ECONOMIC STATISTICS CENTRE OF EXCELLENCE

A COLLABORATION WITH Office for National Statistics

Communicating the Uncertainty of Estimates of International Comparisons of Productivity

Ana Beatriz Galvão

ESCoE Discussion Paper 2022-02

January 2022

ISSN 2515-4664

DISCUSSION PAPER

Communicating the Uncertainty of Estimates of International Comparisons of Productivity
Ana Beatriz Galvão
ESCoE Discussion Paper No. 2022-02
January 2022

Abstract

Different methodologies for computing total annual hours worked are adopted across G7 statistical offices to estimate productivity, and, therefore, we cannot be sure about international comparisons of productivity. This paper describes the design, implementation, and results of an online randomised controlled experiment to assess the impact of communicating this uncertainty in comparing productivity between the UK and the other G7 countries. The online survey results support the proposed communication tools as an effective way of conveying the uncertainty on the estimates of the international comparison of productivity for the UK public. They are effective even for respondents with limited knowledge of what productivity is. But communication tools are likely to be more helpful to members of the public that are familiar with the concept as they are better at making an inference based on the communicated data.

Keywords: randomised online experiment, uncertainty communication, productivity measurement, G7 countries.

JEL classification: C92, E70, 047

Ana Beatriz Galvão, University of Warwick, ana.galvao@wbs.ac.uk

This research has been funded by the Office for National Statistics as part of the research programme of the Economic Statistics Centre of Excellence (ESCoE).

Published by:
Economic Statistics Centre of Excellence
National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
United Kingdom
www.escoe.ac.uk

ESCoE Discussion Papers describe research in progress by the author(s) and are published to elicit comments and to further debate. Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Economic Statistics Centre of Excellence (ESCoE), its partner institutions or the Office for National Statistics (ONS).

© Ana Beatriz Galvão

Communicating the Uncertainty of Estimates of International Comparisons of Productivity

Ana Beatriz Galvão[†]
University of Warwick; ONS and ESCoE Fellow.

January 2022

Abstract

Different methodologies for computing total annual hours worked are adopted across G7 statistical offices to estimate productivity, and, therefore, we cannot be sure about international comparisons of productivity. This paper describes the design, implementation, and results of an online randomised controlled experiment to assess the impact of communicating this uncertainty in comparing productivity between the UK and the other G7 countries. The online survey results support the proposed communication tools as an effective way of conveying the uncertainty on the estimates of the international comparison of productivity for the UK public. They are effective even for respondents with limited knowledge of what productivity is. But communication tools are likely to be more helpful to members of the public that are familiar with the concept as they are better at making an inference based on the communicated data.

Keywords: randomised online experiment, uncertainty communication, productivity measurement, G7 countries.

JEL codes: C92, E70, 047

[†] This research has been funded by the Office for National Statistics as part of the research programme of the Economic Statistics Centre of Excellence (ESCoE). The project was led by Josh Martin (ONS). We are grateful to the ONS fellowship scheme for financial support. I would like to thank several ONS staff for their inputs on visualisation designs, survey questions and comments and suggestion to an early version of this paper. They are (in alphabetical order): Peter Broad, Sumit Dey-Chowdhury, Zoe Harland, Charles Lound, Stuart Mackenzie, Sanjiv Mahajan, Josh Martin, Gentiana Roarson and Dan Taylor. This paper has also benefited from comments and suggestions by Johnny Runge (NIESR). Corresponding address: Prof. Ana Galvao, Warwick Business School, University of Warwick, Coventry CV4 7AL, UK. Email: ana.galvao@wbs.ac.uk

1. Introduction

For international comparisons, productivity is usually measured as output per labour input. As the output is measured using GDP, the measurement of the labour input needs to be consistent with it, and there are a set of methods to obtain estimates of the labour input. There is a wide variation across statistical offices of methodologies to estimate the labour input (hours) when calculating productivity (Ward, Zini and Marianna, 2018; Duhn, 2019). This variation in methods to estimate hours when measuring productivity makes it harder to compare productivity across countries. As we observe differences in productivity levels across countries, we are unsure if these are true because country X is more productive than country Y or because country X measures labour input using a different methodology from country Y.

This paper describes how we use an online randomised controlled experiment to assess the impact of communicating the uncertainty in comparing productivity between the UK and the other G7 countries. The design of the experiment follows Galvão and Mitchell (2021) study on the communication of the uncertainty on GDP estimates. We have, however, to consider that the nature of the uncertainty on the international comparison of productivity differs from the one observed in GDP estimates.

The uncertainty on productivity estimates arises from the fact that we consider four different ways of estimating the labour input: (i) a direct method using the labour force survey, (ii) a simplified component method using the labour force survey combined with the usual number of hours per week worked, (iii) a component method, and (iv) a measure of hours using the total working population and a set number of working hours per week. These methods are described in Duhn (2019), who reports that the direct method (i) may overestimate worked hours as total hours are computed by summing the average weekly hours declared by households in the labour force survey. The component method (iii) instead uses information on jobs and contractual hours to estimate the annual number of hours worked. The simplified component method, proposed by the OECD, helps to reduce the gap between the direct and the component method. Duhn (2019) describes the simplified components method as useful for countries with limited access to administrative data required to compute hours using the components approach (iii). The last method (iv) is a rough estimate of total hours using the total number of people employed in a given year and assuming a 30-hour worked week over 52 weeks.

This method is convenient when no alternative estimate of total hours consistent with aggregate output (gross value added) is available.

For this research project, the ONS computed productivity for international comparisons with hours estimated using the last method (iv) for all G7 countries. This was the only type of estimate obtained for Canada and Japan. For France, Germany, Italy and the UK, a direct estimate of hours (i) was also available. A components-based estimate (iii) was obtained for France, Germany and Italy and the US. For the UK and the US, productivity with hours estimated using the simplified component method (ii) was also accessible.

Instead of choosing a specific estimate of productivity for each country, we exploit the effects of communicating a range of productivity estimates in this report. The aim of an international comparison of productivity (ICP) is to publish statements as "country G was X% more (or less) productive than country Y in 20XX". For this quantitative statement, we have up to nine different estimates for "X%". These arise from the fact that we have three different estimates for the UK (the reference country) and up to three different estimates for the country the comparison is made. In particular, we have nine estimates for how the productivity of the US, Germany, France and Italy is compared to the UK productivity, and three estimates for similar comparison for Japan and Canada. In some of the visualisation tools described in section 2, we indicated all pairwise comparisons (up to nine) mixing methods both for the reference country (UK) and the remaining six G7 countries.

More details on the communication tools designed for this experiment are described in section 2. In section 3, we provide details of the survey implementation. The analysis of the survey results is in section 4. The last section summarises the main empirical results and discusses their implications for the publication of international comparisons of productivity.

2. Communication Tools

As we aim to assess the effect of communicating the uncertainty on international comparison of productivity estimates, we consider seven different communication tools.

The first communication tool is the information provided to the control group – that is, the group of respondents that receive no information on the variability of estimates for the productivity comparisons across countries. This method of communication is the one applied by the ONS in October 2018 to the release of the 2016 ICP estimates.‡

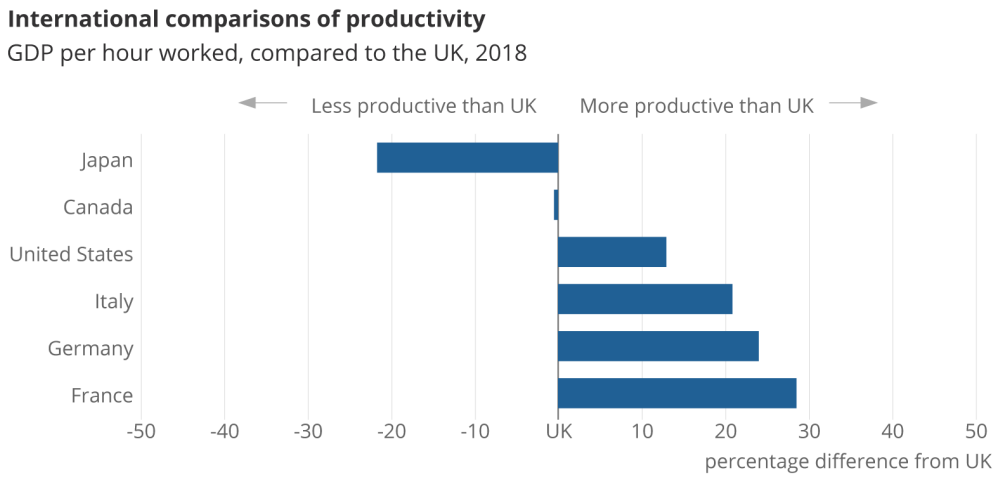
The second communication tool (group 2) describes in words a range of values for productivity but does not include a visualisation of the estimates. The remaining five communication tools combine different ways of visualising the productivity estimates with the description in words of the estimates, as in the case of group 2. These figures -- visualisation tools-- have been created by an ONS visualisation team. They convey the information of up to nine pairwise comparisons depending on the method used to estimate hours for the UK and the G7 country.

The communication tools are all presented in Figure 1. As described in more detail in the next section, each respondent is randomised into a group and only sees the communication tool allocated to their group, as displayed in Figure 1.

As commented earlier, the reason we are not certain about international productivity comparisons is not due to the sampling measurement error, which has a probabilistic interpretation, of the labour market surveys employed to compute hours worked. Instead, the main issue is the variety of hours worked methodologies used as described earlier. The visualisations in Groups 3, 5 and 7 may be better to convey this type of uncertainty than the visualisations in groups 4 and 6. In the case of the visualisation tool of group 6, the dot marks the average across estimates, and the limits are set of the maximum and the minimum.

‡ See https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/bulletins/internationalcomparisonsofproductivityfinalestimates/2016

Figure 1: The seven Communication Tools
Group 1 (Control Group)



Source: Office for National Statistics

The figure shows that UK productivity (output per hour worked) was:

- 29% lower than that of France
- 24% lower than that of Germany
- 21% lower than that of Italy
- 13% lower than that of the United States
- 1% higher than that of Canada
- 22% higher than that of Japan

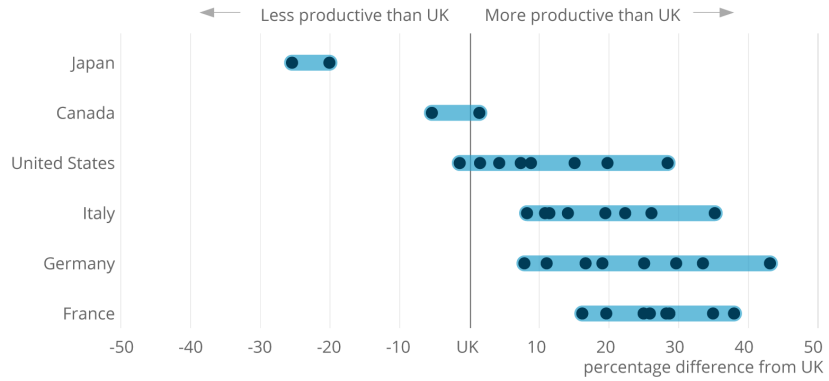
Group 2 (Visualisation Control)

There are many ways to measure and compare productivity across countries. Depending on how you measure hours worked, it is estimated that the UK's productivity (output per hour worked) was:

- between 17% and 39% lower than that of France
- between 8% and 44% lower than that of Germany
- between 9% and 36% lower than that of Italy
- between 29% lower and 1% higher than that of the United States
- between 2% lower and 5% higher than that of Canada
- between 20% and 25% higher than that of Japan

Group 3 (Dot Plot with a Shaded Area)
Text as Group 2 +

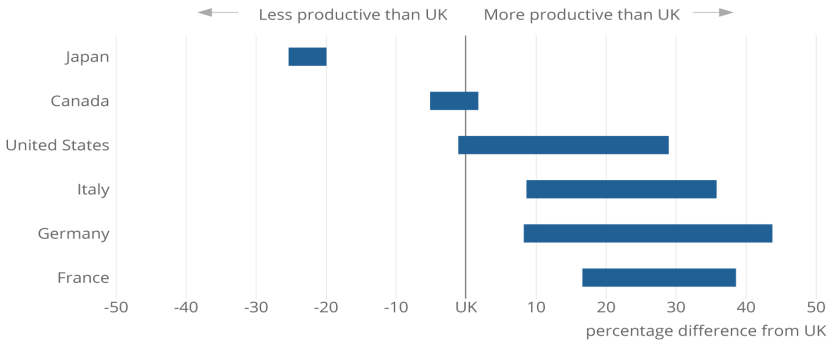
International comparisons of productivity
GDP per hour worked, compared to the UK, 2018



Source: Office for National Statistics

Group 4 (Bar Plot for Estimated Range Limits)
Text as Group 2 +

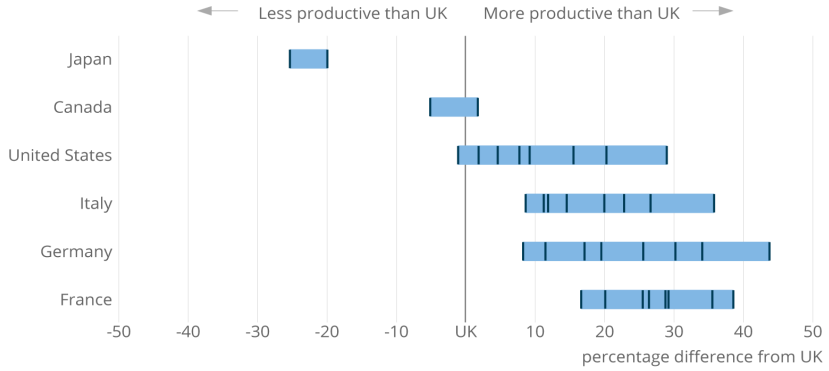
International comparisons of productivity
GDP per hour worked, compared to the UK, 2018



Source: Office for National Statistics

Group 5 (Bar Plot with Estimated Values)
Text as Group 2 +

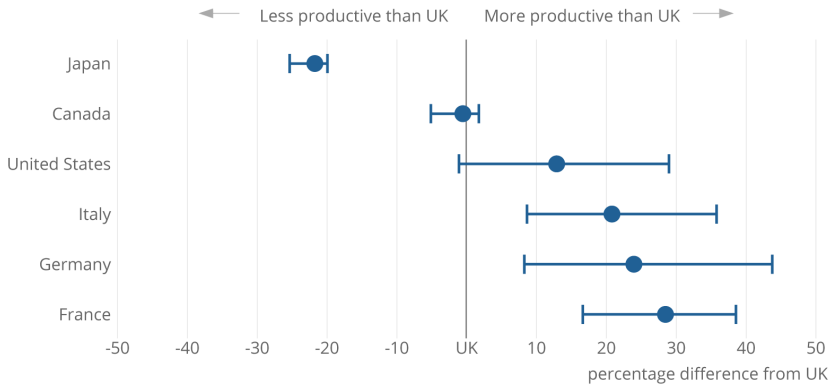
International comparisons of productivity
GDP per hour worked, compared to the UK, 2018



Source: Office for National Statistics

Group 6 (Whisker Plot)
Text as Group 2 +

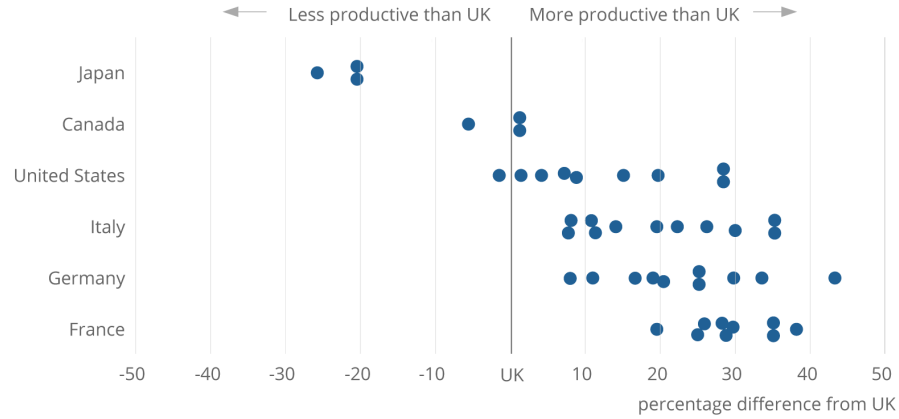
International comparisons of productivity
GDP per hour worked, compared to the UK, 2018



Source: Office for National Statistics

Group 7 (Dot Plot)
Text as Group 2 +

International comparisons of productivity
GDP per hour worked, compared to the UK, 2018



Source: Office for National Statistics

3. Survey Design and Implementation

The online survey was performed by dynata. The pilot survey was carried out on October 7th, 2021, and the results we are reporting are from the main survey executed between October 9th-19th, 2021. Dynata implemented quotas on respondents to represent the UK population in terms of gender, age (only above 18) and location (across 12 UK regions). The first three survey questions asked about these three individual characteristics. In Appendix Table A1, we show the average answer for each of the 20 questions in the survey over the 3,499 respondents. Based on average responses to questions 1 to 3, UK population representation seems to be well achieved.

Dynata have also administered the randomisation of the respondents across the seven groups (target of 500 respondents per group). Appendix Table A2 shows how four individual characteristics, which could potentially affect our results, are distributed across groups. As our comparisons set group 1 as the control group, we indicated in bold the cases in which the proportion of respondents with the indicated characteristic is statistically different than the proportion in the control group. The results in Table A2 suggest that the groups generally have similar composition except that group 5 has a higher proportion of respondents that know what productivity is, and group 4 has a higher proportion of full-time employees. We account for this limited group heterogeneity by adding these individual characteristics in the regression that measures the causal effects of the communication tools to each one of the outcomes considered. Estimated impacts and their significance are in general robust to the inclusion of these controls.

The survey has two main parts described in Table A1. In the first part, we asked questions regarding individual characteristics. These include questions that measure respondents' knowledge about the ONS and productivity before seeing one of our communication tools. The questions from this first part are in line with the survey in Galvão and Mitchell (2021) to evaluate the effect of communication uncertainty on GDP estimates.

The questions in the second part of the survey were asked after the respondents were randomised in each group and observed their allocated communication tools. We use the questions in this second part to define our outcome variables. Table A1 presents

the sample proportion of respondents for each answer. Questions 17-20 are in line with the survey in Galvão and Mitchell (2021), but questions 11-16 have been specially designed to measure respondents' perceptions on comparisons of productivity across countries.

3.1 Evaluated Outcomes

This subsection describes how questions 11-20 are used to define our set of outcome variables.

Our outcome variables include an evaluation of the public behaviour regarding three different ways of using the information on a release of the international comparison of productivity. First, we assess whether the respondents can say whether one country is more productive than another, that is, the probability that they will correctly answer questions 11 and 12. In the case of the comparison between the UK and Italy, the estimates in Figure 1 suggest that Italy is more productive than the UK as all nine estimates of their relative performance indicate that (that is, Q11C is the correct answer). In the case of the comparison between the UK with the United States, one cannot be sure as there are estimates that support the possibility that the US is less productive, even though most of the estimates suggests the US is more productive (that is, Q12B is the correct answer). The sample proportion in Table A1 indicates that only 26.3% of the respondents selected Q12B.

Second, we evaluate whether the respondent can address a quantitative comparison using the numbers communicated by correctly answering questions 13 and 14. We start with the case of Germany, where all % changes are positive. The correct answer is Q13B. We then move to the ambiguous US case, where the correct answer is also the range of values (Q14B). The majority of the respondents have selected Q13B and Q14B (50% and 47%).

Third, we evaluate whether the respondent can correctly infer using the information provided the relative performance between countries in questions 15 and 16. In question 15, we ask how the French standard of living is compared with the UK after explaining that productivity is relevant for living standards. As all estimates indicate France is more productive than the UK, the correct answer is Q15B. In contrast, in question 16, we go back again to the case of the US, where the correct answer is then

Q16D (not sure it depends on how productivity is measured). Interestingly, 20% of the respondents chose Q16D, while 26% replied the equivalent Q12B.

The survey also includes questions to evaluate respondents' trust in the ONS, their understanding of why productivity estimates vary, and their perceptions of communication informativeness and uncertainty awareness. The alternatives in question 18 are reasons for the variety of international comparisons of productivity estimates. We evaluate the effect of the different communications tools on answering each one of these alternatives.

We use the answers to questions 17, 19 and 20 to define outcome variables that measure the degree of trust in the ONS (question 17), awareness that estimates are uncertain (question 19), and the degree of informativeness of the communication tool observed (question 20).

4. Respondents Reaction to Communication Tools

In this section, we analyse how respondents' answers differ depending on which communication tool they observe. Our primary interest is to evaluate the treatment effect of groups 2 to 7 tools (in Figure 1) compared to group 1—the control group. The treatment is evaluated on the set of outcomes described in section 3.1. There are three types of outcome variables. The first is a set of binary variables that indicate whether the respondent has correctly answered questions 11 to 16. The second is again a set of binary variables indicating the answers to question 18 on the reasoning behind the many estimates. The third employs the information on which category was chosen for questions 17, 19 and 20 to measure effects on the public trust in the ONS, awareness of the uncertainty of productivity estimates, and the degree of informativeness of the communication tool. By evaluating whether there is statistical evidence that groups 2 to 7 deliver a different set of outcomes than group 1, we are assessing whether there is a causal effect from the productivity communication tools to these three different types of outcome variables, which describe the UK public behaviour.

In the remainder of this section, we provide evidence on how the communication tools affect the outcomes in Tables 1 and 2 and how individual characteristics affect the

outcomes. In these tables, we have one outcome per row, with outcomes split into three panels to present the three types described above.

Table 1 – Measurement of the Treatment Effects of ICP Communication tools.

These estimates are obtained using multiple regressions for each outcome response listed in each row. Differences between the indicated and control groups (group 1) that are statistically significant at the 5% level are shown in bold. t-statistics computed using robust standard errors. The sample is size is 3,499.

		Group 1	*Group 2*	*Group 3*	*Group 4*	*Group 5*	*Group 6*	*Group 7*
			Visualisation Control	*Dot Plot with a shaded area*	*Box Plot*	*Box Plot with Estimates*	*Whisker Plot*	*Dot Plot*
	Outcome responses:	Mean	How the comm. tools change the probability of answering the indicated **correct response** compared to the control group (in p.p.)					
Q11 C	Italy is more productive	66.1%	**-9.5**	**-7.3**	**-8.4**	-3.4	-3.3	-1.5
Q12 B	Not sure -it depends on how productivity is measured	15.4%	**17.4**	**12.2**	**8.7**	**12.2**	**12.4**	**14.2**
Q13 D	Germany is between 8% and 44% more productive than the UK	5.8%	**58.8**	**51.0**	**46.1**	**52.9**	**47.6**	**56.4**
Q14 D	The United States is between 1% less productive and 29% more productive than the UK	7.4%	**53.8**	**41.2**	**43.1**	**46.5**	**43.0**	**49.6**
Q15 C	higher than in the UK	72.3%	**-9.5**	-2.3	-5.3	-0.6	**-6.1**	-5.5
Q16 D	Not sure -it depends on how productivity is measured	12.8%	**10.0**	**6.8**	**5.5**	**7.2**	**7.2**	**6.0**
	Outcome responses:	Mean	How the communication tools change the probability of answering the indicated r**esponse** compared to the control group (in p.p.)					
Q18 A	The methods used to measure productivity in other countries are more appropriate than the ones adopted by the ONS for the UK.	16.0%	2.4	2.0	2.1	0.0	**-4.2**	-1.6
Q18 B	The ONS method to measure productivity in the UK is more appropriate than the ones adopted by other countries.	17.0%	3.8	**6.2**	**4.9**	1.8	3.8	0.4
Q18 C	The output of the ONS method to measure productivity is a range of values.	25.2%	4.7	**6.5**	4.7	**7.2**	**8.7**	**8.5**
Q18 D	Differences in how productivity is measured across countries.	35.9%	2.9	-1.5	-1.9	-0.8	-0.9	-1.5
Q18 E	Mistakes at the ONS.	4.0%	-1.0	-0.8	1.8	1.2	0.6	-1.8
Q18 F	The ONS has vested interests in results / manipulates production or collection	6.6%	-1.0	1.8	2.4	1.8	1.2	-0.8
Q18 G	The Government has vested interests in the results / interferes in production or collection	13.4%	-0.6	-3.2	-1.3	0.0	-2.6	-3.8
	Outcome responses:	Mean	How **trust**, **awareness** of uncertainty, and perceived **informativeness** of the communication change in comparison with the control group					
Q17	Degree of trust of the ONS (5=trust greatly)	3.67	0.03	0.03	0.04	0.05	0.01	0.03
Q19	Awareness of uncertainty (4 = not at all surprised)	2.56	-0.01	-0.02	-0.05	-0.01	0.06	0.04
Q20	Informativeness (4 = a lot of info on unc.)	2.40	**0.25**	**0.26**	**0.20**	**0.29**	**0.30**	**0.25**

Table 2: Comparing Communication Tools: How Key Visualisation Tools compare with words+numbers communication

These estimates are obtained using a single regression for each outcome response in each row. Differences between the indicated group and group 2 – visualisation control-- that are statistically significant at the 5% level are shown in bold. t-statistics computed using robust standard errors. The sample size of each single regression is 1,000.

		Group 2 *Visualisation Control*	*Group 3* *Dot Plot with a shaded area*	*Group 5* *Box Plot with Estimates*	*Group 7* *Dot Plot*
		Mean:	How the comm. tools change the probability of answering the indicated correct response in comparison with group 2 (p.p.)		
Q11 C	Italy is more productive	56.6%	2.20	**6.10**	**8.00**
Q12 B	Not sure -it depends on how productivity is measured	32.8%	-5.20	**-5.14**	-3.20
Q13 D	Germany is between 8% and 44% more productive than the UK	64.6%	**-7.80**	**-5.80**	-2.40
Q14 D	The United States is between 1% less productive and 29% more productive than the UK	61.2%	**-12.60**	**-7.20**	-4.20
Q15 C	higher than in the UK	62.8%	**7.20**	**8.90**	4.00
Q16 D	Not sure -it depends on how productivity is measured	22.8%	-3.20	-2.75	-4.00
			How the comm. tools change the probability of answering the indicated response in comparison with group 2 (p.p.)		
Q18 A	The methods used to measure productivity in other countries are more appropriate than the ones adopted by the ONS for the UK.	18.4%	-0.40	-2.36	-4.00
Q18 B	The ONS method to measure productivity in the UK is more appropriate than the ones adopted by other countries.	20.8%	2.40	-1.96	-3.40
Q18 C	The output of the ONS method to measure productivity is a range of values.	30.0%	1.80	2.46	3.80
Q18 D	Differences in how productivity is measured across countries.	38.8%	-4.40	-3.73	-4.40
Q18 E	Mistakes at the ONS.	3.0%	0.20	2.21	-0.80
Q18 F	The ONS has vested interests in results / manipulates production or collection	5.6%	2.80	2.82	0.20
Q18 G	The Government has vested interests in the results / interferes in production or collection	12.8%	-2.60	0.60	-3.20
			How trust, awareness of uncertainty, and perceived informativeness of the communication change compared to group 2.		
Q17	Degree of trust of the ONS (5=trust greatly)	3.71	0.00	0.02	0.00
Q19	Awareness of uncertainty (4 = not at all surprised)	2.56	-0.01	0.00	0.05
Q20	Informativeness (4 = a lot of info on unc.)	2.65	0.01	0.05	0.01

Table 3 – The effect of knowledge of productivity, attention to economics news and university graduation on outcome measures.

These estimates are obtained using multiple regressions for each listed question/answer, but single regressions give the same results as the group allocations are not correlated due to survey design. Differences that are statistically significant at the 5% level are in bold. t-statistics computed using robust standard errors. The sample is size is 3,499.

		Know what productivity is (Q10A, B)	Reads Economic News (Q7E, Q7F	Graduate (Q4A, Q4B, Q4C)
		How the individual charact. affects the probability of answering the indicated correct response (p.p.)		
Q11 C	Italy is more productive	**4.80**	**4.20**	-0.40
Q12 B	Not sure -it depends on how productivity is measured	0.80	**-4.90**	-0.50
Q13 D	Germany is between 8% and 44% more productive than the UK	**4.50**	1.60	0.60
Q14 D	The United States is between 1% less productive and 29% more productive than the UK	**4.20**	0.70	**3.70**
Q15 C	higher than in the UK	2.27	**5.15**	**5.57**
Q16 D	Not sure -it depends on how productivity is measured	1.19	**-4.40**	-1.80
		How the individual charact. affects the probability of answering the indicated response (p.p.)		
Q18 A	The methods used to measure productivity in other countries are more appropriate than the ones adopted by the ONS for the UK.	1.10	**4.63**	**13.00**
Q18 B	The ONS method to measure productivity in the UK is more appropriate than the ones adopted by other countries.	1.30	**5.80**	**5.07**
Q18 C	The output of the ONS method to measure productivity is a range of values.	**4.48**	**10.50**	**7.40**
Q18 D	Differences in how productivity is measured across countries.	**7.38**	**8.80**	**8.20**
Q18 E	Mistakes at the ONS.	-0.30	1.29	0.58
Q18 F	The ONS has vested interests in results / manipulates production or collection	1.23	**2.60**	1.58
Q18 G	The Government has vested interests in the results / interferes in production or collection	-1.00	0.40	-0.10
		How trust, awareness of uncertainty, and perceived informativeness of the communication is affected by the ind. Characteristic		
Q17	Degree of trust of the ONS (5=trust greatly)	**0.07**	**0.32**	**0.38**
Q19	Awareness of uncertainty (4 = not at all surprised)	**0.07**	-0.06	**-0.17**
Q20	Informativeness (4 = a lot of info on unc.)	0.05	**0.15**	**0.16**

4.1- Analysis of the Treatment Effects

Table 1 presents the estimated difference in the probability of selecting the corrected indicated answer between a treatment group (2 to 7) and the control group (group 1). Differences that are significant at the 5% level are in bold.

For questions 11 and 15, where the correct answer was a statement saying that a G7 country was more productive than the UK, the communication of different productivity estimates decreased the probability of selecting the correct answer. This suggests that some of the communication tools (an exception is Group 7 – the dot plot) lead to excessive respondent under-confidence regarding making a statement with certainty. Instead, when looking at questions 12-14 and 16, we find that all treated respondents were more likely to answer the questions correctly than the control group. This implies that our suggested communication tools effectively convey the information when we have many estimates for comparing productivity internationally.

The effectiveness of the proposed communication tools is also supported by the fact the respondents in Groups 2 to 7 more frequently say that the ONS communication was informative about the fact that the productivity was uncertain (Q20). The treated groups are also more likely to articulate that "the output of the ONS method to measure productivity is a range of values".

In contrast, there is no difference between the treated groups and the control group regarding (i) trust in the ONS, (ii) how surprised the respondents are about the ICP estimates, and (iii) the view that vested interests are affecting the production/collection of the data.

The additional comparisons presented in Table 2 are helpful to disentangle whether there is any gain in adding a visualisation method in comparison with the text included for Group 2 describing the range of values for the productivity estimates. The entries in Table 2 are the estimated differences in the probability of selecting the indicated answer between a group that includes a visualisation tool (3, 5, 7) and the no visualisation group (group 2). Differences that are significant at the 5% level are in bold. For this analysis, we include the visualisation tools that show the values for all different estimates (dot plots (with and without shades) and box plots with lines for each estimate). We consider these as better visualisation tools to characterise that

productivity estimates are uncertain due to the method to compute hours instead of the effects of sampling uncertainty.

The results in Table 2 suggest that the visualisation tools help respondents answer questions 11 and 15 correctly. They make the respondent more confident that a G7 country is indeed more productive than the UK, even if there are different estimates. This increased confidence makes respondents that have seen the dot plot with shade and the box plot with an estimate to make more mistakes in answering questions 13 and 14. A similar effect is not found when using the dot plot as a visualisation tool as in the case of Group 7. Visualisation tools have the same effect on all the other questions presented in Table 2.

4.2– The Effect of Individual Characteristics

As respondents were randomised into groups, individual characteristics should not affect our interpretation of the estimates of change to the control group/group 2 described in Tables 1 and 2 as causal treatment effects on the set of outcomes. Reassuringly if we estimate a multiple regression for each outcome variable in Table 1, including the individual characteristics in Table A2 as controls, we find that the estimates in Table 1 are robust to this inclusion. However, it is interesting to check whether specific individual characteristics correlate with getting the right answers to questions 11 to 16 or affect respondent perceptions in questions 17 to 20.

Table 3 presents the estimates of the effects of some individual characteristics on answering questions 11 to 20. We consider three individual features. First, we examine whether the respondent knows the concept of productivity by answering question 10 correctly (either Q10A or Q10B). Based on Table A1, about 1/3 of the respondents answered this question correctly. Second, we consider whether the respondent is frequently in tune with economic news (chose option Q7E or Q7F in question 7). This is again equivalent to about 1/3 of the respondents. The third characteristic considered is whether the respondent has graduated from university, that is, has answered Q4A, Q4B or Q4C in question 4. Because these individual characteristics are correlated, the estimates in Table 3 are based on a multiple regression that includes dummy variables for each of these individual characteristics. The regression also includes the group indicators with coefficient estimates presented in Table 1.

The entries in Table 3 measure how individuals with the indicated characteristic perform on average in the questions listed compared to individuals who do not have it. For all three characteristics analysed, we find respondents are more likely to select options A to D in question 18, that is, the options that compare the quality of the ONS productivity estimates with those from other countries.

Respondents that know the concept of productivity are more likely to answer questions 11 to 16 correctly. They are also more likely to say they trust the ONS and are more aware of the uncertainty in the productivity estimates. Respondents who are attentive to economic news are even more trustful of the ONS, but they are more likely to make mistakes when the correct answer is "not sure" (questions 12 and 16) and to say the productivity measures are affected by the government vested interests. Respondents with a university degree are more likely to get questions 14 and 15 correctly, and they are more trustful of the ONS than individuals who do not have a university degree.

5 – Summary and Conclusions

In this section, the main empirical results from section 4 are summarised to establish the main conclusions of the randomised online experiment.

The statistical analysis described in section 4 suggests that:

A. There is strong statistical evidence that the new communication tools, described in section 2 and applied to groups 2 to 7, lead to an improved assessment by the UK public of how productivity estimates are compared across countries.
B. Some of the visualisation tools, described in section 2 and applied to groups 3 to 7, improve respondents' confidence in comparison with a simple description of the range of ICP estimates when answering how a G7 country compares to the UK. The dot plot is particularly recommended as it does not create additional confusion when answering questions 12, 13, 14 and 16 while improving questions 11 and 15.
C. There is no evidence that the new communication tools influenced the trust of the respondents in the ONS and their perception of vested interests in the collection/production of the data.

D. Respondents who know the concept of productivity are more likely to make correct assessments based on ICP estimates, be more aware of the lack of certainty in productivity estimates, and better understand the variety of estimates.

We can then conclude that the proposed communication tools effectively convey the uncertainty on the estimates of the international comparison of productivity. They are effective even for respondents with limited knowledge of what productivity is. But the communication tools will likely be more helpful to members of the public that are familiar with the concept as they are, in general, better when making an inference based on the communicated data.

References:

Duhn, M. (2019) Improving estimates of labour productivity and international comparisons. ONS article available at https://www.ons.gov.uk/economy/economicoutputandproductivity/productivitymeasures/articles/improvingestimatesoflabourproductivityandinternationalcomparisons/2019-01-09

Galvão, A. B. and Mitchell, J. (2021) Communicating Data Uncertainty: Multi-Wave Experimental Uncertainty for the UK. ESCoE Working Paper Series.

Ward, A., Zinni, M. B. and Marianna, P. (2018) International productivity measures: are labour input measures comparable? OECD Statistics Working Papers, 2018/12.

Appendix

Table A1: Sample Proportion for each Survey Question

The total number of survey respondents is 3,499.

Part 1: Questions before the ICP communication

Question number	Proportion (%)	Questions
Q1		What is your gender?
Q1A	49.2%	Male
Q1B	50.4%	Female
Q1C	0.1%	Other
Q1D	0.3%	Prefer not to say
Q2		What is your age?
Q2A	11.4%	18-24
Q2B	16.7%	25-34
Q2C	18.1%	35-44
Q2D	18.0%	45-54
Q2E	15.3%	55-64
Q2F	20.5%	65 and above
Q3		Where do you live?
Q3A	4.2%	North East
Q3B	11.3%	North West
Q3C	8.6%	Yorkshire and The Humber
Q3D	7.1%	East Midlands
Q3E	8.9%	West Midlands
Q3F	9.3%	East of England
Q3G	13.2%	London
Q3H	13.9%	South East
Q3I	8.8%	South West
Q3J	7.9%	Scotland
Q3K	4.9%	Wales
Q3L	2.0%	Northern Ireland
Q4		What is your highest educational qualification?
Q4A	2.7%	PhD or equivalent doctoral level qualification
Q4B	15.0%	Masters or equivalent higher degree level qualification (MA, MSc, PGCE etc.)
Q4C	23.4%	Bachelors or equivalent degree level qualification (BA, BSc etc.)
Q4D	7.1%	Post-secondary below-degree level qualification
Q4E	21.8%	A Level / NVQ Level 3 /Scottish higher or equivalent
Q4F	22.8%	GCSE / O Level / NVQ Level 1 / NVQ Level 2 or equivalent
Q4G	2.3%	CSE or equivalent
Q4H	1.4%	Any other qualification
Q4I	3.3%	None of the above
Q5		Please select the option that best describes your current employment situation:
Q5A	43.8%	Employed full-time
Q5B	13.1%	Employed part-time
Q5C	0.3%	Furloughed from full-time work

Q5D	0.3%	Furloughed from part-time work
Q5E	4.1%	Unemployed and currently looking for work
Q5F	6.2%	Not working and not currently looking for work
Q5G	3.1%	Full-time Student
Q5H	21.3%	Retired
Q5I	3.3%	Self-employed
Q5J	4.6%	Unable to work
Q6		In which, if any, have you ever studied economics? (select all that apply)
Q6A	31.2%	At school
Q6B	26.1%	In higher education (e.g. university, college)
Q6C	7.2%	Through self-directed study (books)
Q6D	6.1%	Self-motivated study (course)
Q6E	2.0%	Other
Q6F	4.0%	Don't know / can't recall
Q6G	38.7%	Not applicable – I have never studied economics
Q7		How frequently, if at all, do you access news stories related to economics or the economy?
Q7A	7.1%	Never
Q7B	19.4%	Rarely
Q7C	12.6%	Monthly
Q7D	25.0%	Weekly
Q7E	20.7%	Almost every day
Q7F	10.8%	Every day
Q7G	4.5%	Not sure
Q8		The Office for National Statistics (ONS) is the UK's largest independent producer of official statistics and the recognised national statistical institute of the UK. Before answering this survey, had you ever heard of the ONS?
Q8A	59.2%	Yes, I had heard of them, and knew what they did
Q8B	24.6%	Yes, I had heard of them, but didn't know what they did
Q8C	11.7%	No, I had never heard of them
Q8D	4.5%	Not sure / don't know
Q8_part2		When did you first hear about the ONS?
Q8_part2_A	66.6%	Before the coronavirus pandemic
Q8_part2_B	19.3%	During or since the start of the coronavirus pandemic
Q8_part2_C	14.1%	I don't remember.
Q9		Personally, how much trust do you have in economic statistics produced by the Office for National Statistics (ONS)? For example, on unemployment, inflation or economic growth?
Q9A	18.8%	Trust them greatly
Q9B	52.4%	Tend to trust them
Q9C	10.3%	Tend not to trust them
Q9D	1.6%	Distrust them greatly
Q9E	16.9%	Not sure / don't know
Q10		To the best of your knowledge, which option most accurately describes what the economic term "productivity" refers to?
Q10A	17.6%	Productivity measures how much output is produced per hour worked

Q10B	16.4%	Productivity measures how much output is produced per worker
Q10C	31.8%	Productivity measures how much output is produced in the country
Q10D	10.4%	Productivity measures how much output is produced compared to other countries
Q10E	7.8%	Productivity measures how much prices have increased during the last year.
Q10F	6.1%	I don't know what productivity is
Q10G	9.9%	I have heard about productivity but not sure what it is.

Part 2: Questions after ICP communication. For questions 11-16, the communication tool is visible while answering the question.

Question n.	Proportion (%)	Questions
Q11		Which country is more productive? The UK or Italy?
Q11A	20.3%	The UK is more productive
Q11B	18.4%	Not sure – it depends how productivity is measured
Q11C	61.3%	Italy is more productive
Q12		Which country is more productive? The UK or the United States?
Q12A	22.6%	The UK is more productive
Q12B	26.4%	Not sure – it depends how productivity is measured
Q12C	51.0%	The United States is more productive
Q13		What is the difference in productivity between Germany and the UK?
Q13A	10.7%	Germany is 8% more productive than the UK
Q13B	24.8%	Germany is 24% more productive than the UK
Q13C	14.1%	Germany is 44% more productive than the UK
Q13D	50.5%	Germany is between 8% and 44% more productive than the UK
Q14		What is the difference in productivity between the United States and the UK?
Q14A	9.8%	The United States is 1% less productive than the UK
Q14B	24.8%	The United States is 13% more productive than the UK
Q14C	18.5%	The United States is 29% more productive than the UK
Q14D	47.0%	The United States is between 1% less productive and 29% more productive than the UK
Q15		A country's productivity is important for living standards. Based on the information given, would you expect that living standards in France are:
Q15A	9.8%	lower than in the UK
Q15B	68.1%	higher than in the UK
Q15C	9.2%	about the same as the UK
Q15D	13.0%	not sure – it depends how productivity is measured
Q16		Based on the information given, would you expect that living standards in the United States are:
Q16A	13.1%	lower than in the UK
Q16B	54.1%	higher than in the UK
Q16C	13.9%	about the same as the UK
Q16D	18.9%	not sure – it depends how productivity is measured
Q17		Based on the information given, how much trust do you have in the estimates of productivity by the Office for National Statistics (ONS)?
Q17A	16.4%	Trust them greatly
Q17B	46.0%	Tend to trust them
Q17C	30.7%	Neither trust nor distrust them

Q17D	5.3%	Tend not to trust them
Q17E	1.6%	Distrust them greatly
Q18		ONS publishes a range of different estimates of how the UK's productivity compares to any given country. Why do you think they do this? (select all that apply)
Q18A	16.1%	The methods used to measure productivity in other countries are more appropriate than the ones adopted by the ONS for the UK.
Q18B	20.0%	The ONS method to measure productivity in the UK is more appropriate than the ones adopted by other countries.
Q18C	31.0%	The output of the ONS method to measure productivity is a range of values.
Q18D	35.4%	Differences in how productivity is measured across countries.
Q18E	4.0%	Mistake at the ONS.
Q18F	7.4%	The ONS has vested interests in results / manipulates production or collection
Q18G	11.8%	The Government has vested interests in the results / interferes in production or collection
Q18H	24.0%	Don't know / not sure
Q18I	0.5%	Other [please write any other reasons]
Q19		Are you surprised that estimates of productivity are uncertain?
Q19A	12.3%	Very surprised
Q19B	31.9%	Fairly surprised
Q19C	42.5%	Not that surprised
Q19D	13.2%	Not at all surprised
Q20		Thinking back to the ONS statement about how productivity varied between countries, how much information did it give that the estimate may be uncertain?
Q20A	6.2%	None at all
Q20B	38.3%	Very little
Q20C	42.6%	Some
Q20D	13.0%	A lot

Table A2: Allocation of Individuals into Groups: percentage of respondents in each group that have the characteristic indicated in the first column

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7
know what productivity is (Q10A, Q10B)	30.9%	28.2%	33.4%	32.7%	**40.9%**	35.8%	35.8%
reads economics news (Q7E, Q7F)	29.5%	30.0%	34.6%	30.1%	31.5%	31.4%	32.8%
Graduate (Q4A, Q4B, Q4C)	40.4%	39.6%	41.0%	41.3%	43.1%	40.6%	42.0%
Full-time employee (Q5A)	39.4%	44.8%	45.4%	**48.9%**	40.9%	46.2%	41.2%

Note: Cases in which the proportion is statistically different at the 5% level from Group 1 (control group) are indicated in bold. The measurement of the treatment effects in the regressions presented in Tables 1 and 2 do not change when these individual characteristics are included as control variables.