

Methodological Improvements to National Accounts for Blue Book 2015: Exhaustiveness

Authors: Joshua Abramsky and Richard Wild, Office for National Statistics

Date: 19 May 2015

Executive summary

This article covers some improvements to estimates of Gross Value Added (GVA) and Gross Domestic Product (GDP) that will be incorporated when revised figures for the UK National Accounts, consistent with Blue Book 2015, are published in September 2015. ESA 1995 changes improve the measurement of GNI and ensure increased comparability of GNI across the European Union (EU). They will have an impact on the estimation of the UK's GNI, which is used to calculate the UK's contribution to the EU budget.

The changes cover improvements to the estimation of concealed income/activity and also under-coverage of small businesses, both of which form part of wider exhaustiveness adjustments that are applied to GVA/GDP to account for any production, income and expenditure that is not picked up in source data. These exhaustiveness adjustments are classified in Eurostat's "tabular approach" (see Eurostat, 2005) and for completeness these classifications are included here.

In the National Accounts:

- hidden economic activity generating concealed income forms part of output and intermediate consumption in the production approach to measuring GDP [GDP(P)], while concealed income itself forms part of compensation of employees (COE), mixed income (MI) and gross operating surplus (GOS) in the income approach [GDP(I)]. No specific adjustments are made to the expenditure approach to GDP prior to balancing, but coherence adjustments are made during balancing to bring this measure into alignment with GDP(P) and GDP(I). These adjustments to GDP(P) and GDP(I) constitute the vast majority of the adjustments to UK GDP for Eurostat exhaustiveness categories N1 (underground production, where the producer should have registered, but has not) and N6 (producer misreporting).
- unincorporated small business under-coverage adjustments form part of output and intermediate consumption in the production approach to ensure consistency with data already feeding into the income and expenditure approaches. Enterprises falling below the VAT and/or PAYE thresholds, by construction, are not recorded on the Inter-Departmental Business Register (IDBR), which provides the sample frame for the Annual Business Survey (ABS), which itself forms the largest source for the production approach to GDP. These adjustments to GDP(P) constitute most of the adjustments to UK GDP for Eurostat exhaustiveness category N3 (producer has no obligation to register administratively).

The changes that will be implemented in Blue Book 2015:

- replace the existing National Accounts concealed income/activity model with a new, more comprehensive model based, in part, on analysis by Her Majesty's Revenue and Customs (HMRC) of evaded corporation and income tax and
- update existing estimates of missing production by unincorporated businesses below the VAT and/or PAYE thresholds using updated administrative data provided by HMRC.

This article provides an overview of the key changes, associated methodology and impacted transactions, but it does not aim to provide a numerical assessment of the impacts. This can be found, along with the impact of other methodological improvements to ensure comparability in the measurement of GNI across EU Member States, in this article ['Impact of ESA 1995 Changes on Current Price Gross National Income Estimates, 2002 to 2010'](#) also being published today.

1 Introduction

Gross National Income (GNI) is an important statistic within the National Accounts, and it is used in the calculation of a Member State's contribution to the EU budget. Many users of official statistics will be more familiar with Gross Domestic Product (GDP). GNI or, as it was previously known, Gross National Product (GNP), describes the total primary income received by residents of a country and links the economic activity described by GDP with the destination of the income so generated.

Due to the administrative importance of the GNI statistic, the EU statistical office (Eurostat) carries out regular audits of the methods and data used to estimate GNI. In 2012, following a comprehensive audit of the methods used across EU countries, a number of areas for improvement were identified which all Member States needed to address. These improvements are known as 'reservations'. The UK National Accounts addressed a number of these reservations in [Blue Book 2014](#), and all the remaining outstanding UK reservations are being addressed in the Blue Book 2015 consistent update of the Quarterly National Accounts, that will be published in September 2015.

One of the improvements is to the estimation of exhaustiveness adjustments relating to concealed income/activity and under-coverage of unincorporated small businesses. This article provides a technical overview of the current and new methodology. It does not provide a numerical assessment of the impact; this can be found in the article ['Impact of ESA 1995 Changes on current price Gross National Income Estimates, 2002 to 2010'](#).

GNI Reservation (3) states:

A review of the methods to ensure exhaustiveness has to be carried out with a view to updating the benchmarks and/or modifying the methods where needed (applicable to years from 2002 to 2010).

Exhaustiveness adjustments cover a broad range of measures to supplement estimates derived directly from survey data and administrative records. Some are minor, such as estimates of cash tips in the income and production approaches, while others are more substantial, such as estimates of income-in-kind (e. g. fuel costs paid by employers) that form part of compensation of employees, or estimates of smuggling that impact on the production and expenditure approaches. While exhaustiveness adjustments may impact on all three approaches to GDP, in most cases the relevant activity is naturally recorded in at least one approach, to which the other approach(es) must be made consistent.

Two major elements of the UK National Accounts' exhaustiveness adjustments are the subject of Reservation (3): concealed income/activity and under-coverage of unincorporated small businesses.

In the National Accounts:

- hidden activity generating concealed income forms part of output and intermediate consumption in the production approach to GDP [GDP(P)], while concealed income itself forms part of compensation of employees (COE), gross operating surplus (GOS) and mixed income (MI) in the income approach [GDP(I)]. It is assumed that components of the expenditure approach [GDP(E)] are already exhaustive on a National Accounts basis in this respect. These adjustments to GDP(P) and GDP(I) constitute the vast majority of the adjustments to UK GDP for Eurostat exhaustiveness categories N1 (underground production, where the producer should have registered, but has not) and N6 (producer misreporting).

- small business under-coverage adjustments form part of output and intermediate consumption in the production approach to ensure consistency with data already feeding into the income and expenditure approaches. Enterprises falling below the VAT and/or PAYE thresholds, by definition, are not recorded on the Inter-Departmental Business Register, which provides the sample frame for the Annual Business Survey, which itself forms the largest feed into the production approach to GDP. These adjustments to GDP(P) constitute most of the adjustments to UK GDP for Eurostat exhaustiveness category N3 (producer has no obligation to register administratively).

The changes that will be implemented in Blue Book 2015:

- replace the existing National Accounts concealed income/activity model with a new, more comprehensive model based, in part, on 'tax gaps' analysis by HMRC of corporation and income tax not charged because of illegal mis- or non-reporting, whether deliberate or accidental, i.e. tax evasion. The new model was derived following a comprehensive review of the existing model, including an assessment of whether it could be updated or should be replaced. The relatively recent availability of tax gaps analysis by HMRC has allowed for more detailed estimates of missing factor incomes to be generated than was previously possible.
- update existing estimates of missing production by unincorporated businesses below the VAT and/or PAYE thresholds using updated administrative data provided by HMRC. The new data are very similar in construction to the data they replace, but are based on much more recent bespoke analysis and are on a genuine Standard Industrial Classification 2007 (SIC2007) basis, rather than having been converted from SIC2003 to SIC2007 via a mapping process.

2 National Accounts concepts

In the UK National Accounts, in compliance with our legal requirements under the European System of Accounts 2010 (ESA 2010) (Eurostat, 2012) and, before that, ESA 1995 (Eurostat, 2007), estimates of GDP via the production, income and expenditure approaches should be exhaustive, meaning that adjustments must be applied to ensure that any relevant economic activity that is not captured within source data is accounted for. Recent work has focused on two key elements of exhaustiveness, concealed income/activity and unincorporated small business under-coverage. The following sections explain each element in more detail.

2.1 Concealed income and activity

Measures of each approach to GDP must include estimates of "underground production" – legal activities that are not legally recorded in part or in full – as this activity falls within the production boundary (ESA 1995, para. 1.13h). Therefore, in the case of concealed income, adjustments must be made to compensation of employees (COE), gross operating surplus (GOS) and mixed income (MI) to transform the raw wages & salaries and profits data to the National Accounts concepts of COE, MI and GOS. In turn, for GDP(P), estimates must also be made of any associated output and intermediate consumption that is not recorded on the production approach, and, if necessary, final consumption missing from the expenditure approach. The use of mostly distinct sources to provide the raw data underpinning the National Accounts estimates produces differences in their respective coverage of aspects of exhaustiveness. As such, a 'one-size-fits-all' approach is invalid and individual specific adjustments to production, income and expenditure must be made to ensure their accuracy.

Consistency between the three measures must then be ensured using coherence adjustments.

For this aspect of exhaustiveness, while the estimates are generated using the Tax Gaps that are compiled by HMRC and therefore relate to taxable variables – loosely speaking, salaries and profits – that conceptually are included in GDP(I), it is the impact on the production approach which determines the overall impact on balanced GDP. This is because GDP(I) ‘...does not provide the same robust balancing exercise’ as GDP(P), within which elements of GOS may be estimated as residuals (ESA 2010, para 9.17). As noted, once all three approaches have been adjusted if and where appropriate for differences in concepts and coverage, subsequent coherence adjustments must be applied to GDP(I) and GDP(E) to align them with GDP(P).

2.2 Under-coverage of unincorporated small businesses

The production approach to GDP should cover all economic activity that falls within the production boundary (ESA 1995, paras. 3.07 – 3.09). Where source data from business surveys or other sources do not cover certain areas of production, for example by industry, sector or transaction, secondary adjustments must be made so that estimates of GDP are exhaustive. In this specific area, ESA 1995 requires that any activities not registered for tax purposes (ESA 1995 para. 3.08) must be included. In this case, the income data provided by HMRC to ONS to form estimates of GDP(I) is already exhaustive as it includes all recorded income, regardless of business size, while estimates of expenditure in GDP(E) are implicitly based on all income and are also exhaustive, so it is solely GDP(P) which requires conceptual adjustment in this respect.

3 Data sources and methodology

3.1 Current approach to concealed income and activity

The existing model used by ONS dates back to 1994. It is based on the difference between estimates of GDP(P) and GDP(I) (prior to the application of coherence adjustments) in 1994, which was assumed to be equal to tax evasion. The basis of the approach was that GDP(P) was considered to be more exhaustive, based on a much-improved IDBR and some explicit conceptual adjustments for the hidden economy. The difference between GDP(P) and GDP(I) was split between mixed income (85%) and COE (15%). At that time, it was assumed that although corporations had incentives to conceal profits, the practice was thought to be very minor. On that basis, outside of this model an additional small adjustment of around +0.3 per cent was applied to corporate profits.

On the basis of the 1994 benchmark values for mixed income and COE, to estimate a series before and after 1994, the growth in the relevant income/corporation tax rates multiplied by the growth in IDBR turnover of sole traders/partners by industry (after 1994) or the growth in mixed income or partner income (before 1994) is applied to the 1994 value. An additional adjustment factor is also included to account for cyclical effects on producers’ propensity to evade taxation, which should average a value of 1 over the whole business cycle. The model is therefore dependent on four variables:

1. the previous year’s income concealment adjustment for industry i
2. the changes in the basic rate of tax for self employed or employees in previous years.
3. the change in turnover for industry i
4. a cyclical adjustment factor

The basic formula to calculate the concealed income adjustment is defined as:

$$CI_{n,i} = CI_{n-1,i} * \rho * \Delta turnover_i * \Delta Tax$$

where CI represents concealed income within compensation of employees or mixed income.

In reviewing the strength of the existing model, two aspects stood out. Firstly, the method as it stands uses a benchmark from 1994 and therefore if retained it would need to be updated under the reservation. Secondly, the method makes several strong assumptions, which are that:

- there is little tax evasion on salaries;
- there is a proportional relationship between the tax rate and evasion; and
- there is a proportional relationship between IDBR turnover (for years after 1994) and the level of tax evasion in all institutional sectors, including those poorly covered by the IDBR.

A number of further points should also be noted:

- The use of the then-improved GDP(P) as the yardstick has some justification but is at odds with international guidance at the time (see OECD, 2002), which suggested using the discrepancy between GDP(I) and GDP(E). Both the use of GDP(P) or GDP(E), however, can be challenged: as both GDP(P) and GDP(E) are predominantly based on surveys, differences between them and to GDP(I) may simply represent sampling errors rather than conceptual differences. Consider the case where either GDP(E) or GDP(P) were lower than GDP(I): this would imply an unlikely scenario of respondents, in aggregate, overstating their incomes for tax purposes.
- Currently, only a small number of industries, dominated by construction, retail and hotels/restaurants, are adjusted for concealed activity on the production approach, yet it seems highly probable that concealment – which can arise from accounting errors as well as deliberate behaviour - occurs to some degree in almost all industries.
- Furthermore, various system and classification changes since the model's inception have made many of the original assumptions opaque and difficult to test. Without this information it is difficult to provide a full justification for the method, nor can we update the benchmark in a consistent fashion.
- Finally, the method depends on IDBR data that has known coverage issues regarding the small businesses that are key (as is addressed in the second element of ONS's review of exhaustiveness). Although adjustments were made to resolve this, this area of activity seems likely to have been under-represented given more recent information.

3.2 New approach to concealed income and activity to be implemented in September 2015

The new method replaces the GDP(P) – GDP(I) discrepancy approach currently used with estimates of concealed income based on Tax Gaps analysis, which HMRC has published annually for the last five years. In this analysis, HMRC undertakes random audits of Corporation Tax (CT), Pay-As-You-Earn (PAYE) and Self-Assessment (SA) tax returns and establishes where erroneous or deliberate misreporting has taken place (HMRC, 2014).

From these cases, estimates of missing income and corporation tax at the UK level are established and these form the basis of the ONS model.

3.2.1 HMRC Tax Gaps analysis

HMRC produces estimates of various types of missing tax revenue due to under-declaration of taxable liabilities, including taxes on pay, income and profits.

The tax schemes of interest here are SA, PAYE and CT. Separate estimates are made for gaps relating to different taxes; some schemes, such as CT, correspond to a single tax, but others, such as PAYE and SA, cover multiple taxes. The sample results for observed missing tax revenue are multiplied by the number of taxpayers in the relevant business population, as well as being supplemented by separate estimates for 'ghosts' (individuals who do not declare any of their earnings to HMRC) and 'moonlighters' (individuals who have more than one source of income but only declare one for tax purposes) to produce UK-level tax gap estimates. In more detail:

3.2.1.1 Self-assessment (SA) returns method

The estimate is constructed using a random sample of SA individuals sent notice to file a return. "Individuals" includes sole traders, pension taxpayers, and partners in partnerships with 4 or fewer partners. Notice to file is issued to all individuals who HMRC believes should pay tax through the SA system but not to individuals "new" to the tax system who may have started a business: this latter category will be excluded from the sample. It is possible that under new legislation allowing withdrawal of a notice to file, taxpayers who are actually dormant will be included in the sample. The sample is stratified by income; in recent financial years the sample size has averaged around 3,000 returns. Selected returns are subjected to a full investigation and the estimate of the tax gap for the selected records is multiplied by the number of taxpayers sent a notice to file; this is the total SA tax gap.

An adjustment is applied by HMRC for partnerships with more than four partners ("large partnerships"). It is assumed that the ratio of tax gap to tax liability is the same for large partnerships as for individuals, so the ratio is applied to the tax liability and added to the total SA tax gap.

3.2.1.2 Pay-As-You-Earn (PAYE) method

The estimate is constructed using the Employer Compliance random sampling programme. This is a random sample of PAYE schemes by small and medium employers (SMEs) only. The sample is stratified by number of employees and by legal status (sole trader, partnership or corporation); sample sizes in recent financial years have averaged around 1,500 returns. Selected returns are subject to investigation and the total tax gap is multiplied by the number of PAYE schemes to calculate the total estimated tax gap.

An adjustment is applied for large PAYE employers and the method is exactly the same as for large partnerships.

3.2.1.3 Corporation Tax (CT) method

The CT estimate is constructed using a random sample of small and medium enterprises sent a notice to file; sample sizes in recent financial years have averaged around 400 returns. Selected returns are subject to investigation and the total tax gap is multiplied by the number of live SME trader schemes to get the total tax gap.

An adjustment is applied for large businesses. This is based on a 100 per cent sample of cases and derived from administrative information from HMRC's personalised interactions with each business.

3.2.1.4 Further adjustments

In spite of any random audit process, some misreporting will remain undetected. To account for this HMRC uses a modelling approach devised by the US Inland Revenue Service (IRS) to generate multipliers to scale up the point estimates. These range from 1.3 to 1.9 for non-business and business SA estimates, while for CT estimates a factor of 1.4 is used.

Adjustments are also made for individuals who do not declare any earnings of any form to HMRC, known as 'ghosts', and for 'moonlighters' who have two or more sources of income but only declare one for tax purposes. Ghosts' tax gap estimates use information on Jobseeker's Allowance claimants, income reported to HMRC by known individuals and estimates of illegal immigration flows. HMRC's moonlighters' tax gap estimates are based on work by Statistics Denmark looking at the shadow economy in Germany, Great Britain and Scandinavia (HMRC, 2014).

3.2.2 ONS model for estimating concealed income and activity

The ONS model transforms the HMRC tax gaps data in four key stages:

1. Estimates for any missing tax that do not relate to production and the generation of income are removed;
2. next, relevant tax rates to scale the missing tax levels into estimates of concealed income are applied;
3. those concealed income series are then backcast and converted from financial to calendar years; and
4. finally, the UK-level data by type of missing tax are pro-rated down to industry level using a new ONS model.

Each stage is explained below.

3.2.2.1 Removal of missing taxes not relevant to GDP

SA tax gaps include capital gains tax (CGT) and taxation of occupational pensions, which are not charged on components of GDP(I). It is assumed that there is no tax evasion relating to pensions, as these cannot be managed without full documentation; and capital gains tax is not charged on produced income (the criterion for relevance to GDP(I)), so both these taxes must therefore be removed.

Further, all tax gap estimates include debt (unpaid tax accrued in previous years), avoidance (tax not paid because of actions considered to be outside the spirit of the law) and legal interpretation (similar to avoidance). To perform avoidance or legal interpretation, the income must first be reported, and so it will not be concealed. Debt is also not relevant as the income was reported, but circumstances intervened such that the tax was not paid. All these elements are therefore removed from the totals.

Even after the above is taken into consideration, work is required to account for the fact that liability to corporation tax, and income tax for SA taxpayers, is not calculated on GOS or MI but on "profits liable to corporation tax" and the corresponding variable for income tax. There are very large conceptual differences between GOS and MI and the latter two variables. For example, corporations are allowed to deduct some losses incurred in past years from their profits to calculate "profits liable to corporation tax," but no such deduction is applied to calculate GOS. As such, "profits liable to corporation tax" can be understood as the sum of a number of positive and negative variables, some of which apply to GOS, some of which don't. All tax evasion by corporations, and therefore Tax Gaps relating to CT after removing debt, avoidance and legal interpretation, must relate to misreporting one of these variables. (Of course, the same corporation may misreport more than one variable.) In estimating concealed income to adjust GDP, only the Tax Gaps related to those variables that apply to GOS should be included. Exactly the same argument applies to the SA Tax Gaps. (It does not apply to the PAYE Tax Gaps, because the concept of COE very closely matches that for taxable salaries).

Unfortunately, a breakdown of the Tax Gaps by “variable” in the sense of the preceding paragraph is not available. However, HMRC provided ONS with data from the random enquiry programme that break misreporting down by the reason given by tax inspectors for estimating the Tax Gaps. For SA, this breakdown is by the value of yield (i.e. the reclaimed tax); but for CT and PAYE, by the number of cases. In both cases the studies aggregate data for a number of years.

This information, while very useful, was not designed to inform GDP estimates, so analysis had to be carried out on these reasons to determine whether they applied to GOS/MI or not.

Using these studies two models were derived:

- (1) One in which all categories determined by the analysis to be likely to relate to GOS/MI were included. This resulted in about 75% of the tax gap left in scope after the removal of CGT, debt, etc. being used for the GDP adjustment.
- (2) A stricter version of (1), whereby only the categories that specifically mentioned concepts related to production in the current year, and therefore determined by the analysis to be very likely to relate to GOS/MI. This resulted in 40-45% of the tax gap left in scope after the removal of CGT, debt, etc. being used for the GDP adjustment.

After comparing the results with existing evasion estimates, we opted to use model (1), as (2) implied implausibly low levels of evasion for the self-employed and partnerships, where national and international evidence points to the bulk of evasion taking place (United Nations, 2008). No reduction was made to COE tax evasion as this component sits entirely within the production boundary.

3.2.2.2 Scaling of relevant missing tax to concealed income

After the removal of missing taxes not relevant to the measurement of GDP, what remains are the tax gaps relating to GDP(I) components, which must then be scaled to the concealed income to which they relate. Data are not available on the characteristics of tax evaders (e.g. income) relative to the whole population, so it is assumed that they face the same effective income tax rate as the population paying that tax (it was possible to estimate a sub-division to partners and employees/sole traders based on an industrial analysis). Dividing the tax gap by this rate gives the concealed income. Similarly, for corporation tax the tax gap is divided by the average effective corporation tax rate. Missing income data are already available for ghosts and moonlighters, because these estimates are already compiled by multiplying concealed income by tax rates.

3.2.2.3 Backcasting and conversion to calendar years

The tax gaps estimates are available for financial years from 2008-09 to 2012-13. To backcast the tax gap for earlier years, the growth rates for PAYE receipts (COE), CT liabilities (CT) and SA income tax receipts (partnerships' profits and individuals' mixed income) are used. Quarterly values are then estimated from the financial year annuals in a common statistical process called 'splining', using X-13 ARIMA. The quarterly values are then aggregated to form calendar year totals. This approach assumes, in the absence of earlier data on tax liabilities, that tax gaps are proportionate to the amount of tax collected.

3.2.2.4 Derivation of concealed income and activity by industry

Derivation of concealed income by industry

In the absence of explicit industry information from the tax gaps data, a model has been derived by ONS based on the likely prevalence of evasion based on industry characteristics.

It is informed by reasoning set out in earlier work carried out by the Australian Bureau of Statistics, reported to the United Nations in their review of national approaches to shadow economy measurement (United Nations, 2008).

In which industries is tax evasion concentrated? Three factors are considered to make tax evasion in an industry more likely:

- a) A larger proportion of total profits in the industry is made by smaller units (classified by profits or employment) as opposed to larger ones.
- b) A larger proportion of total profits in the industry is made through sales to households as opposed to units in other sectors.
- c) The characteristics of the industry are such that enforcement is relatively difficult for other reasons.

There are three principal arguments for a). One is that smaller units typically face less regulation than larger ones, with less stringent financial reporting requirements (e.g., smaller units may not be required to register for VAT or PAYE), less return per prosecution for the tax authorities and, to some extent, more difficulty in detecting evasion. Another is that larger units are likely to be publically traded corporations that distribute their profits to shareholders rather than directly to private owners. If the profits are understated in the accounts for tax evasion reasons, the full amount cannot be distributed to shareholders without making detection trivial. Finally, larger units are better able to afford the transaction costs (e.g., specialist tax accountants) for tax avoidance rather than tax evasion; and avoidance does not run into the detection problem previously mentioned.

An argument for b) is that households (except those trading as businesses) are not required to account for their purchases in any way. Therefore, under-statement of income cannot be detected by examining the records of the purchaser, as would be the case for sales to other sectors where purchasers are not complicit in the evasion. This is compounded by the fact that households are likely to pay in cash for many transactions, which greatly increases the scope for evasion because there is no automatic record of the sale or money transfer. Cash transactions are especially important for PAYE (COE) evasion.

An argument for c) is that the main other characteristic mentioned in the international guidance is the degree of industry regulation. Any regulatory activity additional to tax enforcement multiplies the opportunity for checking the financial accounts and activity of the business and therefore potentially detecting evasion. This is especially the case where, as in the financial sector, the regulation involves checking the accounts of the business.

The reporting units and income to which these patterns of behaviour should be applied to are the levels of sole trader or partnership, and corporate profits and COE. Multiplicative factors can then be applied to create the final adjustment. The general formula for the pattern value $P_{i,c,t}$ for industry i and component v (where v is GOS of corporations, GOS of partnerships, MI or COE) for year t , will then be, where $B_{i,c,t}$ is the basis value (profits or COE) for industry i in year t , and $F_{i,x,v,t}$ is the multiplicative factor applied to industry i to take into account theoretical factor x (where $x = a, b$ or c , from the previous section):

$$P_{i,v,t} = B_{i,v,t} * F_{i,a,v,t} * F_{i,b,v,t} * F_{i,c,v,t}$$

As usual for prorating, the final evasion adjustment $E_{i,v,t}$ is then, where $E_{T,v,t}$ is the total evasion adjustment calculated from the Tax Gaps for component v in year t , and n is the number of industries,

$$E_{i,v,t} = E_{T,v,t} * P_{i,v,t} / S(P_1 \dots P_n)$$

All that remains is to specify the calculation of the three F factors for each industry and evasion category (mixed income, gross operating surplus of partnerships, gross operating surplus of corporations and COE).

$F_{i,a}$ for sole traders and partners can be calculated using published HMRC data on the income (profits) earned by the self-employed by broad SIC2007 industry. These data have income bands showing the number of businesses and the total amount earned by all businesses within the industry/band cell. These industries are broad groups which will each be composed of one or more of the 114 SIC2007 industry groups (SICs) used in Supply Use, and the values for each broad industry can be applied to all of the 114 SICs that compose it. The new model defines $F_{i,a}$ to be the proportion of total income earned by all businesses with profits of below £30,000, as businesses of this size are unlikely to have more than one or two employees.

$F_{i,a}$ for corporate profits can be calculated using the proportion of ABS profits (approximate GVA (aGVA) minus employment costs) by legal status 1 entities of employment size band 1-200. This involves the assumption that pension contributions follow the same pattern as salaries. Again, broad industry groups have to be translated to their member SICs.

In the absence of weights based on the relative prevalence of the three business types within the industry, $F_{i,a}$ for COE has been calculated as the product of the three $F_{i,a}$ adjustments set out above.

$F_{i,b}$ for all four categories should be the same. It is calculated as the proportion of the total expenditure on each industry's principal product that is household final consumption expenditure (HHFCE). The construction industries – 41, 42 and 43 – are treated differently, as these are not final consumption products but capital ones. Data exists for the value of dwellings produced and $F_{41,b}$ and $F_{43,b}$ are set as the ratio of this value divided by total gross fixed capital formation (GFCF) of the construction product. $F_{42,b}$ is zeroed as industry 42 is civil engineering, where it is assumed that a negligible amount of income is concealed.

$F_{i,c}$ clearly cannot be a consistently calculated ratio as the other two are as there is no numerical measurement of the level of regulation in an industry. This means that the determination of values for $F_{i,c}$ must to some extent be a matter of judgment. The new model simplifies by applying values of less than 1 to industries the ONS considered to be heavily regulated and for the remainder the $F_{i,c} = 1$. Due to high levels of regulation in legal and accountancy services (69.1, 69.2) and human health services (86) these industries were assigned C factors of 0.5 for sole-traders/partners; for corporations, legal services was assigned zero and accountancy 0.5. No evasion is assumed to occur within the financial services industries (64, 65, 66), only some avoidance, so these are assigned C factors of zero.

In summary, the full equations for the all-industry adjustment to the GDP(I) components in the years for which Tax Gaps are available are as follows:

Taking all terminology from above, and in addition defining $TG_{S,t}$ as the Tax Gap excluding debt, avoidance etc for category S (S is Corporation Tax, SA (total), SA (partners), SA (sole traders), PAYE, Ghosts (employees), Ghosts (self-employed), or Moonlighters) for year t, A_S as the proportion of Tax Gaps in category S relevant to GDP derived from the multi-year tables mentioned above, and TR_{TAX} as the average effective tax rate for tax TAX (TAX is Corporation Tax, Income Tax for employees and sole traders, and Income Tax for partners):

- (1) $E_{T, \text{GOSCORPS}, t} = (TG_{\text{CORPORATION TAX}, t} \times A_{\text{CORPORATION TAX}}) / TR_{\text{CORPORATION TAX}}$
- (2) $E_{T, \text{GOSPARTNERS}, t} = (TG_{\text{SAPARTNERS}, t} \times A_{\text{SATOTAL}}) / TR_{\text{INCOMETAXPARTNERS}}$
- (3) $E_{T, \text{MI}, t} = [(TG_{\text{SASOLETRADERS}, t} + TG_{\text{SELFHOSTS}, t} + TG_{\text{MOONLIGHTERS}, t}) \times A_{\text{SATOTAL}}] / TR_{\text{INCOMETAXEMPSOLE}}$
- (4) $E_{T, \text{COE}, t} = (TG_{\text{PAYE}, t} + TG_{\text{EMPGHOSTS}, t}) / TR_{\text{INCOMETAXEMPSOLE}}$

The total adjustment to GDP(I) is given by the formula

$$(5) E_{T, \text{GOSCORPS}, t} + E_{T, \text{GOSPARTNERS}, t} + E_{T, \text{MI}, t} + E_{T, \text{COE}, t}$$

These adjustments can then be placed into the industry allocation formulae above to give the results by industry.

Derivation of concealed activity (adjustment to GDP(P) components) by industry

Having derived missing income by industry, missing activity must also be modelled. As production estimates in the National Accounts are derived from different sources than the income estimates, at least part of the missing activity associated with missing income may already be captured within those sources, though this can only be estimated indirectly under present conditions.

In the existing model, it is assumed that production data from the ABS are essentially exhaustive once adjustments are made to the construction, wholesale and retail and hotel/restaurant industries; a handful of very minor adjustments are also made to other industries that are of little economic consequence. In deriving the new model, this assumption was dismissed as no longer plausible. Equally, though, an upper limit of 100 per cent of misreporting to HMRC manifesting as misreporting to the ABS is assumed to be implausible, as ABS returns are collected much earlier than HMRC returns and at least some businesses may only decide to misreport as the HMRC deadline approaches.

Ideally, it would be possible to compare individual reporters' data provided to HMRC and the ABS to test if, for cases where HMRC had identified income concealment, the values reported to the ABS were higher once any conceptual and reporting period differences had been accounted for. However, current data sharing frameworks between ONS and HMRC do not allow for individual records to be compared, while high enough levels of aggregation to permit comparison yielded inconclusive results as unobservable conceptual differences between tax and business survey returns may have explained some or all of the observed differences.

An alternative approach was therefore adopted, which compared HMRC and ABS reporting period overlaps over several calendar years, on the assumption that if exactly the same reporting period was used to report to both HMRC and the ABS, it was likely the same set of accounts were used and any income concealment to HMRC would fully manifest in the ABS return as well, while at the other extreme, no overlap was assumed to imply no income concealment manifesting in the ABS return.

The bulk of HMRC tax returns filed cover standard financial years from April to March, though the second most frequent reporting period covers a standard calendar year. The reverse is true for ABS returns, while both sources receive a minority of returns for various non-standard reporting periods. Notably, in more recent years the ABS has allowed respondents to submit forms covering periods of up to two years. This was borne out, as when comparing reporting period frequencies between the two sources, around 30 per cent of ABS returns had no overlap at all with HMRC returns, which allow for a more limited reporting window. Taking this into account, the overall correlation coefficient between ABS and HMRC returns' reporting periods was just under 50 per cent. Given the conceptual limitations of this indirect approach, this was rounded to 50 per cent and is applied as a scaling factor of 0.5 to the concealed income estimates to form estimates of concealed

activity. In short, the new model assumes 50 per cent of activity generating income concealed to HMRC is already captured in the production measure while the other 50 per cent is missing from it.

For the majority of industries, the scaled concealed income estimates by industry were then split into adjustments to output and intermediate consumption using information on types of misreporting from the tables provided by HMRC mentioned in 3.2.2.1. However, existing estimates of hidden construction activity were used in place of those implied by the new model, as they come from a robust, industry-specific model.

3.3 Current approach to under-coverage of unincorporated small businesses

In the 1990s, ONS worked alongside the Inland Revenue (now part of HMRC) and other government departments to derive estimates of unincorporated small businesses that were outside of the sample frame for the then Annual Business Inquiry, which was and still is the IDBR. The IDBR automatically includes any enterprise above the government-determined VAT threshold in force in any given year and/or which is registered on a PAYE scheme. Any business with turnover at or above the VAT threshold must register for VAT collection purposes, though some businesses below the threshold may also choose to register. Small business which neither register for VAT nor participate in a PAYE scheme will not be on the register.

The cross-departmental work derived estimates of the missing turnover generated by small business not on the IDBR using a mixture of annual tax assessments of small businesses made by the Inland Revenue prior to the SA system coming into force in 1996, plus SA returns thereafter, plus finally a combination of various items of research and modelling.

A set of adjustment factors by SIC1992 industry were produced; the unique factor for each industry was then applied both to ABS estimates of total sales and total business purchases (excluding employment costs), which form the primary feeds into the National Accounts concepts of output and intermediate consumption. As the conceptual adjustments applied to total sales differ from those applied to total business purchases, the final percentage impacts on output, intermediate consumption and GVA by industry tend to differ to the original equal raw impact on sales and purchases.

Following its inception, the model was updated annually by ONS to reflect new industry information and also to reflect the move from the SIC1992 to SIC2003 industry classification. However, since its last update in 2004, the industry uplifts have not been re-estimated, only remapped from SIC2003 to SIC2007. The under-coverage estimates therefore require updating

3.4 New approach to under-coverage of unincorporated small businesses to be implemented in September 2015

The new approach is essentially an update of the existing method, as it is the timeliness of the data rather than the validity of the model that requires resolution. ONS commissioned HMRC to analyse turnover by SIC2007 industry group for unincorporated small business, comprising sole traders and partnerships neither registered for VAT (by legal requirement or voluntarily) nor participating in a PAYE scheme. HMRC provided data for financial years 2010-11 and 2011-12, from which calendar year 2011 estimates are derived.

For each of the industries defined in the Supply and Use tables (a mix of SIC2007 2-4 digit level groups), a single factor unique to each industry is derived as follows:

- (1) The HMRC estimate of turnover for enterprises within that industry with no VAT/PAYE record for 2010-11 and 2011-12 is converted to a calendar year 2011 value, V_S . In the absence of a time series, this was done simply by multiplying each year by factors of 0.75 and 0.25 respectively ;

- (2) next, the calendar year value, V_s , is added to the turnover value for that industry collected by the ABS, V_L , covering enterprises registered for VAT and/or participating in a PAYE scheme, to create total turnover by industry, $V_S + V_L = V_T$;
- (3) finally, the ratio of V_T/V_L gives the uplift factor applied to both total sales and total purchases values estimated for that industry by the ABS.

As only a single calendar year has so far been estimated, the industry factors were backcast to 2009 using a simple linear process and existing adjustment factors were used for prior years. The decision to backcast only to 2009, the first new year required in order to meet Eurostat guidelines on benchmark assumptions, was made because it was assumed the impact of the 2009 recession on enterprise characteristics limited the modelled estimates' validity for earlier years. Labour market data shows that, following the recession, there has been a significant increase in the number of self-employed, as many people made redundant have set up their own businesses.

As GDP(I) and GDP(E) are already exhaustive in this respect, the changes to GDP(P) were aligned by applying coherence adjustments to components of income and expenditure.

3.5 Clarity on the potential interaction between concealed income/activity and under-coverage adjustments

It should be noted that the new under-coverage method does not present any double-counting issues when combined with the new concealed income/activity method, as the two measure different areas of exhaustiveness. The under-coverage model supplements ABS-derived estimates of production activity in GDP(P) for enterprises in the IDBR sample frame with those from outside. However, the additional output, intermediate consumption and GDP(P) is derived from turnover data supplied to HMRC by small businesses that are as likely to misreport their income and production activity as similar businesses that are already included. It is therefore mutually consistent and necessary to uplift GDP(P) for missing small business activity based on HMRC turnover data, and then to uplift GDP(I) and GDP(P) for concealed income and activity based on separate analysis.

4 Impact of the changes

4.1 Impact of changes to estimates of concealed income and activity

The new method impacts on GDP(I) via exhaustiveness adjustments in the following areas:

- GOS via changes to Private Non-Financial Corporations (PNFCs') profits (sector S.11)
- Income earned by partnerships (sector S.11)
- Households' mixed income via changes to income earned by sole traders (sector S.14)

No impact occurs for financial corporations or quasi-corporations (see 3.2.2.4), nor for the Non-Profit Institutions Serving Households (NPISH) (S.15) or Government (S.13) sectors.

It impacts on GDP(P) via exhaustiveness adjustments to:

- market output and intermediate consumption produced/demanded by PNFCs and non-financial quasi-corporations (sector S.11)

Finally, it impacts on GDP(E) via coherence adjustments only (to align the three approaches to GDP) to Household Final Consumption Expenditure (HHFCE) (S.14)

Numerical impacts can be found in the article ['Impact of ESA1995 Changes on current price Gross National Income Estimates, 2002 to 2010'](#).

4.2 Impact of changes to estimates of under-coverage of unincorporated small businesses

The new method impacts on GDP(P) via exhaustiveness adjustments in the following areas:

- market output and intermediate consumption of unincorporated household businesses (S.14)

Coherence adjustments to align measure of GDP(I) were applied to PNFC profits (S.11PR) and self-employment income (S.11;S.14), while for GDP(E) they were applied to HHFCE (S.14), GFCF(S.11) and imports and exports of services (S.11).

5 Conclusions

The changes described here, which will be implemented in Blue Book 2015, improve the exhaustiveness of approaches to GDP by:

- Introducing a comprehensive, up-to-date set of industry-level estimates for concealed income and activity in GDP(I) and GDP(P), based on audits of administrative tax data
- Updating estimates of activity by unincorporated small businesses with contemporary estimates based on administrative tax data

These changes have been implemented to improve comparability in estimating Gross National Income (GNI) across the EU. They will have an impact on the estimation of the UK's GNI, which is used to calculate the UK's contribution to the EU budget.

6 Contact

Richard Wild
Office for National Statistics
sut@ons.gsi.gov.uk
+44 (0)1633 456379

7 References

Eurostat (1996). *European System of Accounts ESA 1995*, Luxembourg: Office for Official Publications of the European Communities.

Eurostat (2005). 'Eurostat's tabular approach to exhaustiveness: Guidelines', ref. Eurostat/C1/GNIC/050 EN, paper from the 5th Meeting of the GNI Committee, 5-6 July 2005.

Eurostat (2013). *European System of Accounts ESA 2010*, Luxembourg: Publications Office of the European Union.

HMRC (2014). 'Measuring tax gaps 2014 edition', available at:
<https://www.gov.uk/government/statistics/measuring-tax-gaps>

ONS (2014). *The United Kingdom National Accounts: The Blue Book*, 2014 edition, available at:
<http://www.ons.gov.uk/ons/rel/naa1-rd/united-kingdom-national-accounts/the-blue-book--2014-edition/index.html>

ONS (2015). 'Impact of ESA1995 changes on current price GDP estimates', available at:
<http://www.ons.gov.uk/ons/guide-method/method-quality/specific/economy/national-accounts/changes-to-national-accounts/blue-book-and-pink-book-2015-changes/index.html>

United Nations (2008). *Non-observed economy in National Accounts: Survey of Country Practices*, New York: United Nations Publications.