Productivity of tax collection in the UK, 1850 to 2019

Authors:
Josh Martin[x]
King's College London

Date:
April 2024

The Productivity Institute
Working Paper No.043

[x] Economic Statistics Centre of Excellence (ESCoE) and The Productivity Institute

Key words
Public service productivity, government output, tax collection, economic measurement

JEL codes
D24, H11, H21

Authors' contacts:
joshua.l.martin@kcl.ac.uk

Acknowledgements
I am grateful to former ONS colleagues Joe Ball, Connor Marsland and Sreeparna Bose for excellent research assistance. Thanks also to Beth Lawrence for her help and support in this work. I received helpful comments from participants of the ESCoE annual conference 2023 (May 2023), TPI Productivity Revolutions: Past and Future conference (June 2023), and a HMRC webinar (Oct 2023), in particular from John Van Reenen, Bart van Ark, Diane Coyle, Richard Heys and Ed Palmer. Thanks to an anonymous referee. The work would not have been possible without archival material made available by The National Archives. This work is heavily influenced by work by David Lagakos and Anders Jensen for the US. The work was largely carried out while the author was employed at the UK Office for National Statistics (ONS). Any views expressed are solely those of the author and cannot be taken to represent those of their employer or affiliations. Any errors are my own.

Suggested citation

J. Martin (2024), *Productivity of tax collection in the UK, 1850 to 2019,* Working Paper No.043, 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

Tax collection is an unambiguous and consistent function of government, with little change in its objective over time – it thus offers the possibility to estimate, by proxy, the productivity of government administration on a fairly consistent basis over a very long time series. We have extracted archival data on the costs and outputs of the tax collection process in the UK from 1850 to 2019, from annual reports by Inland Revenue (the predecessor to HMRC). This allows us to build numerous productivity measures for the last 170 years. The amount of tax collected is not only a function of the effort of Inland Revenue but also the available taxes – the tax base, which we include as an 'input'. We estimate the productivity of tax collection in the UK has increased by around 0.6% per year on average between 1850 and 2019, and by 0.9% per year on average between 1950 and 2019: less than the rate for the economy as a whole, but clearly not flat. We find a considerable pick up in productivity growth from around 1984 onwards, consistent with continuous digitalisation of the tax system from then onwards. We briefly document the history of UK tax collection, which exhibits regular innovation, consistent with productivity growth of the sort we estimate.

1. Introduction

The productivity of the public sector is often not measured, or where it is, it is often assumed to be constant over time. This occurs when the output of the services is measured by its inputs, as is common in the National Accounts and GDP of most economies. The UK is unusual in this regard in that the UK Office for National Statistics (ONS) produce bespoke productivity measures for public services. These allow for more accurate measurement of public service productivity, through the use of superior measures of inputs and output of various public service areas.

This paper explores the productivity of one function of government which does not (at time of writing) have direct output measurement in the ONS public service productivity statistics – tax collection.

We see three main advantages of this work. First, tax collection is an unambiguous and consistent function of government, with little change in its objective over time – it thus offers the possibility to estimate, by proxy, the productivity of government administration on a fairly consistent basis over a very long time series (indeed the rest of the economy suffers from a lack of such consistent long-run measures as well). Our measures may also prove useful in assessment of current productivity in the UK government. Second, as above, this is an area of government expenditure which is not currently well measured. This paper therefore offers inroads for official statistics. And third, it illuminates some interesting challenges and opportunities when trying to estimate productivity of a specific sector or activity over time: working with archival material, without need for classification conversion or (much) macroeconomic data, as is often the case for more broadly defined industries.

Taxation is a quantitatively important topic. Total tax receipts are equal to about 30% of GDP in the UK in recent years, while Income Tax and Corporation Tax alone (studied in this paper) are around 12% of GDP. The productivity of collecting such a large share of national income is clearly economically meaningful. Tax administration costs (i.e. the costs of running the tax administration agency) account for about 0.2% of UK GDP, and about 1% of government final consumption expenditure. While these are markedly smaller figures, they are not insignificant, and as above over a proxy for government productivity over a long time horizon.

While the broad function and objective of the UK tax system has been essentially constant throughout its existence, the scope of taxes collected, tax rates, and operations of the tax collection agencies have varied considerably. Section 3 offers a brief history of UK tax collection since 1850, the start of our measures.

We consider three measures of the productivity of tax collection, ranging from simple to more complex. These are:

1. Tax receipts to expenditure ratio – the ratio of total tax receipts to total expenditure of Inland Revenue (or HMRC, as appropriate), without adjustment for price changes, input mix or tax base.
2. Output per worker – using data on numbers of employees (and deflated expenditure on labour where necessary), and output as the (deflated) receipts of Income Tax and Corporation Tax (and other taxes on income in the past), adjusted for the tax base.
3. Output to total input – comparing Income Tax and Corporation Tax receipts (deflated), against a weighted inputs index comprising ‘conventional’ inputs of labour, capital, and goods and services, and the tax base.

In our methods we are heavily influenced by Jenson and Lagakos (2019), who carry out a similar exercise for the US.

We use a range of historic data. Much we have extracted from annual reports and other publications of Inland Revenue (IR) and Her Majesty's Revenue and Customs (HMRC), stored at The National Archives and more recently online. The nature of tax collection means that good data were collected and recorded historically, enabling such long-run measures as presented in this paper. We supplement these data with other historic sources, modelling, and assumptions, as described in Section 5 and the Data Appendix.

We find that UK tax collection exhibits significant productivity growth over long periods of time, especially since the digitalisation of the tax system began in 1984 with the computerisation of the Pay As You Earn (PAYE) Income Tax system. Our three productivity measures produce somewhat different growth rates over time, reflecting their different coverages and omissions. However, all follow similar ups and downs, and all display considerable productivity growth – starkly at odds with the implicit assumption in the National Accounts of no productivity growth in this area.

The paper proceeds as follows: section 2 provides a summary of the relevant literature; section 3 gives a brief account of UK tax history from 1850 to 2019; section 4 sets out our conceptual framework and methods; section 5 summarises our data (with more in the Data Appendix); section 6 presents our results for the three measures outlined above; section 7 offers a selection of examples and anecdotes of productivity related activities at Inland Revenue; and section 8 concludes.

2. Literature

This section provides a review of other similar studies, and a brief summary of two other related strands of literature.

There are (to our knowledge) only a small number of studies on measuring the productivity of tax collection over time. We are heavily influenced by Jensen and Lagakos (2019), who estimate the productivity of tax administration in the US over time up to 2017. They present two measures of productivity: first, "tax collections per unit of expenditure"[2] between 1867 and 2017, which they find grew at an average annual rate of 1.4%. In a second measure, they use a composite inputs index of expenditures by the Inland Revenue Service (IRS) and the tax base, and an output measure of total taxes collected. They do not distinguish between types of collection expenditures (labour, capital, intermediates), and do not apply price deflators to numerator or denominator. They find US tax collection productivity rises at an average annual rate of 1.2% between 1929 and 2017, with a plausible range of 1.1-1.4% based on a sensitivity analysis. On both measures, the majority of the productivity growth occurs in the earlier part of the time series.

For the UK, Carrera, Dunleavy and Bastow (2009) and Dunleavy and Carrera (2013) construct multiple productivity measures for UK tax collection up to 2007. Their preferred output measure is a cost-weighted index of the number of tax returns processed for a range of taxes.[3] The weights are based on unit costs of each tax. They also produce some measures with using total tax receipts as an output measure. Their input measures cover labour and intermediate inputs only. Carrera, Dunleavy and Bastow's (2009) preferred productivity measure (with the cost-weighted activity index as the output measure, and labour and intermediate expenditures as the inputs measure) grows around 4% per year on average between 1997 and 2007, while their labour productivity measure grows slightly faster. Using total tax receipts as the output measure yields about 1.5% growth per year on average over the same period. Dunleavy and Carrera (2013) present the measure based on tax receipts back to 1988, showing average annual growth in labour productivity of around 3% between 1988 and 2007,

[2] This is broadly equivalent to our first productivity measure, which we refer to as the yield-to-cost ratio, or the inverse of the cost of collection.

[3] They capture data on income tax, corporation tax, capital gains tax, inheritance tax, VAT, excise duties, and other indirect taxes. This is much broader than our second and third productivity measures, which relate only to Income Tax and Corporation Tax (and some earlier taxes on income).

and around 1.5% when also accounting for intermediate inputs. These authors also note various potential 'quality' measures, including on complaints and online tax filing, though suggest the data are not complete enough to be used systemically.

Various other studies have considered tax collection productivity at a point in time, or over a few years, but rarely over an extended period. The OECD publishes regular reports on Tax Administration via their Forum on Tax Administration (FTA) and a Tax Administration Series (TAS) database, providing comparable cross-country data on a range of aspects of tax administration.[4] Amongst this include some measures akin to productivity, including the cost of collection. However, to our knowledge, they have not considered long-term productivity growth in tax administration. Dohrmann and Pinshaw (2009) of consultancy firm McKinsey conducted a benchmarking study of tax administrations in 13 mostly high-income countries, not including the UK. This was principally a point in time (fieldwork in 2008-09) analysis, rather than considering growth over time.

There are also several accounts of the history of UK taxation, including Kay and King (1990), Daunton (2001, 2002), Clark and Dilnot (2002), and Hoppit (2021). While informative on a range of taxation issues, none of them consider productivity explicitly to our knowledge.

A second area of related literature is on the measurement of public sector output, particularly in official statistics and the National Accounts. The seminal "Atkinson Review" (Atkinson, 2005) set out key principles for measuring public sector output in the absence of market prices. In short, these recommend differentiation of services (i.e. separating out areas and functions of government as much as possible, and producing measures of input and output for each), using cost-weighted activity indices to measure the volume of output, and adjusting that volume for changes in the quality of services provided so as to mirror market valuation as closely as possible.

The UK Office for National Statistics (ONS) follow these principles in their estimates of public sector output in the National Accounts, although these measures do not include quality adjustments, which are prohibited under the European System of Accounts 2010.[5] The ONS also produce bespoke "public service productivity" statistics which do include quality adjustments, and also estimate the volume of inputs in each 'service area', such that they can estimate productivity trends since 1997. For instance, in the measurement of education services, primary and secondary school education would be separated, numbers of pupils and attendance rates would be used, in combination with unit costs per pupil, to measure output, and (only in ONS public service productivity statistics) this is adjusted for quality based on attainment, student well-being, and post-education outcomes.[6]

As of 2023, only some two-thirds of UK government expenditure has measures of output that are not based solely on cost in the UK National Accounts and ONS public service productivity statistics. In general, the outputs of individually-consumed services like education, healthcare, and social care are easier to measure. Collectively consumed services, like public administration (the activities of the civil service), are harder to conceptualise, and the relevant data are less likely to exist. Tax administration is one area that is measured by cost in these statistics, at time of writing.

Finally, there is a literature in the field of development economics that considers tax collection productivity in relation to "state capacity". That is, for a government to invest and support economic

[4] Reports in 2004, 2006, 2008, 2010, 2013, 2015, 2017, 2019 and 2021.

[5] All EU member countries are required to follow the European System of Accounts 2008 by law. Since the UK has left the EU, the ONS may choose to deviate from these rules in future. At the time of writing, the UK National Accounts are still in line with ESA 2010, and thus do not include quality adjustments in their public sector output measures, though the ONS is exploring the introduction of quality adjustments into public sector output measures in future (see NSCASE, 2023). The international System of National Accounts (SNA) 2008, followed by many non-European countries, does allow for quality adjustments to be included in public sector output measures, though very few countries do in practice.

[6] See ONS (2022) for a comprehensive review of data and methods used in these statistics, ONS (2023) for the latest statistics, and Foxton, Heys, Grice and Lewis (2021) for an account of the development of these statistics.

growth in its country, it needs money, and one of the main ways a government raises money is through taxation – the efficiency of the collection process thus, in part, may determine the state's ability to support development. See, for instance, Besley and Persson (2009) or Dincecco and Katz (2016). The literature tends to focus on lower-income countries, and consider cross-country comparisons or short-term changes, rather than long histories as we do in this paper – as such, we do not speak to this literature directly.

3. History of UK tax collection

There are several accounts of the history of UK taxation, including Kay and King (1990), Daunton (2001, 2002), Clark and Dilnot (2002), and Hoppit (2021). We do not attempt to provide such a detailed account here, but we draw from these and other sources[7] to provide a very brief summary of key changes in the UK tax collection process.

Taxes on income are a relatively modern phenomenon, having been introduced in the UK to fund the Napoleonic Wars in the late 18th century before being abolished, reinstated, and abolished again. Income Tax was reinstated, as it turned out permanently, in 1842. These early income taxes were initially levied on most types of income, but generally paid only by those with very high incomes.

The headline rate of Income Tax varied very often through the late 1800s, though was generally low by modern standards and still applied only to a small proportion of the population. The so-called "People's Budget" of 1909 sought to expand taxes on the wealthy, but due to legislative hold-ups resulted in very little tax being collected in 1909/10, and almost twice as much as usual in 1910/11. The responsibility for excise duties was passed from Inland Revenue to HM Customs and Excise in 1909 (a process that would be reversed with the creation of HMRC in 2004).

The incidence of Income Tax expanded very greatly with the onset of World War I in order to fund the war effort. Tax rates were increased sharply and individual exemptions were reduced, expanding the tax base markedly. Various profits taxes were also applied for the first time, which businesses paid in addition to Income Tax.

Tax rates were reduced somewhat in the inter-war years, although remaining above pre-War levels. Taxes on profits in excess of Income Tax were largely removed, although residual payments from those charged during WWI trickled in. With the onset of World War II, tax rates and coverage increased again.

In 1944, the Pay As You Earn (PAYE) system was established, with Income Tax deducted by employers directly from the pay-cheque of most workers; that is, by withholding income from workers. It was in part motivated by the ongoing war effort, and the need for government to collect tax more reliably, frequently, and productively. It was also partly the consequence of the large expansion in the number of liable taxpayers brought on by the increase in taxes during the war. Many of these people had never paid tax before, so withholding income allayed fears of people having problems with budgeting for their tax liabilities.

After WWII, tax rates remained high. Tax rates changed fairly frequently between the 1940s and 1970s, in part as the government used these as levers to manage the economy.[8] Great complexity in the tax system arose in the 1950s and 1960s, with many exemptions, allowances and peculiarities introduced. Corporation Tax as we know it today was introduced in 1965, with businesses no longer being liable for Income Tax. Decimalisation of the currency in 1971 also had major implications for taxation rates, calculations and collections.

[7] Including a blog by David Gauke: https://www.taxation.co.uk/articles/2011-09-21-280681-paye-story
[8] These days that is a job given more to monetary policy.

The PAYE system was an entirely paper-based and mostly local system until 1984, when it was computerised and replaced by 12 regional databases, based (principally) on the locations of each taxpayer's employer. This began the digitalisation of the tax system, with electronic data interface for businesses introduced in 1998, online filing for Corporation Tax introduced in 2003, universal online filing introduced in the mid-2000s, and paper forms being phased out entirely over 2008-2011.

HM Revenue and Customs (HMRC) was formed in 2004 as the merger of Inland Revenue (responsible for taxes on income and wealth) and HM Customs and Excise (responsible for taxes on products and imports).

We know relatively little about changes in organisational and corporate arrangements of the tax collection agencies, which are likely to be just as consequential for trends in productivity as changes in the legislative landscape. Some details are contained in the annual reports of the Inland Revenue and HMRC over the years, though these rarely consider changes over more than a few years. A book published in 1949 describes changes in Inland Revenue over the preceding 100 years (Griffith, 1949), and another in 1965 describes in some detail the arrangements at that time (Johnston, 1965). We draw on these materials in section 7.

4. Conceptual framework and methods

Section 4.1 describes the conceptual framework, and the three productivity measures employed in this paper. Section 4.2 describes our approach to parameterising the production function, and thus underpins our measurement of productivity. Section 5 describes the data.

4.1. Conceptual framework and productivity measures

We use three measures for the productivity of tax collection, each constructed in a different way with different data. All can be related in the same conceptual framework, which we set out here. In all cases, variables are for year t, but we drop time subscripts for brevity.

First, we define total taxation receipts as T. Total taxation receipts are the sum of taxation receipts from each taxation type, including Income Tax T_I, Corporation Tax T_C, and all other taxes T_O. Thus we can write:

$$T = T_I + T_C + T_O$$

Receipts of each tax type can be expressed as the product of the tax base (B_I, B_C, B_O) and the collection success rate (C_I, C_C, C_O). In turn, the tax base is the product of taxable income (I_I, I_C, I_O) and the tax rate (or rates, depending on the schedule, where the tax rate is also assumed to capture any allowances and special rules) (R_I, R_C, R_O).

Thus: $T_I = B_I C_I$ and $B_I = I_I R_I$ and so $T_I = I_I R_I C_I$ and equivalently for T_C and T_O.

The tax collection agency uses 'conventional' inputs of labour L, capital K, and intermediate goods and services G. Payments for each factor can be expressed as the product of the input volume and a corresponding price. That is, payments for labour are the product of L (which we assume is a headcount measure) and the average price of labour P_L (i.e. the cost of the average unit of L), and so total payments for labour are $P_L L$. Similarly, payments for capital are $P_K K$, where P_K is the price of a unit of capital services (which we approximate by depreciation), and payments for intermediate goods and services is $P_G G$, where P_G is the price of a unit of goods and services.

Thus total costs are given by:

$$P_L L + P_K K + P_G G$$

We define an inputs index $f(L, K, G)$, which we refer to as a bundle E of 'conventional' inputs of the tax collection agency. We assume a Cobb-Douglas functional form for the inputs index, so write it as:

$$E = L^a K^b G^c$$

where a, b and c are the elasticity of output with respect to labour, capital, and goods and services respectively.

Assuming constant returns to scale and efficient production (as is typical), $a + b + c = 1$ and each can be measured by their shares of total cost (s_L, s_K, s_G). So we set:

$$a = s_L = \frac{P_L L}{P_L L + P_K K + P_G G}$$

and equivalently for b and c. Thus our 'conventional' inputs index is given by:

$$E = L^{s_L} K^{s_K} G^{s_G}$$

Of course, the tax collection agency also 'uses' the legal authority to collect taxes embodied in the tax base B. This could be thought of as a form of institutional capital which facilitates the usability of the 'conventional' inputs of labour, capital, and goods and services. We view the tax base as an imperfect substitute for the bundle of traditional inputs.

To see the role of the tax base, consider some scenarios. If there was no legal ability to collect taxes, then irrespective of the size of the collection agency (i.e. irrespective of how much of the 'conventional' inputs there were), no taxes would be collected. A legal tax base thus enables the use of the 'conventional' inputs. Similarly, if there was a legal tax base but no conventional inputs, no taxes would be collected.

If there was a tax base and some conventional inputs, and then the tax base increased, tax receipts would be likely to increase, i.e. output would increase. This increase in tax receipts (output) should not be viewed principally as a productivity gain, since the additional output was only made possible by additional legal authority to collect more taxes (more inputs). However, if the tax base and tax receipts both increased by 1%, for a tax collection agency of a given size, this could be seen as an increase in productivity.

Increases in the tax base for a given level of conventional inputs are likely to yield progressively smaller increases in tax receipts, due to the fixed capacity of the collection agency to process additional payments. In other words, there are diminishing marginal returns to the tax base. The literature on optimal tax rates (including the "Laffer curve") also applies here – above some level, increases in tax rates may produce limited additional revenue or may even reduce revenue by increasing the incentive to avoid or evade tax. However, such arguments are not necessary for our framework.

Finally, if the conventional inputs increase but the tax base remains the same, tax receipts may rise but there is a finite limit on what can be collected (the tax base) and thus there are diminishing marginal returns to the conventional inputs. At the extreme, if tax receipts fully exhausted the tax base, then additional conventional inputs would yield no additional tax receipts.

Thus, we think it is reasonable to view the bundle of conventional inputs, and the tax base, as imperfect substitutes with diminishing marginal returns. This motivates our view of a composite inputs index as a function of both conventional inputs E and the tax base B, and we shall assume a Cobb-Douglas functional form as follows:

$$I = f(E, B) = E^{\pi} B^{\mu}$$

where π and μ are the output elasticity of the conventional inputs bundle and the tax base respectively, and by assumption $\pi + \mu = 1$ under constant returns to scale.

From here we can proceed to define our three productivity measures in terms.

Productivity measure 1: tax receipts to expenditure ratio

Following our notation, the ratio of tax receipts to expenditure (yield-to-cost ratio) can be written as

$$Productivity_1 = \frac{T}{P_L L + P_K K + P_G G}$$

The numerator (output measure) is total tax receipts, and the denominator (inputs measure) is total expenditure on inputs of labour, capital, and goods and services. Both numerator and denominator are in current prices, i.e. are not adjusted for inflation. It is usually crucial to adjust for price changes when measuring productivity growth. However, in this case, output is money itself (tax receipts), and the appropriate deflator is somewhat unclear. Given that, we can measure productivity as the ratio of two nominal figures. An advantage of this measure is thus that we do not need to determine the appropriate deflator for numerator or denominator.

The inverse of this measure (the "cost-to-yield" ratio, or "cost of collection") was included in Inland Revenue annual reports as early as 1888, and is still published annually by HMRC. It has thus clearly been of interest and relevance for the management of the tax administration system for a long time.

A drawback of this measure is that T can increase for reasons unrelated to the productivity of the tax collection agencies. Recall that T is partly dependent on the tax rates R. If the government increases R, then T will increase (assuming no change in taxable income I or the collection success rate C), for reasons unrelated to the efficiency of the collection agency. This measure is thus susceptible to changes in the tax base over time (due either to changes in income or changes in tax rates), which will reduce the accuracy of the measure. OECD (2022) describe how the cost of collection is not always a good measure of the efficiency of tax administration, especially when comparing across countries. Differences across countries in the tax base, and in the extent to which the burden of tax administration falls on the taxpayer rather than the tax collection agency, will affect comparisons.

Productivity measure 2: labour productivity with tax base adjustment

The second productivity measure we use is a more conventional labour productivity measure, where we use the volume of labour in the denominator as well as adjusting for the tax base.[9] The volume of labour is given by a composite measure, using data on full-time equivalent staff numbers between 1938 and 2019, and deflated expenditure on staff prior to 1938 – see section 4.2 for details. We do not make any adjustment for changes in the skills, experience of education of the workforce due to lack of data.

We also account for the tax base, following the framework set out earlier in this section. We restrict the output measure to just Income Tax and Corporation Tax (and some earlier taxes on income), since we are able only reliably to estimate the tax base for these taxes historically. This measure adjusts for changes in the tax base for Income Tax and Corporation Tax by including the tax base in the composite inputs measure. Labour inputs and the tax base are weighted together in a composite inputs index, discussed in greater detail below.

Following our notation, we write this productivity measure as

[9]An earlier version of this paper did not make an adjustment for the tax base for the labour productivity measure. This resulted in very fast labour productivity growth, due principally to the expansion of the tax base. We have now adjusted this measure for the tax base, to give a more meaningful set of figures.

$$Productivity_2 = \frac{\left(\frac{T_I + T_C}{P_{GDP}}\right)}{L^{\pi} B^{\mu}}$$

where P_{GDP} is the GDP deflator, used as the deflator for Income Tax and Corporation Tax revenues – see section 5.3 for discussion of deflators. Note that in this productivity measure we express both numerator and denominator in volumes (as is typical of productivity indices)

The main benefit of this measure relative to the first productivity measure (the yield-to-cost ratio) is that it accounts for the tax base. Changes in tax rates or coverage that increase the tax base will tend to increase tax revenues, but this will be adjusted for with the tax base input measure. We discuss this in more detail below.

A drawback of this measure is that it accounts for only labour inputs of the 'conventional' inputs (i.e. does not account for capital or intermediate inputs), while the first measure was relative to total expenditure and thus included these. If there is a substitution from labour to other inputs (e.g. through automation or outsourcing), then a labour productivity measure may appear to rise when in reality it is only a change in the type of input. Measuring the volume of total 'conventional' inputs requires suitable deflators for each input type, which is difficult for two reasons: first, the breakdown of costs between capital and intermediates is not known before the 1980s; and second, suitable historic deflators are scarce. The labour measure used here is a more direct input measure, which does not require the identification of a good deflator (at least, not since 1938). Our third productivity measure accounts for these other inputs.

Productivity measure 3: total factor productivity with tax base adjustment

Our third productivity measure is the most comprehensive, accounting for 'conventional' inputs of labour, capital, and goods and services, alongside the tax base. As for the second measure, we restrict the output measure to just Income Tax and Corporation Tax (and some earlier taxes on income), for consistency with our tax base measure.

This productivity measure, which we shall refer to as a total factor productivity (TFP) measure, can be expressed as

$$Productivity_3 = \frac{\left(\frac{T_I + T_C}{P_{GDP}}\right)}{E^{\pi} B^{\mu}}$$

The inputs measure accounts for labour, capital, and goods and services. The volume of labour is as in the second productivity measure (labour productivity), based on staff numbers back to 1938, and before that deflated expenditure on staff. The non-labour inputs are measured in two ways. From 1980 to 2019, we can identify capital and intermediate inputs separately, and so account for each separately, weighting them with labour according to their shares of current price expenditure. The volume of these inputs is estimated by applying deflators to the current price expenditure data (see section 5.3), and the growth rates of each input are weighted by their shares of expenditure. Prior to 1980 we can only identify total non-labour costs, and so use this instead of separate capital and intermediate input indices. Thus, we still account for 'total' inputs, but with less precision prior to 1980.

We use natural log changes and rolling two-period average weights to construct a Tornqvist index, as is preferred in productivity measurement, both for the conventional inputs bundle E, and the total inputs measure.

The advantages of this measure are that it accounts for all inputs, unlike the labour productivity measure. It also accounts for the tax base, which the yield-to-cost ratio did not. It is thus the most

comprehensive measure in this paper. The main drawback is the complexity involved in the estimation, including in estimating the tax base and parameterising the production function, which introduce uncertainty and potential measurement error.

We turn to parameterising the production function next.

4.2. Parameterising the production function

We follow Jensen and Lagakos (2019) (henceforth J&L) to parameterise the production function of tax collection.

First, we exploit the insight that marginal products are proportional to average products under constant returns to scale. That is, inputs at the margin have the same impact on output as that input has on average. This is intuitive: the tax collection agency will allocate its budget optimally, such that the marginal expenditure produces the average return; if they did not, they would be better served allocating their budget differently. The tax collection agency is budget constrained, competing as it does with other government departments for funding, and governments in turn are incentivised against inefficiency by the electoral system. Indeed, the HMRC annual report 2013-14 shows that compliance staff are deployed roughly proportionally to the estimated size of the tax gap across taxpayer groups (HMRC, 2014).[10]

J&L identify audits as the marginal expenditure of the Inland Revenue Service (IRS) in the US. They draw on two studies of the marginal value of audits in the US to estimate the marginal return. First, Malanga (1986) estimates that every $1 spent on audits returns $13.9 in extra tax collections on average. Second, Tauchen, Witte, and Beron (1993) estimate that for every $1 of direct benefits of audits, namely extra tax collections from those taxpayers who are audited (i.e. those identified by Malanga (1986)), an additional $2 of indirect benefits accrue through "deterrent effects" on non-audited taxpayers. For instance, if a taxpayer observes that another taxpayer is audited, then they may be less likely to commit fraud in their subsequent tax returns – the benefits to the collection agency are thus any extra collections from the audited taxpayer, and the prevented fraud from any non-audited taxpayers. Combining the direct and indirect effects gives a total return of approximately $42 (three times $14) for each $1 of audit expenditure.

We replicate this approach for the UK drawing on data published by Inland Revenue and HMRC. On the direct returns, we could not identify an overall yield ratio from audits for the UK. However, various Inland Revenue reports contain the yield ratio of different audit activities across a number of years. We obtained over 100 such ratios for years between FYE 1989 and FYE 2004, which are summarised in Table 1. Clearly there is a wide range of ratios, reflecting a range of audit activities.[11] Without information on the resource allocated to each type of audit it is not possible to determine an overall ratio. However, we judge that the overall median of 9 is fairly representative of overall audit activities.

Our estimate of £9 of direct benefits for every £1 of audits is a little lower than the ratio of 13.9 found by Malanga (1986) and used by J&L. We judge this to be appropriate given the greater onus on taxpayers to file tax returns in the US than in the UK, and thus greater scope for fraud and error. J&L also report other figures in the literature, including a ratio of 6 in the US (Dubin, Graetz and Wilde, 1986) and a range between 3.7 and 22 in Belgium (De Neve et al., 2021), among others. Our estimate of 9 is thus consistent with this range.

[10] Figure 14, Report by the Comptroller and Auditor General, Part Three, R37. HMRC Annual Report and Accounts 2013-14.

[11] The names of audit activities change across each report, such that it is difficult to compare individual activities over time. Audit activities with the highest yield ratio relate to "special investigations", "large businesses" and "international", while activities with the lowest yield ratios tend to be "full enquiry work", namely "Self Assessment business enquiries".

Table 1 – Summary of yield-to-cost ratios of various audit activities of Inland Revenue

Year	Range	Median	Mean (unweighted)
1988/89	4.5 – 86	17.0	22.4
1989/90	5 – 89	16.5	24.6
1997/98	3.2 – 206	12.5	38.9
1998/99	0.3 – 131	6.1	26.5
2000/01	0.9 – 96.4	8.4	23.9
2001/02	1.3 – 148	8.9	29.4
2002/03	1.3 – 350.1	8.7	47.8
2003/04	1.3 – 419.3	9.4	53.2
All years		9.0	33.6

Source: Inland Revenue, authors' calculations.

For the indirect effects of audits, HMRC report estimate "revenue protected" and "future revenue benefit" through audit activities, equivalent to the concepts in Tauchen, Witte, and Beron (1993) for the US. HMRC describe revenue protected as "the value of our activities where we have prevented revenue from being lost to the Exchequer; for example, by stopping a fraudulent repayment claim… it also includes the impact of our compliance work to disrupt criminal activity" and describe future revenue benefit as "the effects of our compliance interventions on customers' behaviour".

Using the estimate of extra tax receipts collected from audits[12], and the indirect benefits described above, we calculate that indirect effects are on average around £2.40 for every £1 of direct benefits. The exact ratio varies from year to year, from around 2 to 2.6 in the latest estimates. Before explicit estimates were made of these indirect benefits, Inland Revenue reports in the 1990s referred to these as "substantial" and "effectively doubles the direct yield from the investigations" – this would be consistent with our estimate using more recent data.

Our indirect-to-direct benefit estimate of 2.4 is a little lower than the estimate of 3 found by Tauchen, Witte, and Beron (1993) and used by J&L for the US. As for the direct estimate, we think this is appropriate given the greater onus on taxpayers to file tax returns in the US than in the UK, and thus greater potential for taxpayers to be influenced in their behaviour when filing their tax returns relative to the UK.

Taking our direct and indirect benefit estimates together gives an overall yield of approximately £21.6 (2.4 times £9) per £1 spent on audits in the UK. This is around half of the overall ratio of 42 used by J&L. We judge that a plausible range for the UK is 14 to 35.

With this in hand, we turn to estimating the output elasticity. First, we can write algebraically the insight that marginal products are proportional to average products in constant-returns production functions.

$$\frac{\partial T}{\partial E} = \pi \frac{T}{E}$$

where $\frac{\partial T}{\partial E}$ is the marginal product of tax collection expenditures, and π is the output elasticity of the bundle of 'conventional' inputs E.

[12] HMRC describe this as "cash expected" since some of the extra receipts are expected to accrue in future years, and some will not be collectable (e.g. due to insolvency) and so the figure is partly estimated.

Recall that our empirical estimate of $\frac{\partial T}{\partial E}$ is 21.6, which we can denote $\hat{\gamma}$. Note that we observe T and E since this is our first productivity measure (the inverse of the cost of collection, or yield-to-cost ratio). With these in hand, we can estimate π.

Assuming that this equation is measured with additive error that is independently and identically distributed over time, we can write this equation as a linear regression problem, and estimate it using the method of moments. Our estimate of π is thus:

$$\hat{\pi} = \hat{\gamma} \frac{\sum_{t=1}^{T} \frac{T_t}{X_T}}{\sum_{t=1}^{T} \left(\frac{T_t}{X_T}\right)^2}$$

This mirrors the standard estimator in a linear regression problem.

We obtain an estimate of π of 0.30, on the basis of estimating this equation on data covering 1850 to 2019. However, the studies we use to estimate the direct and indirect effects are from only the past couple of decades, so it may be more appropriate to use only that data to estimate the equation. We find that varying the time horizon alters the estimate of π little, as shown in Table 2. The estimate of π is clearly much more sensitive to $\hat{\gamma}$, our empirical estimate of the marginal product (of audits). Overall, we judge that a plausible range for π is between 0.2 to 0.5.

Table 2 – Sensitivity of output elasticity of conventional inputs bundle to time horizons and marginal product estimates

Time range used in estimator	**Lower estimate of $\hat{\gamma}$ (=14)**	**Central estimate of $\hat{\gamma}$ (=21.6)**	**Upper estimate of $\hat{\gamma}$ (=35)**
1850-2019	0.20	0.30	0.49
1950-2019	0.19	0.30	0.48
1980-2019	0.18	0.28	0.46

Source: Authors' calculations.

Jensen and Lagakos (2019) find $\pi = 0.19$ for the US, which is at the lower end of our plausible range. As described previously, more of the onus is on taxpayers to file their own tax returns in the US than in the UK, and thus the IRS arguably has a relatively smaller role in tax collection in the US than HMRC/Inland Revenue do in the UK. It would thus be appropriate for the output elasticity of the 'conventional' inputs to be larger in the UK than in the US. It is, however, reassuring that their estimate lies on the edge of our plausible range for the UK.

Assuming constant returns to scale, and thus $\pi + \mu = 1$, implies $\mu = 0.70$ (J&L have $\mu = 0.81$), with a plausible range for μ between 0.5 and 0.8. We do not vary these parameters over time given a lack of information on the variation of marginal products over time.[13]

Thus, the central estimate for our third productivity measure is:

$$Productivity_3 = \frac{\left(\frac{T_I + T_C}{P_{GDP}}\right)}{E^{0.3} B^{0.7}}$$

We provide a sensitivity analysis of this productivity measure to different estimates of π and μ in Appendix A.

[13] It seems plausible that the output elasticity of the conventional inputs bundle has fallen over time, consistent with Table 2, as the tax base has grown. However, we do not make any estimate of this.

5. Data

Much of the data we use comes from annual reports of Inland Revenue (the predecessor to HMRC), stored at The National Archives. Data was extracted and transcribed on three visits in 2019, 2021 and 2023, from the site in Kew in England. We supplement this with various other sources including the Bank of England's Millennium of macroeconomic dataset[14] (Thomas and Dimsdale, 2017), various ONS datasets, Brian Mitchell's British Historical Statistics (Mitchell, 2011), HMRC's annual reports, and various other sources. See the Data Appendix for more details.

The main data variables of interest from the Inland Revenue (IR) and HMRC reports relate to tax revenues and the 'conventional' inputs of the tax collection agency. We summarise these in section 5.1. Sections 5.2 and 5.3 describe the calculation of the tax base, and the choice of deflators, respectively. The Data Appendix provides additional detail.

5.1. Inland Revenue data

Cost of collection ($1/Productivity_1$)

Our first productivity measure ($Productivity_1$) is the inverse of the "cost-to-yield ratio" or "cost of collection" which is included in IR and HMRC reports for benchmark years from financial year ending (FYE) 1848 to 1884, and then every year thereafter. It is broadly the ratio of total revenues from tax collection to total cost of the tax collection activities, though the inclusions in both revenues and costs change somewhat over time. Revenues include tax receipts collected by other agencies/bodies on behalf of IR, and exclude receipts collected by IR on behalf of other bodies. Costs include payments to other government departments/bodies for services, and exclude costs borne by IR in providing services to other government departments.

We splice through a number of changes in the scope of the organisation over time, notably the integration of the Contributions Agency (responsible for the collection of National Insurance contributions) into IR in 2000, and the formation of HMRC in 2004. Both of these events changed the scope of revenues considerably, affecting the aggregate cost of collection. To avoid these leading to discontinuities in our first productivity measure, we splice through them, maintaining the 'level' of the series as it was before FYE 2000, i.e. as close as possible to the scope of our other productivity measures. See the Data Appendix for more details.

Since around 1980, the cost of collection of different types of tax have been reported separately. For greater consistency with our other productivity measures, which include only Income Tax and Corporation Tax in their output measures, we construct an aggregate cost of collection for only these taxes, weighting together their individual costs of collection by their shares total combined tax revenue. This measure is thus largely immune to the organisational changes in 2000 and 2004 described above, and shows a similar trend to our adjusted measure – see Appendix A.

Income Tax receipts (T_I)

Our measure of Income Tax receipts is as reported by Inland Revenue and HMRC in their annual reports. Data exist for almost all years since 1841, with the exception of a handful of years in the 1800s which we impute for.

We separately account for Surtax receipts (also known as Super-tax before 1929). The Surtax was an additional tax on top of Income Tax for high earners. It was introduced in 1910 and abolished in 1973. While it was a separate tax, it was a tax levied on individual income, so we combine this with Income Tax receipts henceforth.

[14] Available from: https://www.bankofengland.co.uk/statistics/research-datasets.

Our modern measures cover specifically Income Tax, so this does not include employees' or employers' National Insurance contributions. It does, however, include Income Tax paid by the self-employed via self-assessment. Prior to the creation of Corporation Tax in 1965, businesses were also liable for Income Tax on their profits, which are included in our measure of Income Tax receipts.

The receipts are recorded against the year in which they are received, rather than the year to which the income relates. This may give rise to minor timing differences relative to our estimate of the tax base, and the 'conventional' inputs, though these should be small and balance out over time.

Corporation Tax receipts (T_C)

We group all taxes on business income and profits under the heading of "Corporation Tax", although in practice there are a wide range with different names and coverage. Before the creation of Corporation Tax in 1965, businesses were liable for Income Tax on their profits, and these revenues are included in Income Tax receipts as described above. Additional taxes specifically on business profits are included under the heading of Corporation Tax receipts and include: Excess profits duty (FYE 1915-1921), Corporation profits tax (FYE 1921-1924), Profits tax (FYE 1938-1965), and Excess profits tax (FYE 1940-1946). Details of the inclusions, and brief descriptions of these various taxes on profits, are in Appendix C. From 1965, our measure includes just Corporation Tax receipts, which has existed in broadly its current form since then.

As for Income Tax described above, Corporation Tax receipts (and other profit taxes) are recorded against the year in which they are received. In the case of the various short-lived profit taxes, this means that some residual revenues are recorded for years after the tax is abolished.

Staff numbers (L)

Our measure of the volume of labour input is a composite measure, using the best possible data over different periods of time. Appendix B provides more detail and we present here just a summary. We use data on full-time equivalent (FTE) staff numbers (of varying methods) back to FYE 1939, and prior to that a measure of deflated expenditure on staff (described below). We also adjust as far as possible for changes in the scope of activities carried out by Inland Revenue and HMRC over time, in order to avoid these changes causing sharp movements in our labour input measure.

With the formation of HMRC in 2004, the number of staff increases sharply. The scope of activities conducted by HMRC is broader than Inland Revenue (encompassing also the trade duties and levies that were previously the responsibility of HM Customs and Excise). For consistency in scope over time, and for consistency with our output measure, we splice on the growth in the HMRC growth rates to the level of Inland Revenue staff in the year of the merger. Over the full time series, we also remove (as far as possible) staff relating to the Valuation Office (VO), since they account for a fairly large and variable number of staff, and our output measure does not include output relating to valuation. We extend the series prior to FYE 1939 using deflated staff costs (i.e. staff costs adjusted for inflation). To this we collect a series of data on nominal staff costs (see more below) and construct a deflator to best represent wages of Inland Revenue staff (see section 5.3).

The *number* of staff is not necessary for our productivity calculations. Only the annual changes in the volume are relevant (put another way, the units are irrelevant). Nonetheless, Figure 1 plots our series of the implied estimate of (relevant) staff on a full-time equivalent basis between 1850 and 2019. The peak staff level was in FYE 1978.[15] Although the combined staff of Inland Revenue and HM Customs and Excise, as combined in HMRC in 2004, was larger, this reflects staff engaged in a broader set of

[15] The IR annual report FYE 1982 reports that the peak was of 85,614 in October 1977, though this relates to all staff including the Valuation Office staff which we exclude. Our series peaks at 75,706 in FYE 1978 (the same year as the reported peak).

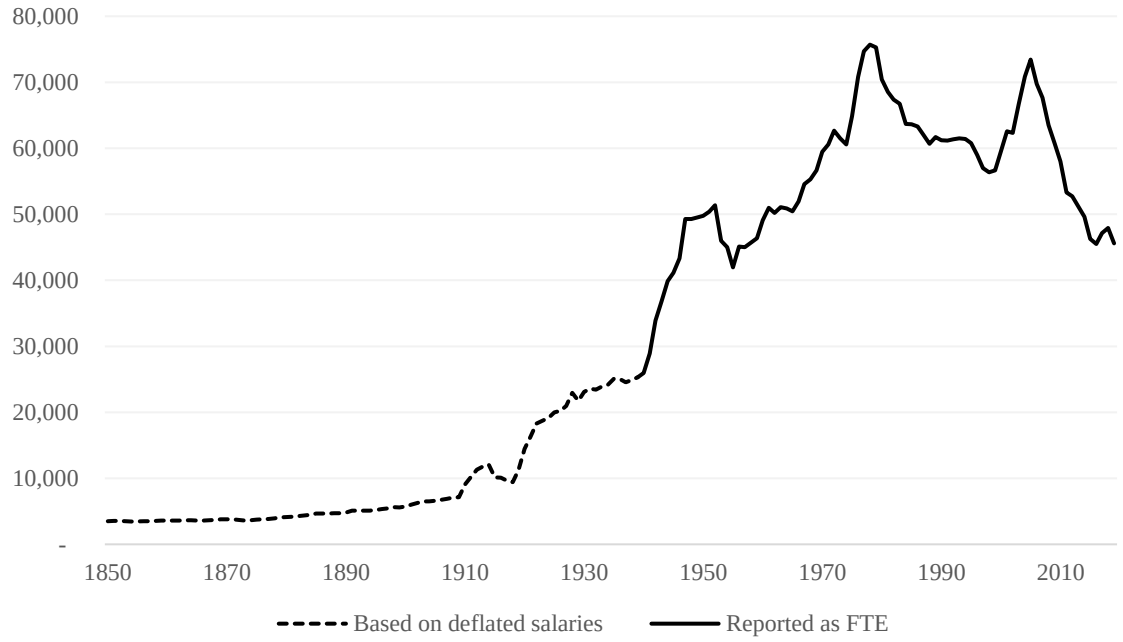
activities. Once splicing through this, as above, FYE 1978 remains the peak of our labour volume series. There series also spikes in FYE 2005, the year of the creation of HMRC, before falling sharply.

Labour costs ($P_L L$)

We use data on labour costs to extend the labour volume series (described above), and to inform the weight on labour in the aggregation of the 'conventional' inputs bundle (with intermediate inputs and capital) in our third productivity measure. As for the data on staff numbers, we adjust for various changes in the scope of Inland Revenue and HMRC over time.

To extend the volume estimate, we need to match some data on labour costs with a suitable deflator. Our labour cost deflator is a wage deflator, constructed from various sources (as described in section 5.3). This is already challenging, but would be far more so if we tried to also account for non-wage labour costs pre-1939. Non-wage labour costs were, in any case, of little importance prior to the expansion of the welfare state in 1948 – recall that we have staff numbers for our volume estimate back to 1939. As such, we use a series on salary costs only (i.e. excluding non-wage labour costs such as pension costs) for this. Salary costs are reported in IR annual reports as "salaries and allowances" from 1850, and we have a near complete time series, with only a few gaps in the mid-1800s which we interpolate for.

Figure 1 – Labour volume estimate, full-time equivalent staff numbers, financial year ending 1850 to 2019



Source: Inland Revenue, HMRC, authors' calculations.
Notes: The level is set by the number of staff at Inland Revenue, excluding the Valuation Office, between 1939 and 2005, extended on either side in various ways. It is therefore lower than the current staff employed at HMRC, since it approximates the volume of staff involved in the scope of activities previously done by Inland Revenue excluding the Valuation Office.

For weighting labour in the 'conventional' inputs bundle (with intermediate inputs and capital) we prefer a measure of total labour costs, including pension costs, employers' National Insurance contributions, and other non-wage labour costs.[16] From FYE 1994 onwards we have a robust measure

[16] A conceptual issue arises from the extent to which these costs relate to contemporaneously employed workers. Employers' pension costs may in part reflect the cost of previously employed workers, which relate little to the current production

of total labour costs from IR and HMRC annual reports. For consistency with our measure of intermediate inputs (see below) we restrict staff costs as far as possible to the scope of Inland Revenue, and adjust for various mergers and changes in responsibilities. Prior to FYE 1994, we splice on a series of combined salary costs (as described previously) and pension obligations (with some empirical challenges, described in the Appendix).

Expenditure on goods and services ($P_G G$)

Expenditure on goods and services (i.e. intermediate inputs) are only separately identifiable in the Inland Revenue records from the mid-1970s onward. From FYE 1976 onwards we have considerable detail on the value and composition of intermediate inputs, albeit with several changes in categories and reporting over time. As for staff numbers and labour costs, we adjust for changes in departmental scope where possible, including most notably for the formation of HMRC.

Prior to FYE 1976 we can identify total non-labour costs (reported as "other charges" in IR annual reports) to a reasonable degree. As we cannot separately identify it, we assume this also includes costs of capital investment and depreciation of existing capital. We thus interpret this measure as total non-labour inputs (comprising intermediate inputs and capital inputs), rather than just intermediate inputs. For this earlier period, we also include the costs borne by other government departments for services provided to IR, such as postage and stationery services by HM Stationary Office, costs for accommodation and associated services (e.g. cleaning) by PSA and earlier the Office for Works.

Finally, we link the two sets of data – detailed intermediate inputs information from FYE 1976 onwards, and less detailed prior to this – with the more detailed data taking precedent.

Depreciation ($P_K K$)

The preferred measure of capital input for productivity analysis is capital services, which measures the flow of productive services from a capital stock. Constructing such as measure for tax collection would be difficult and not worthwhile, given that capital is a relatively small input. As such, we follow the approach of ONS public service productivity statistics in using the deflated value of consumption of fixed capital (depreciation) as the capital input measure.

Data on capital assets is recorded inconsistently over time, and we combine various measures, making some imputations and judgements to produce a time series that starts in FYE 1980. Depreciation is first reported in the FYE 2000 IR annual report, with data for FYE 1999 and 2000, and reported annually thereafter. As usual, we splice through the creation of HMRC, maintaining the level reported by IR up to then. We include amortisation (depreciation of intangible assets) where reported.

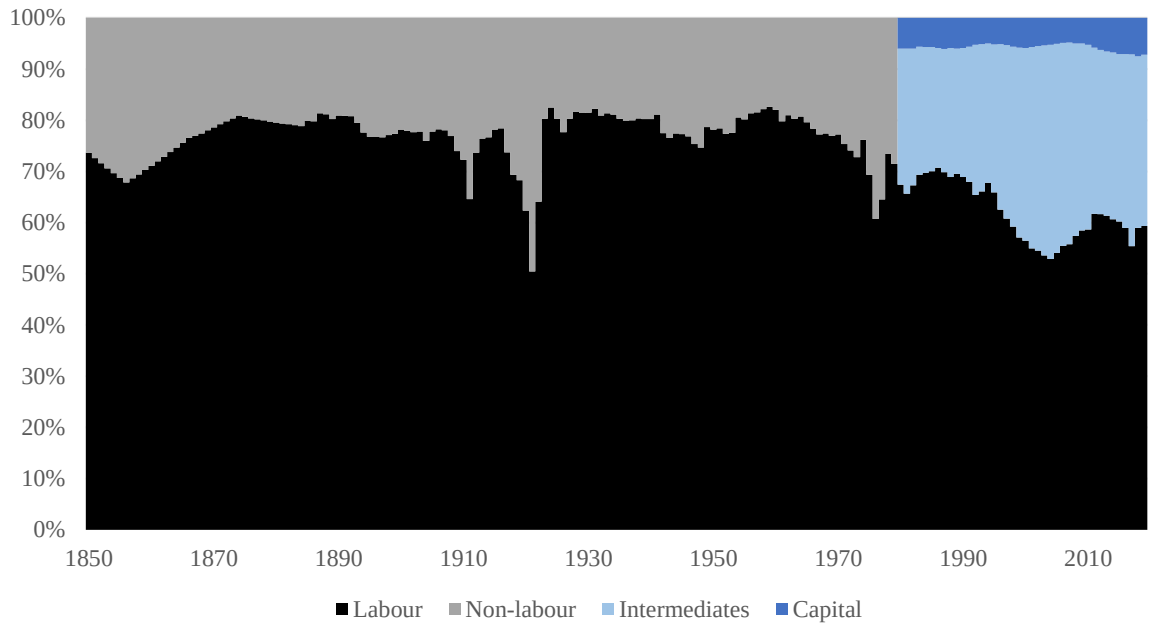
Prior to FYE 1999, with no available data on depreciation, we make the assumption of a steady state capital stock in order to impute depreciation. If the capital stock is a constant volume, then inflows to the capital stock (investment) match outflows (depreciation). As such, depreciation can be estimated as equal to investment. In years where we have data on both investment and depreciation, the annual values are similar on average, although the investment series is naturally more variable than the depreciation services. To avoid imparting unrealistic volatility into our capital input measure, we estimate a linear regression through the combined capital investment and depreciation data, and use those fitted values as the estimate for consumption of fixed capital. While a rough measure, we think our estimate is reasonable, and preferable to the unadjusted volatile series. It is, in any case, a small input, so improvements would make little difference to the result.

Figure 2 shows the composition of total costs in our 'conventional' inputs bundle of labour, capital and intermediate inputs. Prior to FYE 1980, we identify only total non-labour inputs (comprising

function. Given empirical challenges, we take the more straightforward route of using total labour costs as incurred at the time in the weighting of the production function. In practice, we expect this should make relatively little difference.

capital and intermediate inputs combined). By this measure, tax collection has always been a labour-intensive activity, with labour costs accounting for over 70% of total cost on average between 1850 and 2019. That has fallen over time, with labour accounting for around three-quarters of total cost on average up to the mid-1975s, after which other costs begin to grow, and labour's share falls, to finish at around 60% in the past decade. Capital accounts for around 6% on average since 1980, with intermediates around a third of the total. Intermediates peaked as a share of total costs at around 40% in the early-2000s, consistent with the large private service contracts around that time, but has since fallen back.

Figure 2 – Shares of total cost, nominal values, financial year ending 1850 to 2019



Source: Inland Revenue, HMRC, authors' calculations.
Notes: Labour costs = total labour costs, including pension costs and other non-wage labour costs. Intermediates and capital only separately identified from FYE 1980 onwards, prior to which total non-labour costs are shown, which are assumed to cover intermediates and capital.

5.2. Tax base (*B*)

We define the tax base as the theoretical amount of tax available to be legally collected given the prevailing tax rates and levels of labour and business income. We approximate the tax base for Income Tax and Corporation Tax by multiplying the levels of aggregate labour and capital income by the average of the main tax rates in each year.[17] While crude, we believe this provides a reasonable approximation for the changes in the tax base over time. Specifically, we are ignoring complexities relating to allowances, variations in rates by income level, and special rates for special groups. For instance, for the Income Tax base in FYE 2019, we average the basic rate of 20% and the higher rate of 40% (giving 30%), and multiply this by total wage and salary income, sourced from the National Accounts.

We repeat this exercise back to 1850 for Income Tax, to 1965 for Corporation Tax, to 1910 for the various other profit taxes levied before Corporation Tax (see Annex C), and between 1910 and 1973

[17] The income of the self-employment – "mixed income" in the National Accounts" – is included in the tax base measure, with the Income Tax rate applied.

for surtax (levied on labour income in addition to Income Tax). The Data Appendix describes our calculations in more detail.

We use National Accounts data on factor incomes in GDP. Data come from ONS National Accounts and Mitchell's British Historical Statistics (Mitchell, 2011) which in turn is based on a range of sources including Feinstein (1972). Using National Accounts data throughout should provide relative consistency, although of course the measure is at best an approximation for the true value.

5.3. Deflators

For two of our productivity measures, we need deflators for either inputs (labour, goods and services, capital, tax base) or output (tax revenue). We source these principally from ONS and the Bank of England's millennium of macroeconomic data (Thomas and Dimsdale, 2017).

For the labour deflator, which we only need in years prior to FYE 1939, we use a combination of sources, obtained from (Thomas and Dimsdale, 2017) and Mitchell (2011). Much historic wage data from this time related to labourers and workers in agriculture and manufacturing, which would not be appropriate for Inland Revenue staff. We have therefore, where possible, used wage indices for relevant industries or occupations to better reflect the likely wage trends of Inland Revenue. Our composite wage index, comprising all of these sources and periods, exhibits less wage growth between 1850 and 1910 than an average index for all workers, but is similar to this average from 1910 onwards. See the Data Appendix for more details.

For the goods and services deflator, we use the GDP deflator. We experimented with a bespoke deflator that accounted for the product mix that we could identify from the breakdown in the Inland Revenue accounts. However, with many changes in the available breakdown of expenditure over time, and limited high-quality services deflators available for years prior to 1997, we could not be confident that this would be higher quality than the GDP deflator.

For the capital deflator we again use the GDP deflator. We experimented with the implied capital stocks deflator for the central government sector from the ONS capital stocks datasets, although this only extends to 1995, and would include to a large degree irrelevant assets such as roads.

Finally, for the deflator on output (tax revenues) we use the GDP deflator. There is no obvious choice here, since the output is money itself. A general price index is therefore appropriate, so we considered a consumer price index (such as the CPI or RPI historically), the GDP deflator, the implied GVA deflator, and the implied general government final consumption expenditure deflator. We chose the GDP deflator since it would not be significantly affected by the prices of imported goods or consumer trends, is available over a very long time period, and should be more consistent over time than CPI.

6. Results

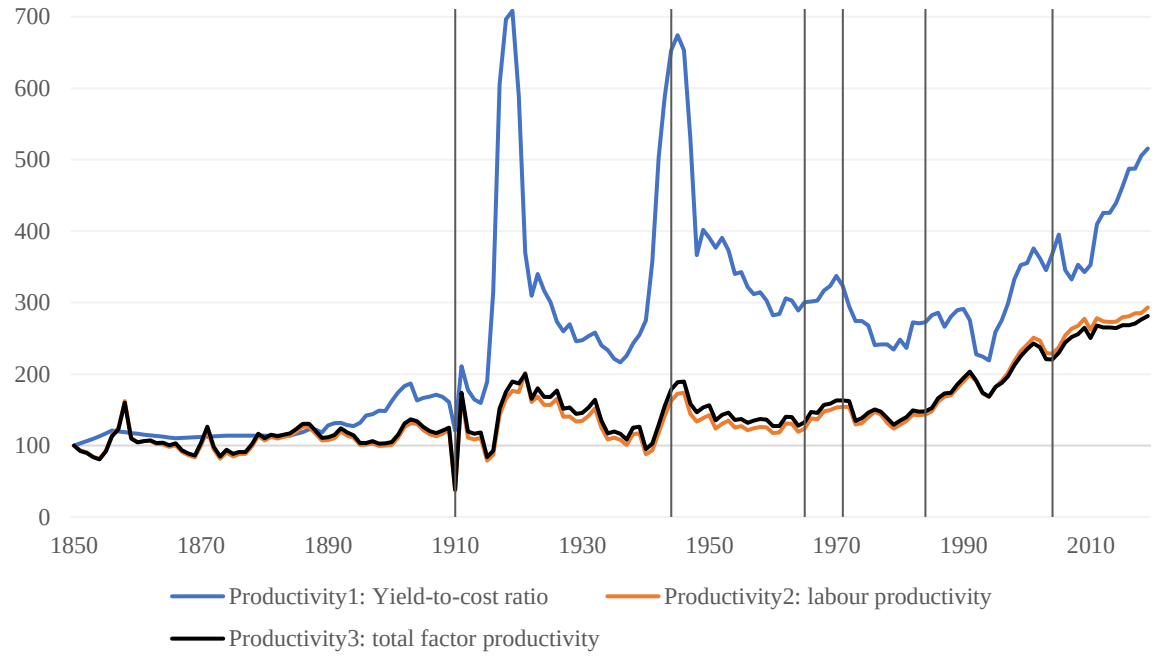
Figure 3 shows our three productivity measures for tax collection, indexed to equal 100 in 1850. Important reference periods are indicated by vertical lines, defined in the notes to the chart. Figures 4 and 5 show the same data, indexed instead to 1950 and 1997 respectively, in order to better show trends over more recent years.

All three measures of productivity increase over time, and in the case of the yield-to-cost ratio, very considerably. The yield-to-cost ratio (Productivity1), which does not adjust for the tax base as an input, shows peaks in productivity during the first and second world wars, consistent with an increase in tax revenues as a result of an increase in tax rates over this period (which goes unaccounted for in this measure). The tax base-adjusted labour productivity (Productivity2) and total factor productivity (TFP) (Productivity3) measures also shows an increase during these periods, albeit a much smaller

one, which likely reflects an increased willingness of the public to pay their taxes in support of the national war effort.

Our preferred measure, tax base-adjusted TFP, shows relatively little growth in productivity of tax collection between 1850 and early 1900s, consistent with relatively little change in organisational practices over this period. The war and inter-war periods were marked by regular changes in tax policy and an expansion of Inland Revenue, so while tax receipts increased, so too does our estimate of the inputs. The introduction of PAYE in 1944 (the second vertical line from the left in Figure 3) appears to leave productivity a step-change higher relative to the period before the wars, across all three measures, although substantial changes in tax policy in the intervening period make it difficult to precisely quantify this effect.

Figure 3 – Tax collection productivity measures, index 1850 = 100



Source: Various sources (see text), authors' calculations.
Notes: Vertical lines correspond to (from left to right): 1910 – People's Budget, 1944 – PAYE create, 1965 – Corporation Tax created, 1973 – Decimalisation, 1984 – PAYE computerised, 2004 – HMRC formed.

Regular changes in tax policy continued through the 1950-80s, with tax policy often used to manage the economy. Tax rates were also high during this period, which may have reduced tax receipts for a given level of the inputs, by incentivising avoidance or evasion (so-called "Laffer curve" effects) – this is another sense in which these productivity measures are 'social' measures. Indeed, we estimate a fall in productivity over this period. The introduction of Corporation Tax in 1965, replacing Income Tax for businesses, sees a small increase in productivity in the subsequent years, suggesting this was a relatively smooth transition. By contrast, decimalisation in 1971 appears to temporarily weaken productivity, perhaps due to increases in error (and/or fraud) in the early years of the new standard.

There is some apparent cyclicality in productivity exhibited in Figure 4, with declines around (and shortly after) the recessions in 1973-74, 1979-80, 1990-91 and 2008-09. This will reflect a relative decline in tax receipts (output) as national income falls, relative to conventional inputs of the tax administration agency, which in the short-run are likely relatively fixed.

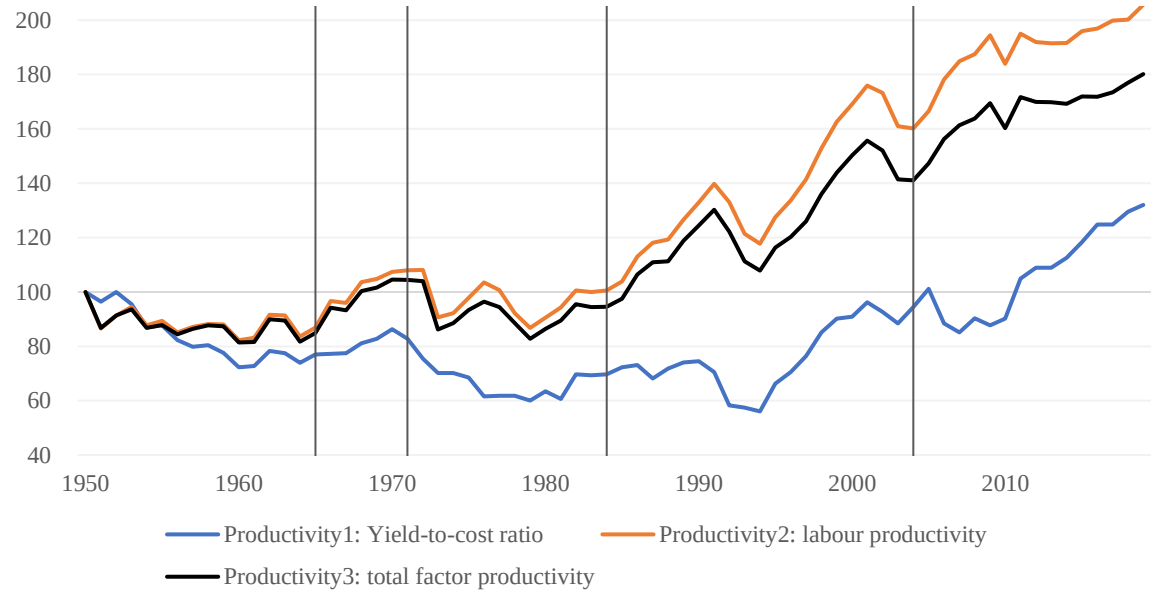
Computerisation of the PAYE system in 1984 coincides with a marked change in trend of all three productivity measures, with productivity growth increasing from this point onwards. Digitalisation

began with this change in 1984 but continued through the subsequent decades, and thus is consistent with a growth effect rather than (only) a level effect. While we capture capital inputs in our preferred productivity measure, computerisation also enables increased efficiency in the use of other inputs. Figure 4 shows that labour productivity growth outpaces total factor productivity growth from 1984 onwards, consistent with capital deepening and a substitution from labour to other inputs over this period.

Productivity continued to grow fairly consistently up to 2019, with continued digitalisation of the tax system, as outlined in section 2 (see also section 7). Our tax collection productivity measures exhibit a slight slowdown in productivity growth after the 2007/08 financial crisis and economic downturn (see Table 3), as is apparent in data for the economy as a whole and the market sector – a trend known as the "productivity puzzle". This is somewhat at odds with trends in public service productivity estimated by the ONS, which show faster growth after 2007 than before.

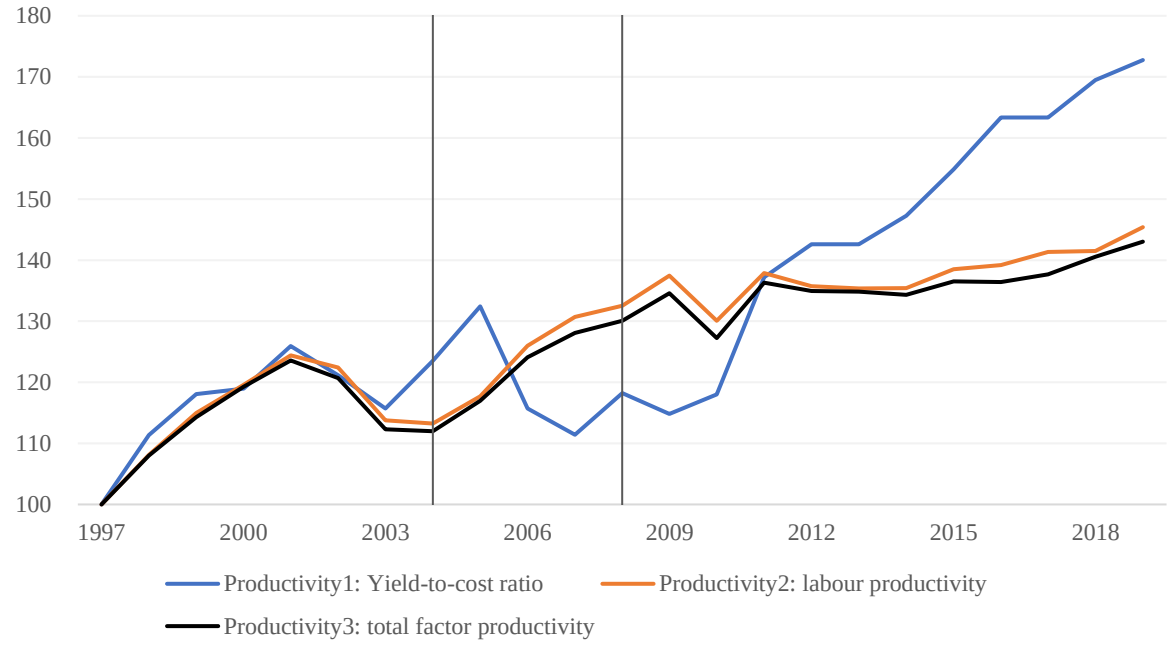
Budget cuts at HMRC from around 2008 onwards, especially in 2011 consistent with the 'austerity' programme, reduced 'conventional' inputs considerably. Our composite input volume index, covering labour, capital and intermediates, peaks in 2005 before falling slightly in 2006 and 2007, and then more sharply between 2008 and 2015, whereafter it flattens off. However, tax receipts changed relatively little over this period, implying an increase in productivity (albeit a modest increase relative to the preceding couple of decades). An increase in productivity over this period is also consistent with a narrowing of the 'tax gap' – the estimated difference between tax receipts and the theoretical tax liabilities, as estimated by HMRC and published annually in their "tax gap" report. This reflects an increase in tax receipts (output) relative to the tax base (an input).

Figure 4 – Tax collection productivity measures, index 1950 = 100



Source: Various sources (see text), authors' calculations.
Notes: Vertical lines correspond to (from left to right): 1965 – Corporation Tax created, 1971 – Decimalisation, 1984 – PAYE computerised, 2004 – HMRC formed.

Figure 5 – Tax collection productivity measures, index 1997 = 100



Source: Various sources (see text), authors' calculations.
Notes: Vertical lines correspond to (from left to right): 2004 – HMRC formed, 2008 – global financial crisis and economic downturn.

Table 3 summarises the average annual growth rates of the various measures of productivity of tax collection, and some comparators, over different sub-periods from 1850 to 2019. Tables A1, A2 and A3 in the Appendix do similarly for the constituent parts (inputs and outputs) of each productivity measure over the same periods.

Tax collection productivity is estimated to grow faster than the productivity of the economy as a whole between 1997 and 2019, due to faster growth over both the pre- and post-2007 period. Since 1950, whole economy TFP is estimated to have grown somewhat faster than tax collection TFP, albeit with a very different pattern over time – while whole economy TFP growth appears fastest between 1950 and 1980, slowing somewhat thereafter, and then slowing considerably after 2007, tax collection TFP is estimated to have fallen between 1950 and 1980, before accelerating sharply. The technological progress which drove whole economy productivity in the 1800s and early 1900s – electrification, production lines, modern business management techniques, etc. – likely applied mostly to manufacturing and allied industries, and less to public administration activities such as tax administration.

The tax collection productivity measures reflect the activity of just one organisation, and so the timing of individual investments and innovations are very important to the time series. This 'micro' narrative is in contrast to the macro trends of the economy as a whole, which is a grand average of all the organisations, industries and regions of the economy, each of which will experience investments, innovations and shocks at different times. It is thus no wonder that the data for the economy as a whole is smoother and grows more continuously than for tax collection alone. That they should be relatively similar over the long-term is, however, reassuring.

Table 3 – Summary of productivity measures, average annual growth rates by sub-period

Variable	1850-2019	1850-1950	1950-2019	1950-1980	1980-2019	1980-1997	1997-2019	1997-2007	2007-2019
Productivity 1 – Yield-to-cost ratio Input = Expenditure (CP) Output = Total tax receipts (CP)	**1.0**	**1.4**	**0.4**	**-1.5**	**1.9**	**1.1**	**2.5**	**1.1**	**3.7**
Productivity 2 – Labour productivity Input = Labour (vol), tax base (vol) Output = Total direct tax receipts (vol)	**0.6**	**0.4**	**1.1**	**-0.3**	**2.1**	**2.7**	**1.7**	**2.7**	**0.9**
Productivity 3 – Total-factor productivity Input = Labour, capital, intermediates, tax base (vol) Output = Total direct tax receipts (vol)	**0.6**	**0.4**	**0.9**	**-0.5**	**1.9**	**2.2**	**1.6**	**2.5**	**0.9**
Memo: Whole economy labour productivity	*1.3*	*0.9*	*1.9*	*2.6*	*1.4*	*1.9*	*1.0*	*1.7*	*0.4*
Memo: Whole economy total factor productivity	*1.1*	1.0	*1.4*	*1.8*	*1.0*	*1.3*	*0.8*	*1.7*	*0.1*
Memo: Public service productivity (with quality adjustments)	*NA*	*NA*	*NA*	*NA*	*NA*	*NA*	*0.2*	*0.0*	*0.4*

Source: Various sources (see text), ONS, authors' calculations.
Notes: "CP" is short for current prices, i.e. not adjusted for inflation. "Vol" is short for volumes, i.e. adjusted for inflation. Whole economy labour productivity is measured by output per worker for consistency with our labour volume measure.

7. Examples of productivity growth in tax collection

In order to support the estimates of productivity growth in tax collection presented in section 6, it is useful to consider examples of innovation and business practices that might have supported this productivity growth. This section briefly reviews some evidence.[18]

Section 3 gave a brief history of UK tax collection, including some examples of innovations that supported productivity. The first major change to tax collection came with the development of the Pay As You Earn (PAYE) system in 1944. Previously taxes were collected of workers annually or biannually, which meant earners liable to tax had to anticipate their tax liabilities, save the necessary money, file the necessary paperwork, and pay their taxes once or twice a year. As the tax burden increased, many more people than before WWII were liable to tax for the first time. IR launched an "Income Tax Quiz for Wage Earners" – a two-penny booklet explaining Income Tax in simple words – in 1942, to support understanding of the tax code (Griffith, 1949). They also tried to simplify the tax forms "to bring their language nearer to the standards we now associate with 'Plain Words'" (Griffith, 1949, p.43).

[18] More recent annual reports from IR and HMRC, available online since annual report FYE 1998, provide much more detail on their organisational changes and innovations than reports in the past. Interested readers should consult these for further details. See also Carrera, Dunleavy and Bastow (2009) and Dunleavy and Carrera (2013).

PAYE allowed money to be withheld (or deducted at source) in anticipation of tax liabilities, reducing the need to budgeting, calculation and paperwork. If employment and incomes were stable throughout the year, the estimate could be very accurate. PAYE is still with us today in a similar fashion to when it was introduced, suggesting it was a good innovation. Indeed, Figure 3 suggests a level-shift in productivity from 1944, consistent with the introduction of PAYE.

Just as WWII motivated the creation of PAYE, so WWI motivated innovation too. Griffith (1949) describes the increase in workload at Inland Revenue in WWI given the widening of the tax base, and what was needed to keep tax revenue coming in: "…And all the time, in all the Branches, the work grew more complicated and greater in bulk, involving much negotiation and an increasing amount of litigation… The effect of all this upon the Board's staff was immediate and far-reaching. It called for new standards of efficiency in knowledge of the law and its practice, and for a wider dissemination of that knowledge" (p.40).

This passage suggests the importance of the IR staff having the necessary knowledge to the law and tax schedules. Several IR reports describe the training delivered to staff its importance. For instance, the FYE 1970 annual report says: "…the specialised nature of much of the Department's work calls for the devotion of substantial resources to training new staff and to management training for supervisory and senior staff" (pp.9-10). It reports also that 4,700 staff were "under training in the Chief Inspector's Branch" in FYE 1970, which makes up nearly 10% of the staff in that part of IR at the time. IR annual report FYE 1990 reports that over 200,000 days of training were undertaken in FYE 1990 (p.36), equivalent to over 3 days in the year for every person employed at IR at the time.

There is evidence that Inland Revenue also used staff to generate ideas for internal reform and improvements. The IR annual report FYE 1982 describes a "staff suggestion scheme" with awards paid for the best suggestions. The highest award in that year was for "a suggested simplification of the register sheet for PAYE taxpayers with some Schedule D income" (p.8). This scheme was still operating in FYE 1991, with a reported increase in participation and £1.7m in savings achieved as a result (IR annual report FYE 1991, p.35).

Inland Revenue also seem to have been very conscious of the need for reform and organisational change. They conducted a "comprehensive review of the organisation of the Department" in 1949 (IR report FYE 1949). Johnston (1965) describes experimentation in work design: "The question of the optimum size of a local office from the standpoint of management has been the subject of much thought in the Revenue as in other departments with a decentralised organisation. The old tax office had a staff of about twenty. In the thirties experiments were made with much larger offices, for example, one tax office for the whole of Bristol or of Leicester. These large offices were not regarded as a success. Experience suggested that a tax office of forty to fifty staff was an efficient unit" (p.49). IR annual report FYE 1990 also reports of "experiments" in amalgamating the work assessing and collecting tax (p.42).

Productivity growth across the period 1950-1980 is estimated to be negative (Table 3). One reason could be an increase in the size and complexity of the tax system over that time, which was largely met by greater labour input rather than productivity gains. The IR annual report for FYE 1969 says: "The sustained increase in staff since 1946 reflects the constant growth in the number of taxpayers and in the complexity of the legislation. The rise from 1946 to 1948 was to cope with work on PAYE which proved heavier than expected…. The main causes of growth since 1965 have been the additional work on capital gains tax and corporation tax [and] the early stages of the introduction of automation" (p.15). As well as decimalisation in 1971 (see section 6), another reason for weak productivity growth in the 1970s (see Figure 4) might be shortages of staff with the necessary training and experience. The Inland Revenue annual report for FYE 1970 says: "many branches continue to carry a large proportion of inexperienced offices and this has again had an impact on the quality of

work" (p.8). IR annual report FYE 1971 also reports "increasing difficulties in obtaining payment from some taxpayers" (p.15).

Mechanisation began as early as 1958 when "an electronic computer" was installed for use in the production of statistics (IR annual report FYE 1958, p.12). In 1965 "three major projects were in hand" regarding automation and the use of "automatic data processing" (IR annual report FYE 1965, p.10). A centre at East Kilbride to centralise and automate PAYE work of Scottish taxpayers, using a computer, was opened in 1970 (IR annual report FYE 1970). This was the start of the full digitisation and centralisation of PAYE which was launched in 1984.

Johnston (1965) describes in delightful detail the savings available from automation: "When this [the production of PAYE Tables] was first done in 1944 it took over sixty people about six months to calculate over ten million figures for about 8,000 pages of tables; over forty different printers were engaged to produce the tables. Subsequently the volume of the tables was greatly reduced (at the cost to employers of some additional arithmetic); even so, in 1953 the calculations and the preparation of the copy for the printer still took several people about six weeks with the use of punched cards and a tabulating machine. Now the job is done by computer in a matter of hours" (p.79).

An efficiency programme began around 1980. The IR annual report FYE 1981 reports that "nearly 300 posts were saved… as a result of the 1980 staffing inspection recommendations" (p.21). Other reviews, including the 1982 Rayner Review of Government Statistical Services also saved "over 900 staff" (p.21). The Conservation government made plans to reduce the size of the civil service in 1983, requiring IR to reduce their staff by about 7,000 by April 1988 relative to April 1984 (IR annual report FYE 1984). At that time, they planned to reduce staffing through efficiency improvements, including computerisation of PAYE (expecting to save around 4,000 jobs), computerisation of accounting work, and simplification of the tax code. They also planned to increase headcount in some areas to meet increased workloads, and for work on 'the black economy'. In IR annual report FYE 1991 they report that between 1981 and 1991, staff savings were made from "administrative and efficiency measures (6,490), computerisation (6,260) and legislative changes (3,200), while increased workload added 8,330, for a net reduction of 7,620 (p.36).

Computer terminals were deployed rapidly through the late 1980s. The IR annual report FYE 1989 reports an increase in the use of computer terminal from near zero in FYE 1985 to nearly 40,000 in FYE 1989 (p.22), covering well over half of all staff at IR at that point. This is a period that coincides with a sharp pickup in productivity growth (see Figure 4). By 1990 there is already reference to replacing older terminals with new equipment (IR annual report FYE 1990, p.25).

Digitalisation continued in the 1990s and 2000s: Electronic Data Interface was introduced in 1998, enabling data to be transferred from larger employers to HMRC; online filing for Corporation Tax was introduced in 2003, and universal online filing was rolled out over 2004-05, with paper forms phased out between 2008 and 2011. During this time, HMRC was formed in 2004, and widespread government spending cuts (so-called "austerity") were implemented in the 2010s.

8. Conclusion

We have developed a framework to define and measure the productivity of tax collection, an area that has received little attention in the past, and is currently measured entirely by cost in official economic statistics. Using archival data and some creative data work, we have estimated three productivity measures for UK tax collection from 1850 to 2019.

Our preferred estimate suggests annual average productivity growth of around 0.6% between 1850 and 2019 of around 0.9% since around the end of WWII, and of around 2% in recent decades, with some possible slowdown since the global financial crisis. A simpler measure suggests even faster growth over the long run. These measures tell a consistent story of significant increases in

productivity over time, with productivity growth since the mid-1980s especially apparent, likely reflecting a continual digitalisation process that started with the computerisation of PAYE in 1984. The ups and downs of productivity growth, albeit general growth over a long period of time, is consistent with a history of innovation, experimentation and (more recently) digitalisation of tax collection.

There are a number of limitations and notes of caution regarding these measures. First, our preferred measure (tax base-adjusted total factor productivity) accounts only for Income Tax and Corporation Tax (and predecessors). As such, our productivity measures do not reflect all of the activity of Inland Revenue, and much less the activity of HMRC, which is responsible for many other taxes and activities. Second, the data we have used come with considerable uncertainty, especially historically, and should be interpreted as only indicative of trends over such long periods of time. Third, we have made a number of assumptions of the production function, most notably in the parameters used in our composite inputs measure. The weighting of 'conventional' inputs (or labour, capital and intermediates) with the tax base is based on a number of assumptions and limited data, and is therefore uncertain. That said, the estimates of productivity growth since around 1950 are fairly robust to different weighting specifications, with estimates before this point more sensitive (see Appendix A).

Fourth, the productivity measures should all be interpreted as "social productivity" measures, since they will reflect the degree of willingness of the taxpayers to pay their taxes. During the World Wars, citizens were more willing to pay their taxes (even at higher rates) in order to contribute to the war effort – this would enable Inland Revenue to collect more in tax receipts without additional input, and thus show up as an increase in tax collection productivity in our measures. This is not necessarily a problem, but should be borne in mind when interpreting the figures.

Fifth, we do not adjust for changes in service quality over time, as advocated in the Atkinson Review (Atkinson, 2005), and as implemented by ONS in their public service productivity statistics (e.g. ONS, 2022). Quality measures that could be considered include: citizen satisfaction with the tax system as a whole; customer satisfaction with direct interaction with the tax agency (e.g. during audits or when seeking guidance), including the ease of contacting the tax agency; the degree of fraud and/or error; the progressiveness of the tax system; and so forth. Changes in quality could be accounted for in the output measure in addition to changes in the 'quantity' of output, as measured by tax receipts in this paper.

Sixth, we do not make any account of the costs borne by taxpayers to pay their taxes. In the US, there is a large industry in tax services to individuals, since most taxpayers are liable to calculate their own tax liabilities. By contrast, in the UK only the self-employed and those with high incomes or complex tax arrangements need to complete tax returns. The UK thus has a lower burden on its taxpayers than the US, and a higher burden on the tax agency itself, which may result in a lower level of tax collection productivity in the UK than the US if the costs to citizens are not accounted for. These costs may also have changed over time in the UK, even if they are relatively small today.

References

Atkinson, T. (2005). "Atkinson Review: Final report Measurement of Government Output and Productivity for the National Accounts". *Palgrave Macmillan, Basingstoke; Her Majesty's Stationery Office (HMSO).*

Besley, T. and Persson, T. (2009). "The Origins of State Capacity: Property Rights, Taxation, and Politics". *American Economic Review, 99, 1218–1244.*

Carrera, L., Dunleavy, P. and Bastow, S. (2009). "Understanding productivity trends in UK tax collection". *LSE Public Policy Group Working Paper, July 2009.*

Chapman, A. and Knight, R. (1953). *Wages and Salaries in the United Kindon, 1920-1938.* Cambridge University Press.

Clark, T. and Dilnot, A. (2002). "Long-term trends in British Taxation and Spending". *The Institute for Fiscal Studies Briefing Note No. 25.*

Daunton, M. (2001). "Trusting Leviathan: The Politics of Taxation in Britain, 1799-1914". *Cambridge University Press; Cambridge, UK.*

Daunton, M. (2002). "Just Taxes: The Politics of Taxation in Britain, 1914–1979". *Cambridge University Press; Cambridge, UK.*

De Neve, J. E., Imbert, C., Spinnewijn, J., Tsankova, T., and Luts, M. (2021). "How to improve tax compliance? Evidence from population-wide experiments in Belgium." *Journal of Political Economy, 129*(5), 1425-1463.

Dincecco, M. and Katz, G. (2016). "State Capacity and Long-run Economic Performance". *The Economic Journal, 126, 189–218.*

Dohrmann, T. and Pinshaw, G. (2009). "The Road to Improved Compliance: A McKinsey benchmarking study of tax administrations – 2008-2009". *McKinsey report.* Available: https://www.mckinsey.com/~/media/mckinsey/dotcom/client_service/public%20sector/pdfs/road_improved_compliance.ashx

Dunleavy, P. and Carrera, L. (2013). "Growing the Productivity of Government Services". *Edward Elgar Publishing Limited; Cheltenham, UK.*

Feinstein, C. H. (1972). *National income, output and expenditure of the United Kingdom 1855-1965*, Cambridge: Cambridge University Press.

Foxton, F., Grice, J., Heys, R. and Lewis, J. (2019). "The measurement of public goods: lessons from 10 years of Atkinson in the United Kingdom". *Eurona, 2019/2.*

Griffith, L. W. (1949). *A Hundred Years, the Board of Inland Revenue, 1849-1949.* Somerset House: London.

Hoppit, J. (2021). "The Dreadful Monster and its Poor Relations: Taxing, Spending and the United Kingdom, 1707-2021". *Allen Lane; London, UK.*

Jensen, A. and Lagakos, D. (2019). "Productivity in U.S. Tax Collection, 1929 to 2017". *Unpublished working paper, proceedings of Brookings Institution conference on productivity, May 2019.*

Johnston, A. (1965). *The Inland Revenue*. The New Whitehall Series No. 13. George Allen & Unwin Ltd: London.

Kay, J. A. and King, M. A. (1990). *The British Tax System (Fifth Edition).* Oxford University Press.

Malanga, F. (1986). “The Relationship Between IRS Enforcement and Tax Yield”. *National Tax Journal*, 39(3), 333–337.

Mitchell, B. (2011). “British Historical Statistics”. Cambridge University Press 1988; Cambridge, UK.

NSCASE. (2023). The Quality Adjustment of Public Services. National Statistician’s Committee for Advice on Standards for Economic Statistics; NSCASE(23)12.

OECD. (2022). “Tax Administration 2022: Comparative Information on OECD and other Advanced and Emerging Economies”. *OECD Publishing, Paris.* https://doi.org/10.1787/1e797131-en.

ONS. (2022). “Sources and methods for public service productivity estimates”. *ONS article.* Available: https://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity/methodologies/sourcesandmethodsforpublicserviceproductivityestimates

ONS. (2023). “Public service productivity: total, UK, 2020”. *ONS article.* Available: https://www.ons.gov.uk/economy/economicoutputandproductivity/publicservicesproductivity/articles/publicservicesproductivityestimatestotalpublicservices/latest

Tauchen, H. V., Witte, A. D. and Beron, K. J. (1993). “Tax Compliance: An Investigation Using Individual Taxpayer Compliance Measurement Program (TCMP) data”. *Journal of Quantitative Criminology,* 9(2), 177–202.

Thomas, R. and Dimsdale, N. (2017). “A millennium of macroeconomic data.” Bank of England research dataset. Available: https://www.bankofengland.co.uk/statistics/researchdatasets

Williamson, J. G. (1982). “The Structure of Pay in Britain, 1701-1911”, *Research in Economic History, 7.*

Appendix A – Additional results

Table A1 – Summary of output measure and constituent parts, average annual growth rates

Variable	**1850-2019**	**1850-1950**	**1950-2019**	**1950-1980**	**1980-2019**	**1980-1997**	**1997-2019**	**1997-2007**	**2007-2019**
Income Tax receipts [including surtax] (CP)	6.3	5.7	7.2	9.0	5.9	7.4	4.7	7.9	2.2
Corporation Tax receipts [including other profit taxes] (CP)	NA	NA	7.9	9.6	6.6	11.1	3.3	4.9	1.9
Total direct tax receipts (CP)	6.5	5.9	7.4	9.1	6.0	8.2	4.4	7.1	2.1
Total direct tax receipts (vol) [deflated by GDP deflator]	**3.5**	**4.5**	**2.0**	**1.4**	**2.4**	**2.5**	**2.3**	**4.8**	**0.3**

Table A2 – Summary of inputs measures and constituent parts, average annual growth rates

Variable	**1850-2019**	**1850-1950**	**1950-2019**	**1950-1980**	**1980-2019**	**1980-1997**	**1997-2019**	**1997-2007**	**2007-2019**
Total expenditure (CP)	4.9	3.4	7.0	11.0	4.1	7.6	1.5	4.3	-0.9
Labour (vol)	1.5	2.7	-0.1	1.2	-1.1	-1.2	-1.0	1.7	-3.2
Non-labour (intermediates and capital combined) (vol)	NA	1.8	NA	3.9	NA	NA	NA	NA	NA
Intermediate inputs (vol)	NA	NA	NA	NA	1.1	3.4	-0.6	3.6	-4.0
Capital (vol)	NA	NA	NA	NA	1.0	1.2	0.9	1.2	0.6
Total expenditure inputs (vol)	**1.6**	**2.4**	**0.5**	**1.7**	**-0.4**	**0.1**	**-0.8**	**2.4**	**-3.3**
Total direct tax base (CP)	6.4	6.2	6.7	9.8	4.4	5.9	3.3	4.4	2.4
Total direct tax base (vol) [deflated by GDP deflator]	**3.4**	**4.8**	**1.4**	**2.0**	**0.9**	**0.3**	**1.3**	**2.2**	**0.5**

Table A3 – Summary of tax-base adjusted labour productivity measure and constituent parts, average annual growth rates

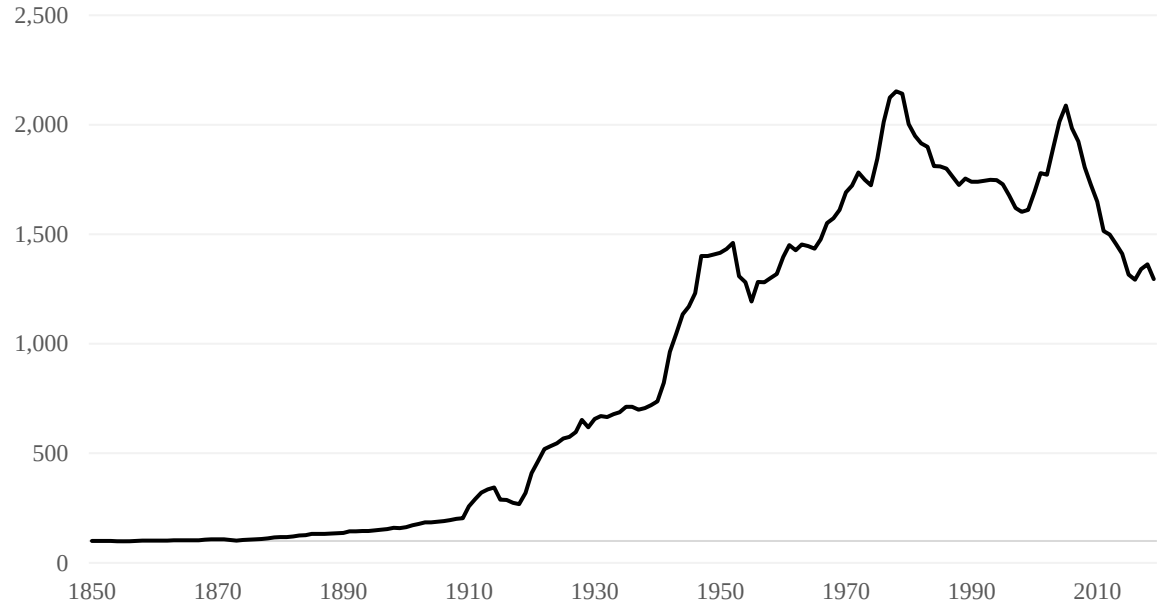
Variable	1850-2019	1850-1950	1950-2019	1950-1980	1980-2019	1980-1997	1997-2019	1997-2007	2007-2019
Total direct tax receipts (vol) [deflated by GDP deflator]	**3.5**	**4.5**	**2.0**	**1.4**	**2.4**	**2.5**	**2.3**	**4.8**	**0.3**
Labour (vol)	1.5	2.7	-0.1	1.2	-1.1	-1.2	-1.0	1.7	-3.2
Total direct tax base (vol) [deflated by GDP deflator]	3.4	4.8	1.4	2.0	0.9	0.3	1.3	2.2	0.5
Combined inputs: labour and tax base (vol)	**2.8**	**4.1**	**0.9**	**1.8**	**0.3**	**-0.1**	**0.6**	**2.0**	**-0.6**
Productivity 2 – Labour productivity Input = Labour (vol), tax base (vol) Output = Total direct tax receipts (vol)	**0.6**	**0.4**	**1.1**	**-0.3**	**2.1**	**2.7**	**1.7**	**2.7**	**0.9**

Table A4 – Summary of base-adjusted TFP measure and constituent parts, average annual growth rates

Variable	1850-2019	1850-1950	1950-2019	1950-1980	1980-2019	1980-1997	1997-2019	1997-2007	2007-2019
Total direct tax receipts (vol) [deflated by GDP deflator]	**3.5**	**4.5**	**2.0**	**1.4**	**2.4**	**2.5**	**2.3**	**4.8**	**0.3**
Total expenditure inputs (vol)	1.6	2.4	0.5	1.7	-0.4	0.1	-0.8	2.4	-3.3
Total direct tax base (vol) [deflated by GDP deflator]	3.4	4.8	1.4	2.0	0.9	0.3	1.3	2.2	0.5
Combined inputs: labour, capital, intermediates and tax base (vol)	**2.8**	**4.0**	**1.1**	**1.9**	**0.5**	**0.3**	**0.7**	**2.3**	**-0.6**
Productivity 3 – Total-factor productivity Input = Labour, capital, intermediates, tax base (vol) Output = Total direct tax receipts (vol)	**0.6**	**0.4**	**0.9**	**-0.5**	**1.9**	**2.2**	**1.6**	**2.5**	**0.9**

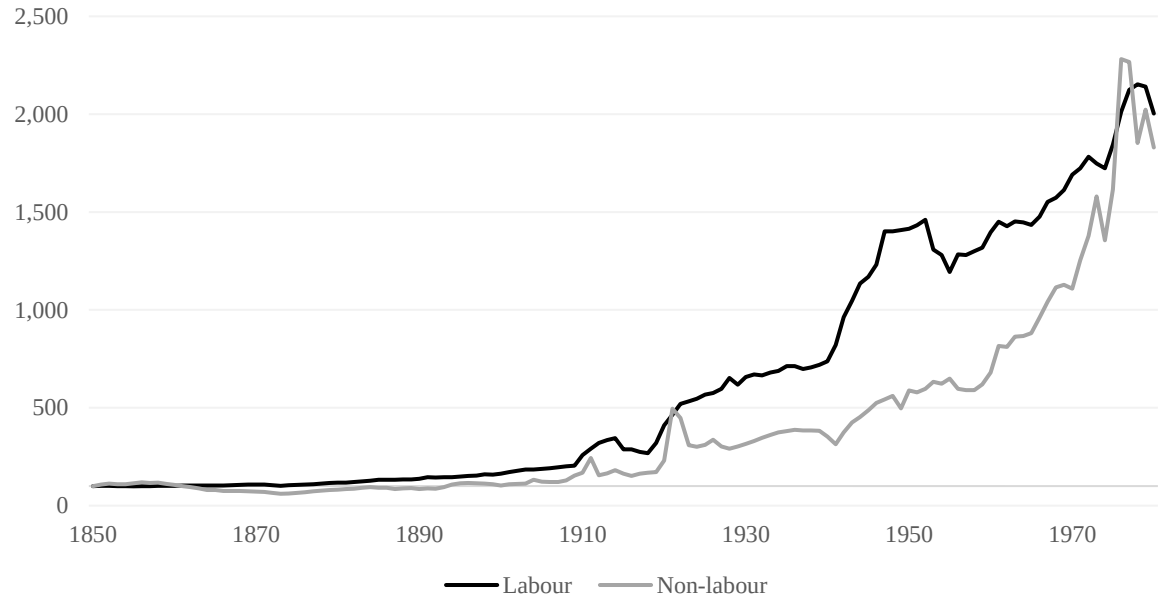
Volume indices of 'conventional' inputs

Figure A1 – Labour inputs, volume index, index 1850 = 100, FYE 1850 to 2019



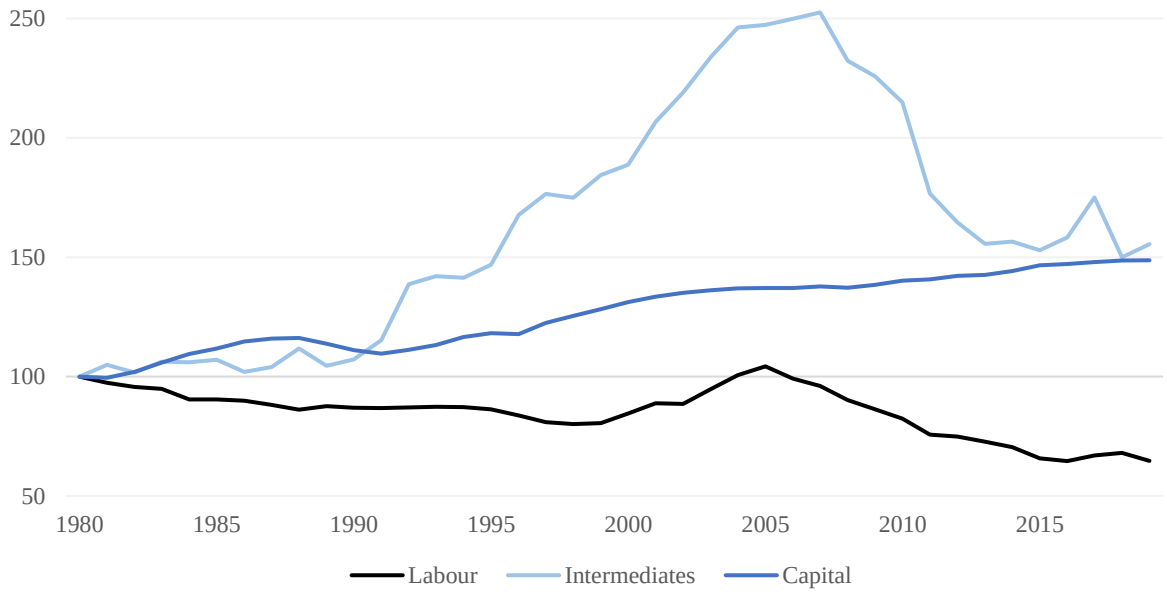
Source: Inland Revenue, HMRC, authors' calculations.
Note: This is the same data as presented in Figure 1, but expressed as an index relative to 1850 = 100.

Figure A2 – Non-labour and labour inputs, volume index, index 1850 = 100, FYE 1850 to 1980



Source: Inland Revenue, HMRC, authors' calculations.
Note: The "labour" series is the same as presented in Figure A1, but shown here for comparison against non-labour inputs. "Non-labour" is assumed to cover intermediate and capital inputs, but are not separately identified over this period – see section 4 for discussion.

Figure A3 – Labour, capital and intermediate inputs, volume index, index 1980 = 100, FYE 1980 to 1980



Source: Inland Revenue, HMRC, authors' calculations.
Note: The "labour" series is the same as presented in Figures A1 and A2, but re-indexed and shown here for comparison against intermediate and capital inputs.

Sensitivity of results to weighting of conventional inputs bundle and tax base

In section 4.2 we described how we weight together the bundle of 'conventional' inputs (labour, capital, intermediates) with the tax base in our third productivity measure, which we call tax base-adjusted total factor productivity. Our central estimate is a weight on the conventional inputs bundle (π) of 0.3, and by constant returns to scale, a weight on the tax base (μ) of 0.7. We argue that a plausible range for the weight on the conventional inputs bundle is 0.2 to 0.5, with a corresponding range for the tax bae between 0.8 and 0.5.

Table A5 shows the average annual productivity growth rates of our tax base-adjusted total factor productivity measure under different assumptions about these weights, π and μ. In each case, the same weight is used throughout the time series. The second data column, $\pi = 0.3, \mu = 0.7$, corresponds to our central estimate. Giving greater weight to the conventional inputs bundle (a greater value of π) results in larger estimates of productivity growth over the long run. This is because the conventional inputs bundle grows slower than the tax base, so giving it greater weight means the aggregate input index grows more slowly, which for given output growth means faster productivity growth. However, in certain sub-periods, especially recently, this makes less difference. Specifically, if the tax base and conventional inputs bundle grow at a similar pace, as they did between 1950 and 2007 (see Table A4) then different weighting will make little difference.

We can also relax the assumption of constant returns to scale – that is, that π and μ do not have to sum to 1. Based on our central estimate of $\pi = 0.3$, Table A5 also shows a scenario with diminishing returns to scale ($\pi = 0.3$, $\mu = 0.6$) and increasing returns to scale ($\pi = 0.3$, $\mu = 0.8$), where π and μ sum to less than 1 and more than 1 respectively. Relative to our central estimate (which assumes constant returns to scale), this makes a considerable difference in the earlier periods, but much less so in more recent periods.

Table A5 – Sensitivity of tax base-adjusted total factor productivity measure with respect to assumed production function weights, annual average growth rates, 1850 to 2019 and sub-periods

Weighting	$\pi = 0.2$ $\mu = 0.8$	$\pi = 0.3$ $\mu = 0.7$	$\pi = 0.4$ $\mu = 0.6$	$\pi = 0.5$ $\mu = 0.5$	$\pi = 0.3$ $\mu = 0.6$	$\pi = 0.3$ $\mu = 0.8$
Returns to scale	Constant	Constant	Constant	Constant	Decreasing	Increasing
1850-2019	0.4%	0.6%	0.8%	1.0%	0.9%	0.3%
1850-1950	0.2%	0.4%	0.7%	0.9%	0.9%	0.0%
1950-2019	0.8%	0.9%	0.9%	1.0%	1.0%	0.7%
1950-1980	-0.5%	-0.5%	-0.4%	-0.4%	-0.3%	-0.7%
1980-2019	1.8%	1.9%	2.0%	2.2%	2.0%	1.8%
1980-1997	2.2%	2.2%	2.3%	2.3%	2.3%	2.2%
1997-2019	1.4%	1.6%	1.8%	2.1%	1.8%	1.5%
1997-2007	2.5%	2.5%	2.5%	2.5%	2.7%	2.3%
2007-2019	0.5%	0.9%	1.3%	1.7%	1.0%	0.9%

Source: Various sources (see text), ONS, authors' calculations.

Notes: First row denotes the weighting of the 'conventional' inputs bundle of labour, capital and intermediates (π), and of the tax base (μ). Second row shows whether the weights corresponding to constant returns to scale (the weights sum to 1), decreasing returns to scale (sum to less than 1) or increasing returns to scale (sum to more than 1). The central estimates in this paper are the second data column ($\pi = 0.3, \mu = 0.7$, constant returns to scale).

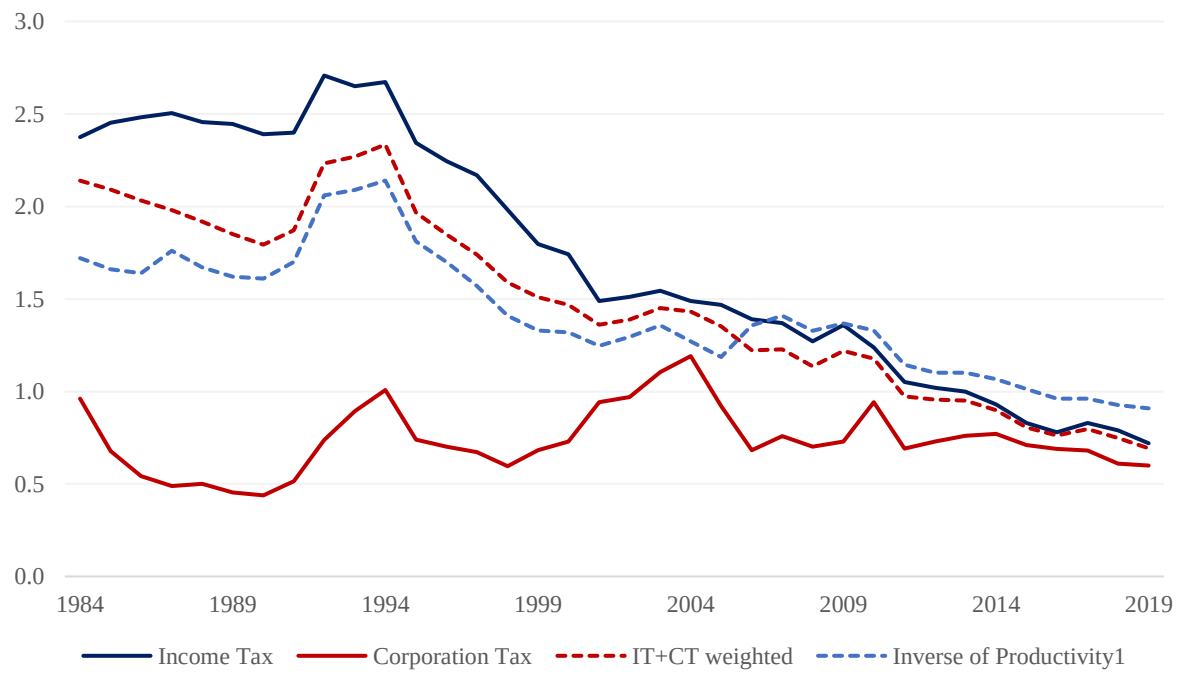
Cost-to-yield ratios for Income Tax and Corporation Tax

Since around 1980, the cost of collection of different types of tax have been reported separately (see sections 3 and 4.1 for more on what is meant by this). For greater consistency with our other productivity measures, which include only Income Tax and Corporation Tax in their output measures, we construct an aggregate cost of collection for only these taxes, weighting together their individual costs of collection by their shares of total combined tax revenue. This measure is thus largely immune to the organisational changes in 2000 and 2004 described in the paper.

We have compiled these figures from the Inland Revenue and HMRC annual reports, adjusting for minor methodological and definitional changes over time; for instance, figures for Income Tax prior to FYE 2011 were given including income from penalties, while later figures exclude that income, giving a higher (more expensive) cost-to-yield ratio – we splice the historic data onto the modern level.

Figure A4 shows the cost-to-yield ratio (cost of collection), the inverse of our first productivity measure (the yield-to-cost ratio) for Income Tax, Corporation Tax, the weighted average of Income Tax and Corporation Tax, and the data underlying our first productivity measure (which has a wider scope than just Income Tax and Corporation Tax). The weighted average measure shows a similar profile to (the inverse of) our first productivity measure. It is also notable that considerably more productivity growth (on this measure) is seen in the collection of Income Tax than in Corporation Tax over the period in Figure A4 (1984-2019).

Figure A4 – Cost-to-yield ratios for Income Tax, Corporation Tax, and aggregates, FYE 1984 to 2019



Source: Inland Revenue, HMRC, authors' calculations.
Notes: The light blue dashed line labelled "Inverse of Productivity1" is the inverse of (one divided by) our first productivity measure, the yield-to-cost ratio. This should be compared with the red dashed line labelled "IT+CT weighted" which is equal to the solid red and solid dark blue lines, weighted by their respective shares of total combined revenue.

Appendix B – Data appendix

Given the long time-horizon and constraint on both our time and the number of documents that can be requested on a single visit, we used Inland Revenue reports from years at regular intervals, such as every 5 years. Usually, the reports would contain 5-10 years' worth of backdata which we extracted, such that we had a complete time series. In a few cases, especially earlier in the period examined, we have some gaps. We often apply linear interpolation in these cases.

Cost of collection ($1/Productivity_1$)

Our first productivity measure ($Productivity_1$) is the inverse of the "cost-to-yield ratio" or "cost of collection" which is included in IR and HMRC reports for benchmark years from financial year ending (FYE) 1848 to 1884, and then every year thereafter. It is broadly the ratio of total revenues from tax collection to total cost of the tax collection activities, though the inclusions in both revenues and costs change somewhat over time. Revenues include tax receipts collected by other agencies/bodies on behalf of IR, and exclude receipts collected by IR on behalf of other bodies. Costs include payments to other government departments/bodies for services, and exclude costs borne by IR in providing services to other government departments.

We splice through a number of changes in measurement over time. Until FYE 1959, tax revenues are given in "gross" terms; that is, without deducting repayments, which might arise due to overpayment, errors, appeals, and so forth. From FYE 1960 onwards, tax revenues are in net terms. There is a discontinuity from FYE 2000 due to the integration of the Contributions Agency (responsible for the collection of National Insurance contributions) into IR, substantially reducing the aggregate cost of collection. For consistency with our other productivity measures which do not include National Insurance in their output measures, we maintain the 'level' of the series from before this change (i.e. excluding National Insurance from the scope), and splice more recent data onto that level. A further change in scope arises due to the creation of HMRC from IR and HMC Customs and Excise in 2004, which we again splice through, maintaining the previous level.

Since around 1980, the cost of collection of different types of tax have been reported separately. For greater consistency with our other productivity measures, which include only Income Tax and Corporation Tax in their output measures, we construct an aggregate cost of collection for only these taxes, weighting together their individual costs of collection by their shares total combined tax revenue. This measure is thus largely immune to the organisational changes in 2000 and 2004 described above.

Income Tax receipts (T_I)

Our measure of Income Tax receipts is as reported by Inland Revenue and HMRC in their annual reports. Data exist for almost all years since 1841, with the exception of a handful of years in the 1800s which we impute for.

We separately account for Surtax receipts (also known as Super-tax before 1929). The Surtax was an additional tax on top of Income Tax for high earners. It was introduced in 1910 and abolished in 1973. While it was a separate tax, it was a tax levied on individual income, so we combine this with Income Tax receipts henceforth.

Our modern measures cover specifically Income Tax, so this does not include employee's or employer's National Insurance contributions. It does, however, include Income Tax paid by the self-employed via self-assessment. Prior to the creation of Corporation Tax in 1965, businesses were also liable for Income Tax on their profits, which are included in our measure of Income Tax receipts.

The receipts are recorded against the year in which they are received, rather than the year to which the income relates. This may give rise to minor timing differences relative to our estimate of the tax base, and the 'conventional' inputs, though these should be small and balance out over time.

Corporation Tax receipts (T_C)

We group all taxes on business income and profits under the heading of "Corporation Tax", although in practice there are a wide range with different names and coverage – full details of the inclusions are in Appendix C. Before the creation of Corporation Tax in 1965, businesses were liable for Income Tax on their profits, and this is included in Income Tax receipts as described above. Additional taxes specifically on business profits are included under this heading and include: Excess profits duty (FYE 1915-1921), Corporation profits tax (FYE 1921-1924), Profits tax (FYE 1938-1965), and Excess profits tax (FYE 1940-1946). From 1965, our measure includes just Corporation Tax receipts, which has existed in broadly its current form since then.

As for Income Tax described above, Corporation Tax receipts (and other profit taxes) are recorded against the year in which they are received. In the case of the various short-lived profit taxes, this means that some residual revenues are recorded for years after the tax is abolished.

Staff numbers (L)

Our measure of the volume of labour input is a composite measure, using the best possible data over different periods of time. We use data on full-time equivalent (FTE) staff numbers (of varying methods) back to FYE 1939, and prior to that a measure of deflated expenditure on staff (described below). We also adjust as far as possible for changes in the scope of activities carried out by Inland Revenue and HMRC over time, in order to avoid these changes causing sharp movements in our labour input measure.

Between FYE 2000 and FYE 2019 we use a FTE measure that is the average over the financial year, treating non-full time staff according to their share of full-time hours. We exclude the number of contingent labour and contractors to avoid double counting with intermediate inputs (since we believe the costs of these workers are captured in our measure of intermediate expenditure). In the late 1990s we use a measure described as "staff usage", which appears to be similar to the FTE measure described above, but is different in overlapping years for unknown reasons.

With the formation of HMRC in 2004, the number of staff increases sharply. The scope of activities conducted by HMRC is broader than Inland Revenue (encompassing also the trade duties and levies that were previously the responsibility of HM Customs and Excise). For consistency in scope over time, and for consistency with our output measure, we splice on the growth in the HMRC growth rates to the level of Inland Revenue staff in the year of the merger. We also adjust for the creation of the UK Border Agency (in 2010) and the treatment of civil service central recruitment (moved to Cabinet Office in 2018).

Over the full time series, we remove (as far as possible) staff relating to the Valuation Office (VO), since they account for a fairly large and variable number of staff, and our output measure does not include output relating to valuation. This is straightforward from FYE 1997 onwards, but somewhat more difficult in earlier years. Additional detail included in some IR annual reports means we have regular benchmark years for non-VO staff numbers prior to 1997. We join between these benchmark years using data on the number of total Inland Revenue staff (i.e. including VO staff). Before 1950, valuation was the responsibility of local government, but those staff were incorporated into Inland Revenue from FYE 1951 onwards; we therefore treat pre-1950 staff numbers as non-VO, and adjust for the discrepancy in FYE 1951.

We have FTE data for FYE 1939 and FYE 1947, but not the intervening years. We had intended to use the pattern of deflated salary costs to inform the pattern in this gap. However, real salary costs increases only modestly, while staff numbers nearly double. This might suggest our wage deflator during WWII is inappropriate, but it is beyond the scope of this paper to create a better one. As such, we use the pattern of nominal staff costs (which approximately double, in line with the FTE data) to fill the gap.

We extend the series prior to FYE 1939 using deflated staff costs (i.e. staff costs adjusted for inflation). To this we collect a series of data on nominal staff costs (see more below) and construct a deflator to best represent wages of Inland Revenue staff (see section 5.3). We account for the separation of customs and excise from Inland Revenue in 1910 by holding constant nominal expenditure into the previous year, and then splicing on the trend for earlier years. This is complicated by the delay in the passing of the Finance Bill in 1910 which disrupted tax collection, and may have had some effects on labour volumes and staff costs in adjacent years.

Labour costs ($P_L L$)

We use data on labour costs to extend the labour volume series (described above), and to inform the weight on labour in the aggregation of the 'conventional' inputs bundle (with intermediate inputs and capital) in our third productivity measure. As for the data on staff numbers, we adjust for various changes in the scope of Inland Revenue and HMRC over time.

To extend the volume estimate, we need to match some data on labour costs with a suitable deflator. Our labour cost deflator is a wage deflator, constructed from various sources (as described in section 5.3). This is already challenging, but would be far more so if we tried to also account for non-wage labour costs pre-1939. Non-wage labour costs were, in any case, of little importance prior to the expansion of the welfare state in 1948 – recall that we have staff numbers for our volume estimate back to 1939. As such, we use a series on salary costs only (i.e. excluding non-wage labour costs such as pension costs) for this. Salary costs are reported in IR annual reports as "salaries and allowances" from 1850, and we have a near complete time series, with only a few gaps in the mid-1800s which we linearly interpolate for.

For weighting labour in the 'conventional' inputs bundle (with intermediate inputs and capital) we prefer a measure of total labour costs, including pension costs, employers' National Insurance contributions, and other non-wage labour costs. There are some empirical challenges here; for instance, pension obligations (referred to as "superannuation" in the annual reports) of all government departments were the responsibility of the Paymaster General's Office from 1973 to 1993, and are therefore difficult to identify for these years.

From FYE 1994 onwards we have a robust measure of total labour costs from IR and HMRC annual reports. For consistency with our measure of intermediate inputs (see below) we restrict staff costs as far as possible to the scope of Inland Revenue, and adjust for various mergers and changes in responsibilities. Prior to FYE 1994, we splice on a series of combined salary costs (as described previously) and pension obligations. The pension costs series is constructed based on data relating to "superannuation" and "other non-effective charges", for which we have a near complete time series from FYE 1850.

Expenditure on goods and services ($P_G G$)

Expenditure on goods and services (i.e. intermediate inputs) are only separately identifiable in the Inland Revenue records from the mid-1970s onward. From FYE 1976 onwards we have considerable detail on the value and composition of intermediate inputs, albeit with several changes in categories and reporting over time. As for staff numbers and labour costs, we adjust for changes in departmental scope where possible, including most notably for the formation of HMRC. We also have to content

with changes in reporting practices, with the level of detail provided reducing materially for a few years around 2000, before returning. We include the costs associated with Private Finance Initiatives (PFIs), Public Private Partnerships (PPPs), and rental and lease payments within intermediate costs, even where they may relate to the use of capital goods. Thus, our capital input measure (see below) relates only to capital owned by IR/HMRC.

The data on goods and services expenditure from FYE 1976 onwards cover a range of services including accommodation (office buildings), postal and telephone services, stationery and printing etc (provided by Her Majesty's Stationary Office for some of the time), travel and subsistence, legal services, consultancy, publicity[19], and other miscellaneous expenditures. There are numerous changes of categories between annual reports.

Prior to FYE 1976 we can identify total non-labour costs to a reasonable degree. These are reported as "other charges" and later "other costs" (i.e. other than labour costs) in IR annual reports. As we cannot separately identify it, we assume this also includes costs of capital investment and depreciation of existing capital. We thus interpret this measure as total non-labour inputs (comprising intermediate inputs and capital inputs), rather than just intermediate inputs.

For this earlier period, we also include the costs borne by other government departments for services provided to IR, reported as "charges met out of the Votes of other Departments" and later "costs borne by other Departments"[20] in IR annual reports. These relate to services provided by certain central government bodies to other government departments, such as postage and stationery services by HM Stationary Office, costs for accommodation and associated services (e.g. cleaning) by the Property Service Agency and earlier the Office for Works, and so on. A complication is that pensions costs (superannuation costs) were also provided by another department (the Paymaster General's Office) between 1973 and 1993, and so are included in that total – we need to exclude this to avoid double counting with the labour costs (described above). Fortunately, we have an estimate of superannuation costs from our estimate of total labour costs, and use that to adjust the "costs borne by other Departments" to leave just non-pension services.

Finally, we link the two sets of data – detailed intermediate inputs information from FYE 1976 onwards, and less detailed prior to this – with the more detailed data taking precedent.

Depreciation ($P_K K$)

The preferred measure of capital input for productivity analysis is capital services, which measures the flow of productive services from a capital stock. Constructing such as measure for tax collection would be difficult not worthwhile, given that capital is a relatively small input. As such, we follow the approach of ONS public service productivity statistics in using the deflated value of consumption of fixed capital (depreciation) as the capital input measure.

Depreciation (of capital assets) is recorded inconsistently over time, and we combine various measures, making some imputations and judgements to produce a time series that starts in FYE 1980. Depreciation is first reported in the FYE 2000 IR annual report, with data for FYE 1999 and 2000. It is reported annually thereafter. As usual, we splice through the creation of HMRC, maintaining the level reported by IR up to then. We use figures for the HMRC "core department" (excluding the Valuation Office) where available. The reported figures are revised materially in the FYE 2010 report

[19] Inland Revenue ran TV adverts earlier than this, including the infamous "Hector the Tax Inspector" adverts, and those featuring "Mrs Doyle" from the TV show Father Ted. These expenditures are likely in the "miscellaneous" category. Spending on "publicity" is only separately identified from 2003/04, with the creation of HMRC.

[20] Historically the budgets of government departments were set by annual 'Votes' in parliament, and so most costs were referred to by which 'Vote' they related to. This terminology is no longer used.

as new financial reporting standards were adopted. We include amortisation (depreciation of intangible assets) where reported.

Prior to FYE 1999, with no available data on depreciation, we make the assumption of a steady state capital stock in order to impute depreciation. If the capital stock is a constant volume, then inflows to the capital stock (investment) match outflows (depreciation). As such, depreciation can be estimated as equal to investment. In years where we have data on both investment and depreciation, the annual values are similar on average, although the investment series is naturally more variable than the depreciation services.

To avoid imparting unrealistic volatility into our capital input measure, we estimate a linear regression through the combined capital investment and depreciation data, and use those fitted values as the estimate for consumption of fixed capital. Specifically, we fit a linear regression line to the combined series of nominal capital investment (FYE 1980 – 1999) and depreciation (FYE 2000 to 2019) and use the fitted values as our estimate of nominal consumption of fixed capital. We then deflate it for the volume estimate.

Tax base (B)

We define the tax base as the theoretical amount of tax available to be legally collected given the prevailing tax rates and levels of labour and business income. We approximate the tax base for Income Tax and Corporation Tax by multiplying the levels of aggregate labour and capital income by the average of the main tax rates in each year. While crude, we believe this provides a reasonable approximation for the changes in the tax base over time. Specifically, we are ignoring complexities relating to allowances, variations in rates by income level, and special rates for special groups. For instance, for the Income Tax base in 2021/22, we average the basic rate of 20% and the higher rate of 40% (giving 30%), and multiply this by total wage and salary income, sourced from the National Accounts.

We repeat this exercise back to 1850 for Income Tax, to 1965 for Corporation Tax, to 1910 for the various other profit taxes levied before Corporation Tax (see Annex C), and between 1910 and 1983 for surtax (levied on labour income in addition to Income Tax).

We use National Accounts data on factor incomes in GDP. Data come from ONS and Mitchell's British Historical Statistics (which is sourced from a range of studies including Feinstein, 1972). Using National Accounts data throughout should provide relative consistency, although of course the measure is at best an approximation for the true value.

For the Income Tax base, we principally use the wages and salaries component of compensation of employees from the National Accounts. This is published by ONS back to 1948. Before 1948, we extend the series used the trend in compensation of employees (which includes non-labour income such as employers' pension contributions, although this is a small component, especially in the past), sourced from Mitchell (2011). Prior to 1965 businesses were also liable for Income Tax, so we include net operating surplus (see below) alongside wages and salaries. We also include the income of the self-employed (mixed income) throughout. In each case, data come from ONS back to 1948, extended with data from Mitchell for earlier periods.

For Surtax (included in Income Tax receipts in our calculations), we construct a composite rate as the average across all income bands. The Surtax was a graduated tax, applied at different rates to different bands of income above an eligibility level. It is far beyond the scope of this paper to model the distribution of incomes and apply the tax rates appropriately, so we take a simple average of the rates across different income bands. We source the rates from various sources including IR annual reports. We apply our composite rate to 10% of total wage and salary income, on the assumption that only

10% of total labour income was liable for Surtax, since it was levied only on people with high incomes.

For the Corporation Tax base (including predecessor taxes before 1965), we estimate net operating surplus (the closest concept to 'profits' in the National Accounts) using data on gross operating surplus and consumption of fixed capital (depreciation), where net operating surplus is gross operating surplus minus consumption of fixed capital. ONS publish data on gross operating surplus back to 1948, but on consumption of fixed capital only back to 1987. We have an estimate of consumption of fixed capital from Mitchell up to 1980. Between 1980 and 1987 we use the growth of gross operating surplus, before reverting to a net operating surplus concept. Prior to 1948 we use gross operating surplus data from Mitchell.

The tax rates for Income Tax and Corporation Tax are relatively easily sourced from government documents and various sources online from around the 1970s onwards, including from the Institute for Fiscal Studies (IFS) "tax lab". Income Tax rates since 1850 are obtained largely from Mitchell, with some corrections based on IR annual reports, and converted from shillings and pence to a percentage. Tax rates on the various profits taxes of the early 19th century are sourced from various (sometimes conflicting) sources including Parliamentary minutes, IR annual reports, Daunton (2001) and others – these are discussed in greater detailed in Annex C.

Deflators

For two of our productivity measures, we need deflators for either inputs (labour, goods and services, capital, tax base) or output (tax revenue). We source these principally from ONS and the Bank of England's millennium of macroeconomic data (Thomas and Dimsdale, 2017).

For the labour deflator, which we only need in years prior to FYE 1939, we use a combination of sources, obtained from the Bank of England's millennium of macroeconomic data dataset and Mitchell's British Historical Statistics. Much wage data from this time related to labourers and workers in agriculture and manufacturing, which would not be appropriate for Inland Revenue staff. We have therefore, where possible, used wage indices for relevant industries or occupations to better reflect the likely wage trends of Inland Revenue. Between 1920 and 1938 we construct a composite wage index based on the wages in industries "insurance, banking and finance", "local government services" and "professional services" (Mitchell, 2011, pp.171-172; sourced from Chapman and Knight, 1953). Between 1851 and 1911, we use decennial census data for the wages of occupation categories "government low-wage", "government high-wage" and "clerks (excluding government)" (Mithell, 2011, p.153; sourced from Williamson, 1982), linearly interpolating between years. In all other periods, in the absence of more relevant options, we use a whole economy average earnings index (Thomas and Dimsdale, 2017). Our composite wage index, comprising all of these sources and periods, exhibits less wage growth between 1850 and 1910 than an average index for all workers, but is similar to this average from 1910 onwards.

As a cross-check, we also constructed a wage index for years after 1938 and used it to deflate staff costs, which we compared with the staff numbers measure (our preferred labour input measure where available). The trend in deflated staff costs since 1938 matches that of staff numbers reasonably well. For the aggregate labour input measure, we use staff numbers back to FYE 1939, and backcast using the trend in deflated staff costs.

For the goods and services deflator, we use the GDP deflator. We experimented with a bespoke deflator that accounted for the product mix that we could identify from the breakdown in the Inland Revenue accounts. However, with many changes in the available breakdown of expenditure over time, and limited high-quality services deflators available for years prior to 1997, we could not be confident that this would be higher quality than the GDP deflator.

For the capital deflator we again use the GDP deflator. We experimented with the implied capital stocks deflator for the central government sector from the ONS capital stocks datasets, although this only extends to 1995, and would include to a large degree irrelevant assets such as roads.

Finally, for the deflator on output (tax revenues) we use the GDP deflator. There is no obvious choice here, since the output is money itself. A general price index is therefore appropriate, so we considered a consumer price index (such as the CPI or RPI historically), the GDP deflator, the implied GVA deflator, and the implied general government final consumption expenditure deflator. We chose the GDP deflator since it would not be significantly affected by the prices of imported goods or consumer trends, is available over a very long time period, and should be more consistent over time than CPI.

Appendix C – Summary of historic profits taxes and tax rates

All information correct to the best of our knowledge. Please note that there may be some inaccuracies.

Excess profits duty (EPD) (FYE 1915-1921)

Introduced in Budget September 1915, charged as a percentage on "war profits", calculated as profits above a pre-war standard. The precise rules to calculate the pre-war standard were complex, with many allowances for businesses of different ages. EPD came into effect in 1917, after some delay to details being agreed after initial announcement in 1915, and applied retrospectively to profits from FYE 1915 onwards. As a result, tax revenues started only in FYE 1917, relating to a tax base from FYE 1915 onwards. The rate was 50% for profits in FYE 1915, 60% for FYE 1916, 80% for FYE 1917 and 1918, 40% for FYE 1919, and 60% for FYE 1920 and 1921. Further payments were received between FYE 1922 and 1926 to settle existing cases. For more, see Billings and Oats (2014).

We model the tax base as the average level of Net Operating Surplus in adjacent years, relative to half the average of NOS in the three years before WWI (FYE 1912-1914), multiplied by the prevailing rate. This attempts to mirror, as far as possible, the application of the tax in practice. The reason for averaging adjacent years is to mimic the accounting treatment of taxes during a time when profits were changing quickly. The reason for halving the pre-war average rate of NOS is to reflect the many exemptions to the tax (see Billings and Oats, 2014).

Corporation profits tax (CPT) (FYE 1921-1924)

Introduced in Budget April 1920, charged at one shilling in the pound (5%). Ran alongside EPD in FYE 1921, with firms paying either EPD or CPT, whichever was higher. The difference is that EPD was charged only on "excess profits" (above pre-war levels) whereas CPT applied on total profits. After EPD was repealed, CPT continued. The rate was reduced to 2.5% for FYE 1923 and 1924, and abolished thereafter. Further small payments received until around the start of World War I to settle existing cases.

We model the tax base as the average level of Net Operating Surplus in adjacent years, multiplied by the prevailing rate. The reason for averaging adjacent years is to mimic the accounting treatment of taxes during a time when profits were changing quickly.

Profits tax (FYE 1938-1965)

Introduced in 1937 as the National Defence Contribution, and renamed "profits tax" after WWII. From FYE 1938 to 1946 it was charged at a flat rate of 5% on total profits. Ran alongside the Excess Profits Tax (EPT) between FYE 1940 and FYE 1946, with firms paying either of the two, whichever was higher (which was usually the EPT). From FYE 1947 the tax rate differed between "retained profits" (taxed at a lower rate) and "distributed profits" (taxed at a higher rate). These two rates were initially 10% and 25% respectively, with the "distributed profits" rate increasing to 30% in FYE 1950 and 50% in FYE 1951. Until this time, this tax could be offset against Income Tax, such that the effective rate was lower. From FYE 1952 onwards, the rates were reduced, but could no longer be offset against Income Tax. From FYE 1952 to 1955 the rates were 2.5% and 22.5%, changing to 2.5% and 27.5% in FYE 1956, and 3% and 30% in FYE 1957 and 1958. From FYE 1959 onwards it reverted to a flat rate on total profits, now charged at 15%. It was abolished in FYE 1965, when Corporation Tax was created.

We model the tax based as the average level of Net Operating Surplus in adjacent years, multiplied by the prevailing rate or average of rates. The reason for averaging adjacent years is as previously. We average rates on "retained profits" and "distributed profits" as roughly half of profits were retained and half distributed, such that effective rate was roughly the average of the two. We additionally half

this average for the years when the tax could be offset against Income Tax (FYE 1947 to 1957), since such offsetting roughly halved the effective rate.

Excess profits tax (EPT) (FYE 1940-1946)

Introduced in 1939 to provide revenue to fund WWII. Operated in a similar manner to the EPD levied during WWI, but at higher rates. It was charged at 60% of "excess profits" in FYE 1940, rising to 100% for FYE 1941 to 1945, and then 60% in FYE 1946, before being abolished. For the years in which the rate was 100%, a 20% refund was after the war. Ran alongside the Profits tax, with only the higher of the two payable.

We model the tax base in the same was as for the EPD, with the reference years for the pre-war average taken over FYE 1938-1939. We use an effective rate of 80% for FYE 1941 to 1945, rather than the actual rate of 100%, given the postwar refund of 20% and in order to keep the calculation straightforward.