resilience", NBER Working Paper No. w29431.

Feldstein, M., 2017. "Underestimating the Real Growth of GDP, Personal Income, and Productivity", *Journal of Economic Perspectives*, Vol. 31, No. 2, p. 145-164.

Fernald, J. and Inklaar, R., 2022. "The UK Productivity "Puzzle" in an International Comparative Perspective", *The Productivity Institute Working Paper No.020.*

Forlani, E., Martin, R., Mion, G., and Muûls, M., 2022. "Unravelling Firms: Demand, Productivity and Markups Heterogeneity", CEP Discussion Paper No 1402.

Foster, L., Haltiwanger, J. and Syverson, C., 2008. "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?" *American Economic Review*, Vol. 98, No. 1, p. 394- 425.

Foxton, F., Grice, J., Heys, R., and Lewis, J., 2019. "The Welfare Implications of Public Goods: Lessons from 10 years of Atkinson in the UK", ESCoE Discussion Paper 2019-11, London: ESCoE.

Frontier Economics. 2017. *Productivity Improvement in the Water and Sewerage Industry in England since Privatisation: Final report for Water UK*.

Galvão A. B., 2022, "Communicating the Uncertainty of Estimates of International Comparisons of Productivity", ESCoE Discussion Paper 2022-02, London: ESCoE.

Garasto, S., Djumalieva, J., Kanders, K., Wilcock, R., and Sleeman, C., 2021. "Developing experimental estimates of regional skill demand", ESCoE Discussion Paper 2021-02, London: ESCoE.

Goodridge, P., Haskel, J. and Edquist, H., 2022. "We see data everywhere except in the productivity statistics." *Review of Income and Wealth*, 68, p.862-894.

Goodridge, P., Haskel, J. and Wallis, G., 2017. "Spillovers from R&D and other intangible investment: evidence from UK industries." *Review of Income and Wealth*, 63, p.S22-S48.

Goodridge, P., Haskel, J. and Wallis, G., 2018, "Accounting for the UK productivity puzzle: a decomposition and predictions." *Economica*, 85(339), p.581-605.

Grice, J., 2016. "National accounting for infrastructure", *Oxford Review of Economic Policy*, Vol. 32, No. 3, p. 431–445.

Grifell-Tatjé, E., Knox Lovell, C. A., and Sickles, R., 2018. (Eds.) *The Oxford Handbook of Productivity Analysis*, Oxford University Press.

Griliches, Z., 1994, "Productivity, R&D, and the Data Constraint", *American Economic Review*, Vol. 84, No. 1, p. 1-23.

Gu, W., Hussain, J. and Willox, M., 2019. "Environmentally Adjusted Multifactor Productivity Growth for the Canadian Manufacturing Sector". Statistics Canada Analytical Studies Branch Research Paper Series, Catalogue no. 11F0019M — No. 425.

Gutiérrez, G. and Piton, S., 2020. "Revisiting the global decline of the (non-housing) labor share", *American Economic Review: Insights*, 2(3), p.321-338.

Guvenen, F., Mataloni, R., Rassier, D., and Ruhl, K., 2022. "Offshore Profit Shifting and Aggregate Measurement: Balance of Payments, Foreign Investment, Productivity, and the Labor Share", *American Economic Review*, Vol. 112, No. 6, p. 1848-84.

Hall, R.E. and Jorgenson, D.W., 1967. "Tax policy and investment behavior", *American Economic Review*, 57(3), p.391-414.

Harris, L., and Martin, J., 2019. "A guide to quality adjustment in public service productivity measures", Office for National Statistics, August.

Haskel, J. and Westlake, S., 2018. *Capitalism without capital: The rise of the intangible economy*. Princeton University Press.

Hearne, D., 2021. "Regional prices and real incomes in the UK", *Regional Studies*, 55:5, p. 951-961.

Heys, R., Martin, J., and Mkandawire, W., 2019. "GDP and Welfare: A Spectrum of Opportunity", ESCoE Discussion Paper 2019-16, London: ESCoE.

Huo, Z., Levchenko, A., and Pandalai-Nayar, N., 2020. "Utilization-Adjusted TFP Across Countries: Measurement and Implications for International Comovement", NBER Working Paper No. w26803.

Jacob, N. and Mion, G., 2022. "The UK's Great Demand and Supply Recession", *Oxford Bulletin of Economics and Statistics*, DOI: https://doi.org/10.1111/obes.12533.

Jones, C., and Klenow, P., 2016. "Beyond GDP? Welfare across Countries and Time", *American Economic Review*, 106(9): 2426–2457.

Jorgenson, D., 2012. "The World KLEMS Initiative", *International Productivity Monitor*, vol. 24, pages 5-19, Fall.

Jorgenson, D., 2018. "Production and Welfare: Progress in Economic Measurement", *Journal of Economic Literature*, 56(3), p. 867–919.

Kalemli-Ozcan, S., Sorensen, B., Villegas-Sanchez, C., Volosovych, V., and Yesiltas, S., 2015. "How to Construct Nationally Representative Firm Level Data from the Orbis Global Database: New Facts and Aggregate Implications", NBER Working Paper No. w21558.

Kim, K., Petrin, A., and Song, S., 2016. "Estimating production functions with control functions when capital is measured with error", *Journal of Econometrics*, vol. 190 (2), p. 267-279.

Kirby, S. and Riley, R., 2008. "The external returns to education: UK evidence using repeated cross-sections", *Labour Economics*, Vol. 15, p. 619-630.

Koop, G., McIntyre, S., Mitchell, J., Poon, A., 2020. "Regional output growth in the United Kingdom: More timely and higher frequency estimates from 1970", *Journal of Applied Econometrics*, Vol. 35(2), p. 176-197.

Koop, G., McIntyre, S., Mitchell, J., Poon, A. and Wu, P., 2023. "Incorporating Short Data into Large Mixed-Frequency VARs for Regional Nowcasting", ESCoE Discussion Paper 2023-07, London: ESCoE.

Levinsohn, J. and Petrin, A., 2003. "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, Vol. 70, No. 2, p. 317–341.

Lui, S., Black, R., Lavandero-Mason, J. and Shafat, M., 2020, "Business Dynamism in the UK: New Findings Using a Novel Dataset", ESCoE Discussion Paper 2020-14, London: ESCoE.

Mackenzie, S. and Baybutt, C., 2022. "International comparisons of UK productivity (ICP), final estimates: 2020", *ONS website article.*

Martin, J. and Baybutt, C., 2022. "The 'F Words': Why Surveying Businesses About Intangibles is so Hard", ESCoE Discussion Paper 2022-20, London: ESCoE.

Martin, J., and Jones, K., 2022. "An occupation and asset driven approach to capital utilisation adjustment in productivity statistics", ESCoE Discussion Paper 2022-11, London: ESCoE.

Martin, J. and Mackenzie, S., 2021. "Productivity puzzle revisited: what the latest official data tells us about the UK's biggest economic issue", *Evidence to the Productivity Commission, October 2021.*

Martin, J., 2019, "Measuring the Other Half: New Measures of Intangible Investment from the ONS", *National Institute Economic Review*, 249, R17-R29.

Martin, J., 2021. "Impact of double deflation on labour productivity: 1997 to 2018", ONS website article.

Mion, G. and Tong, M., 2021. "The Impact of Offshore Profit Shifting on the Measurement of GDP: The Case of the UK", ESCoE Discussion Paper 2021-05, London: ESCoE.

Nakamura, L., 2020. "Evidence of Accelerating Mismeasurement of Growth and Inflation in the U.S. in the 21st Century," Working Papers 20-41, Federal Reserve Bank of Philadelphia.

Nakamura, L., and Hulten, C., 2021. "Expanded GDP for Welfare Measurement in the Twenty-First Century", in Corrado, C., Haskel, J., Miranda, J. and Sichel, D. (Eds.), *Measuring and Accounting for Innovation in the Twenty-First Century*, Studies in Income and Wealth, Vol. 78, NBER.

Nakamura, L., Samuels, J., and Soloveichik, R., 2017. "Measuring the "Free" Digital Economy within the GDP and Productivity Accounts," ESCoE Discussion Paper 2017-03, London: ESCoE.

Nathan, M. and Rosso, A., 2015. "Mapping digital businesses with big data: Some early findings from the UK," *Research Policy*, Elsevier, vol. 44(9), pages 1714-1733.

Neubig, T. and Wunsch-Vincent, S., 2017. "A missing link in the analysis of global value chains: cross-border flows of intangible assets, taxation and related measurement implications", WIPO Economic Research Working Paper No. 37.

Nolan, B., Roser, M., and Thewissen, S., 2019. "GDP per capita versus median household income: what gives rise to the divergence over time and how does this vary across OECD countries?", *Review of Income and Wealth*, Series 65, No. 3, p. 465-494.

Nordhaus, W.D. and Tobin, J., 1972. "Is Economic Growth Obsolete", in *Economic Growth, Fiftieth Anniversary Colloquium, V*, National Bureau of Economic Research, New York.

O'Mahony, M. and Samek, L., 2021a. "Measuring the Volume of Services Industries Output and Productivity: An Audit of Services Producer Price Indices in OECD Countries", *International Productivity Monitor*, No. 40, p. 89-104 (also available as ESCoE Technical Report 10, London: ESCoE).

O'Mahony, M. and Samek, L., 2021b. "Incorporating Health Status in Human Capital Stocks: An Analysis for the UK", ESCoE Discussion Paper 2021-03, London: ESCoE.

O'Mahony, M. and Timmer, M. P., 2009. "Output, Input and Productivity Measures at the Industry Level: The EU KLEMS Database", *Economic Journal*, 119 (538), F374–F403.

ONS, 2019. "Developing experimental estimates of investment in intangible assets in the UK: 2016", *ONS website article.*

ONS, 2020. "Firm-level labour productivity measures from the Annual Business Survey, Great Britain: 1998 to 2018", *ONS website article.*

ONS, 2021. "Investment in intangible assets in the UK: 2018", *ONS website article.*

Oulton, N., 2021. "Measuring productivity: theory and British practice", ESCoE Discussion Paper 2020-01, London: ESCoE.

Oulton, N., 2022. "The Link Between the Standard of Living and Labour Productivity in the UK: A Decomposition", *International Productivity Monitor*, (42), p.187-211.

Poquiz, J. L., 2023. "How much is your video call worth? Measuring the value of free digital goods", mimeo.

Riley, R. and Rosazza Bondibene, C., 2019. "Winners and Losers in the Knowledge Economy: Evidence from Linked Employer-Employee Data", March, mimeo.

Riley, R., Rincon-Aznar, A. and Samek, L., 2018. "Below the Aggregate: A Sectoral Account of the UK Productivity Puzzle", ESCoE Discussion Paper 2018-06, London: ESCoE.

Riley, R., Rosazza Bondibene, C., and Young, G., 2015. "The UK Productivity Puzzle 2008-2013: Evidence from British Businesses", Bank of England Working Paper No. 531.

Li, W., Mizen, P., Riley, R., and Schneebacher, J., 2023. "Managing to Adapt: Structured Management Practices and Firm Resilience", mimeo.

Schreyer, P., 2021. "Accounting for Free Digital Services and Household Production – An Application to Facebook", mimeo.

Sen, A., 1985. Commodities and capabilities. Amsterdam: North-Holland.

Sharpe, A., Sichel, D. and Van Ark, B., 2022. "Introduction to the Symposium on Productivity and Well-being, Part I", *International Productivity Monitor*, (42), p.104-116.

Shimizu, C., Diewert, E. and Nomura, K., 2023. "Improving the SNA: Alternative Measures of Output, Input, Income and Productivity for China". TCER Working Paper E-178.

Sickles, R., and Zelenyuk, V., 2019. *Measurement of Productivity and Efficiency: Theory and Practice*, Cambridge University Press.

Solow, R.M., 1956. "A Contribution to the Theory of Economic Growth", *Quarterly Journal of Economics*, vol. 70 (1), p. 65-94.

Solow, R. M., 1957. "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics*, vol. 39 (3), p. 312–320.

Stiglitz, J., Sen, A. and Fitoussi, J., 2009. *Report by the Commission on the Measurement of Economic Performance and Social Progress*.

Syverson, C., 2011. "What Determines Productivity", *Journal of Economic Literature*, Vol. 49, No. 2, p. 326-365.

Syverson, C., 2017. "Challenges to Mismeasurement Explanations for the US Productivity Slowdown", *Journal of Economic Perspectives*, Vol. 31, No. 2, p. 165-186.

Taylor C., 2023, "UK inclusive income: 2005 to 2019", *ONS website article.*

Teichgräber, A. and Van Reenen, J., 2021. "Have Productivity and Pay Decoupled in the UK?", *International Productivity Monitor*, (41), p.31-60.

Timmer, M., Miroudot, S., and de Vries, G., 2019. "Functional specialisation in trade", *Journal of Economic Geography*, Vol. 19 (1), p. 1-30.

Tol, R. S. J., 2005. "The Marginal Damage Costs of Carbon Dioxide Emissions: An Assessment of the Uncertainties", Energy Policy, Vol. 33, p. 2064-2074.

Vassilev, G., Sidhu, S., Payne, C. and King, W., 2020, "Measuring What Matters: Towards an Indicator-based Approach to Human Capital", Paper Prepared for the 36th IARIW General Conference, Oslo, Norway, August 24-28, 2020.

Ward, A., Zinni, M. B., and Marianna, P., 2018. "International productivity gaps: Are labour input measures comparable?", OECD Statistics Working Papers 2018-12, OECD Publishing, Paris.

White, T. K., Reiter, J., and Petrin, A., 2018. "Imputation in U.S. Manufacturing Data and Its Implications for Productivity Dispersion", Review of Economics and Statistics, vol. 100 (3), p. 502-509.

Wooldridge, J. M., 2009, "On estimating firm-level production functions using proxy variables to control for unobservables", *Economics Letters,* Vol. 104 (3), p. 112-114.

Zella, S., 2021. "Public service productivity: Total, UK, 2018", Office for National Statistics, April.