UNIVERSITY OF OXFORD

Discussion Papers in Economic and Social History

UNIVERSITY OF OXFORD

Discussion Papers in Economic and Social History

Number 2 December 1995

TURNING WATER INTO WINE

NEW METHODS OF CALCULATING FARM OUTPUT AND NEW INSIGHTS INTO RISING CROP YIELDS DURING THE AGRICULTURAL REVOLUTION

LIAM BRUNT

TURNING WATER INTO WINE

NEW METHODS OF CALCULATING FARM OUTPUT AND NEW INSIGHTS INTO RISING CROP YIELDS DURING THE AGRICULTURAL REVOLUTION[1]

LIAM BRUNT
NUFFIELD COLLEGE, OXFORD
ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES, PARIS
TELEPHONE AND FAX + 33 1 44 16 24 62
LIAM.BRUNT@NUF.OX.AC.UK

[1] This research was undertaken as part of my M.Phil. degree, which was kindly funded by the Economic and Social Research Council. I would like to thank Charles Feinstein, James Foreman-Peck, Tim Leunig, Joachim Voth, and especially Lucy White, for helpful comments and discussion. Any remaining errors or omissions are my responsibility.

and 16 per cent for oats). This creates an overestimate of total output, and the overestimate rises as production becomes more skewed in favour of wheat. This proposition is demonstrated in Appendix II.

3. The Solution.

The solution to the externality problem in constructing the output series is to adjust downwards the wheat yield. We purge it of the effect of other crops in the rotation to find the True Wheat Yield. We can then simply multiply the arable acreage by the True Wheat Yield to obtain a correct estimate of total output. We call this the Adjusted Method. This adjustment process has the disadvantage that we need to use price data in order to scale down the wheat yields. However, the process also reveals a great deal of new information. In particular, it allows us to isolate the impact on wheat yields of every other crop in the rotation - for the first time, we can quantify the importance of new crop technology on wheat yields during the Agricultural Revolution.

Let us take the case of a six course rotation of wheat, beans, barley, peas, oats and turnips. We first set up the arbitrage conditions for wheat and the other crops. The example of beans is given below. Let the True Wheat Yield be Y^{W*} and let the extra wheat yield resulting from growing beans be Y^{WBe}. We know that the revenue from the True Wheat Yield ($Y^{W*}.P^{W}$) must be equal to the direct revenue from the crop of beans ($P^{Be}.Y^{Be}$), plus the indirect revenue resulting from higher wheat yields ($Y^{WBe}.P^{W}$). Otherwise the farmer would be better off growing less wheat and more beans. Then,

$$Y^{W*}.P^{W} = P^{Be}.Y^{Be} + Y^{WBe}.P^{W}$$

$$\Rightarrow \quad Y^{W*} = \frac{P^{Be}.Y^{Be}}{P^{W}} + Y^{WBe}$$

$$\Rightarrow \quad Y^{WBe} = Y^{W*} - \frac{P^{Be}.Y^{Be}}{P^{W}} \qquad (6)$$

But we also have an identity. The True Wheat Yield (Y^{W*}) equals the total wheat yield (that is, the *measured* wheat yield, Y^{W}) minus the wheat yield due to beans (Y^{WBe}) and all the other crops. Then,

$$Y^{W*} = Y^{W} - Y^{WBe} - Y^{WB} - Y^{WP} - Y^{WO} - Y^{WT} \qquad (7)$$

By substituting equation (6) and the other arbitrage conditions into equation (7) we easily find the True Wheat Yield:

$$Y^{W*} = \frac{Y^{W} + \frac{P^{Be}.Y^{Be}}{P^{W}} + \frac{P^{B}.Y^{B}}{P^{W}} + \frac{P^{P}.Y^{P}}{P^{W}} + \frac{P^{O}.Y^{O}}{P^{W}} + \frac{P^{T}.Y^{T}}{P^{W}}}{6} \qquad (8)$$

We can now proceed to calculate total wheat equivalent output by multiplying the total acreage of all the crops involved by Y^{W*}. This is the Adjusted Method (AM). So total wheat-equivalent output (Q^{W*}) is:

$$Q^{W*} = Y^{W*}.(A^{W} + A^{Be} + A^{B} + A^{P} + A^{O} + A^{T}) = AM \qquad (9)$$

Using the Adjusted Method, we calculated the total wheat-equivalent output for all the farms in the Arthur Young dataset. (Obviously, we made the calculations using the specific crop rotation relevant to each farm; the 160 farms for which we have all the required data used 44 different rotations between them). We also calculated total wheat-equivalent output using the New Method and the Traditional Method for the same sample of farms. As noted above, the existence of bare fallows - and the complete absence of fallow price data - raise serious problems for the Traditional and Adjusted Methods. Moreover, there is no farm price data for clover (only yield data). So we had to assume that the price of clover was uniform across all farms at 438 pence per load - this approximate valuation is derived from Arthur Young.[9] Due to these problems we have made three different comparisons of the three Methods, based on different samples of farms. Table 1

[9] Young A, *Eastern Tour*, p164.

below gives the estimated output using all three methods for *only those farms (13) which had no clover and no fallows*. Table 2 gives the results for *only those farms (31) which had no fallow* (ie it includes some farms which grew clover). Table 3 gives the results for *all farms (160)*.

Table 1. 'Pure' sample (13 observations).

Method	Mean Output (bu)	SE of Mean Output (bu)	Variance
Traditional	7395	2674	92967007
Adjusted	7354	2679	93300724
New	10781	4065	214828235

Table 2. 'Clover but no fallow' sample (31 observations).

Method	Mean Output (bu)	SE of Mean Output (bu)	Variance
Traditional	5084	1355	56931703
Adjusted	5068	1356	56970643
New	6311	1865	107779694

Table 3. 'Full' sample (160 observations).

Method	Mean Output (bu)	SE of Mean Output (bu)	Variance
Traditional	2630	386	23897823
Adjusted	2602	383	23446804
New	3346	511	41798811

In all cases the three output series are very highly correlated (between 96 and 99 per cent using Pearson's Product Moment Correlation Coefficient). The upward bias of the Traditional Method is apparent in all the calculations, but appears to be slight. The New Method clearly provides an over-estimate, but is it not clear by how much. The data from Table 1 suggest that the New Method is over-estimating by 47 per cent. But in Table 3 (the full sample) we know that the Adjusted Method is giving an underestimate due to the fallow problem - so the New Method must be

over-estimating by less than 29 per cent. We will now consider in more detail the properties of each output series.

4. The Bias, Efficiency and Heteroskedasticity of the Three Methods.

It is generally desirable to use an estimator which is *unbiased* - that is, an estimator which gives a prediction which is correct on average. It is also desirable to have an *efficient* estimator - that is, an estimator which has a low variance. These two properties are combined in the Mean Squared Error (MSE), which is defined as:

$$\text{MSE} = \text{variance} + (\text{bias})^2$$

The estimator with the lowest Mean Squared Error is generally the preferred estimator. In Table 4 below we give the MSE for each of the three Methods, based on the 'pure' sample. The Traditional Method and the Adjusted Method perform virtually identically, and they are both much better than the New Method.

Table 4.

Method	Mean Squared Error	Variance	Bias
Traditional	92967791	92967007	28
Adjusted	93300724	93300724	0
New	226572564	214828235	3427

It would useful to know the MSE for the full sample (ie for those farms which use clover and a bare fallow) because this is the most likely situation faced by a researcher. However, when fallows are included the Adjusted Method and the Traditional Method are both biased downwards. Moreover, the Traditional Method is inherently biased upwards (as shown in Appendix II) so it is impossible to know even the direction of the overall bias. Since *all* the Methods are biased in the presence of bare fallows, we cannot calculate the MSE because we cannot quantify the bias.

In fact, although the Mean Squared Error is the usual tool for discriminating between estimators, it is not clear that it is appropriate in this instance. Let us consider the effect of bias. Precise estimates are sometimes essential in empirical work. If we are comparing the absolute level of output - for example, making international productivity comparisons - then we need a measure which is not biased. Or at least the bias must be the same in all cases, a condition which is very unlikely to hold in practice.[10] But biased estimates are not *necessarily* a problem in empirical work. For many applications, particularly the estimation of production functions, we normally take natural logarithms of all the variables. The overestimate of total output inherent in the Traditional and New Methods then biases the constant downwards but leaves all the coefficients unchanged. This can be shown as follows. The relationship between 'true' total output and the factor inputs K_i has the following form:

$$\ln Q^{W*} = C + \alpha \ln K_i + \mu$$

If we use an estimate of total output which is consistently too high by around 47 per cent (such as the New Method, as suggested by Table 1) then the equation takes the form:

$$\ln(Q^W/1.47) = C + \alpha \ln K_i + \mu$$

$$\Rightarrow \quad \ln Q^W = C - \ln\frac{1}{1.47} + \alpha \ln K_i + \mu \qquad (10)$$

The estimate of α is unaffected by the bias of the output measure.

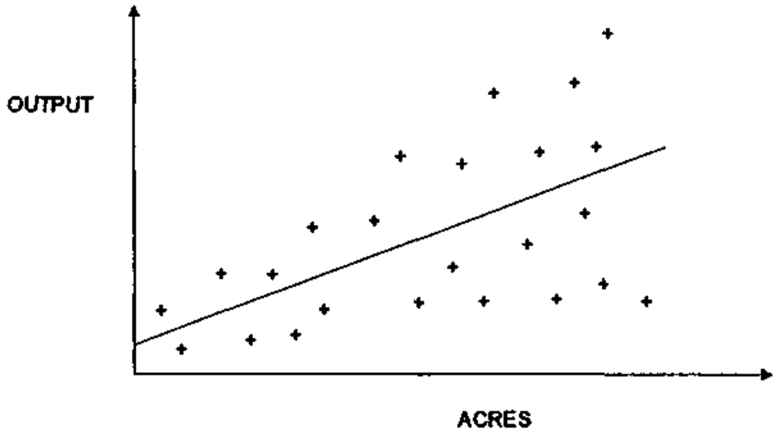
[10] With the Traditional Method the bias is induced by the over-representation of wheat in arable acreage: but the proportion of wheat in total acreage varies drastically between regions and countries. With the New Method the bias is caused by externalities between crops, which vary drastically from farm to farm. So if an unbiased estimator is required then the Adjusted Method must be used.

Let us turn to the question of the variance of the three estimates. The New Method has far higher variance than the other two Methods because there is more measurement error. This is generally undesirable because it makes it less likely that the variable will appear to be significant in a regression (even when it is truly significant). However, in this case the problem is less serious than normal because we are estimating an output series - which will generally be the *dependent* variable in a regression. It is well known that measurement errors in the dependent variable (for example, those caused by excessive variance) do not bias the estimated coefficients of the independent variables.[11]

A more serious problem is heteroskedasticity. A Spearman Rank Correlation test shows that *all* of the Methods are likely to create heteroskedasticity in a regression, as the following example demonstrates. Let us assume that all farms devote more acres to wheat than any other crop and that the degree of bias is the same on all farms (all farms grow 40 per cent wheat, 30 per cent oats and 30 per cent turnips). We have seen that when output is skewed in favour of wheat, then the total wheat equivalent output will be overestimated. But the *absolute* size of the overestimate will rise as the acreage of the farm increases. This creates heteroskedasticity in the errors and biases upwards the t-statistics in a regression. Take the case of land inputs and total output in Figure 1 below. Output rises as the size of the farm increases, but there is also a rise in the measurement errors of the output series.

[11] Kennedy P, *A Guide To Econometrics.*

Figure 1.



We ran the following regression for all three output series:

$$Q^W = C + \alpha ACREAGE + \mu \qquad (11)$$

Although the precise formulation will obviously be more detailed and feature more variables, any production function will be based on the relation between land area and output and all the other variables will be very highly correlated with land input. Hence any heteroskedasticity problem revealed by the above regression signals a general problem. The results of the Spearman Rank Correlation Tests are given in Table 5 below.

Table 5.

	Correlation between ACREAGE and μ
Traditional Method	0.71
Adjusted Method	0.69
New Method	0.63

As expected, there is evidence of heteroskedasticity in all the regressions and the problem is of roughly equal magnitude for all of the output series. Fortunately, the problem of heteroskedasticity can usually be solved fairly easily by transforming the offending variable.[12]

In terms of their bias and efficiency, none of the output measures performs noticeably better than the other two. The New Method is the most biased, but bias is often not a problem. Moreover, although the Traditional and Adjusted Methods are less biased they have the additional problem of ignoring the fallows (which introduces more random measurement error). In practice, the absence of price or output data may compel us to use the New Method. And when price and output data are available, the Adjusted Method will be employed because it enables us to assess the impact of different crops on grain yields. We demonstrate this calculation in the next section.

5. Quantifying the Externalities between Crops.

Having made the calculations required by the Adjusted Method, it is a simple matter to calculate the contribution of each crop in the rotation to the wheat yield. For example, we substituted (8) into (6) to find the beneficial effect of beans:

$$Y^{WBe} = \frac{Y^W + \frac{P^{Be}.Y^{Be}}{P^W} + \frac{P^B.Y^B}{P^W} + \frac{P^P.Y^P}{P^W} + \frac{P^O.Y^O}{P^W} + \frac{P^T.Y^T}{P^W}}{6} - \frac{P^{Be}.Y^{Be}}{P^W} \qquad (12)$$

When applied to the Arthur Young dataset, this method gives the results reproduced in Tables 6, 7 and 8 (again we have used three different sub-samples in order to see the effect of the clover problem and the fallow problem).

[12] Koutsoyiannis A, *Theory Of Econometrics* (MacMillan, 1977), p187.

Table 6.

	Mean (bu per acre)	SE of Mean (bu per acre)	SD (bu per acre)	Increase in True Wheat Yield (%)
Measured Wheat Yield	25.31	1.82	6.55	
True Wheat Yield	18.20	1.15	4.16	
Wheat Yield Due To Barley	-2.47	1.04	3.30	-13.57
Wheat Yield Due To Oats	-0.45	0.63	2.08	-2.47
Wheat Yield Due To Beans	-2.22	2.09	4.18	-12.2
Wheat Yield Due To Clover				
Wheat Yield Due To Peas	3.37	0.79	2.36	18.52
Wheat Yield Due To Turnips	9.15	1.27	4.22	50.27

The results contained in Table 6 above are the most reliable and plausible. Cereal crops tend to compete with wheat for nutrients and thereby reduce overall wheat yields - hence the effects of barley and oats are negative. By contrast, root crops tend to fix nitrogen in the soil and raise wheat yields.[13] The perverse sign on beans is very strange but the overall ranking of crops is reasonable. The new root crops such as turnips were more effective than traditional legumes such as beans and peas, so it is likely that turnips were the most beneficial crop and beans the least.[14] The superiority of turnips was caused by the different mechanism which they used to recycle nitrogen. Rather than fixing nitrogen directly from the atmosphere, the turnips were consumed by animals and the nitrogen returned to the soil in the form of animal dung. Moreover, turnips had two positive effects on wheat yields, rather than just one. Not only did they lead to an increase in nitrogen levels, they

[13] Neal E and K Neal, *Biology For Today* (Blandford, 1983).
[14] Shiel R, 'Improving soil productivity'.

were also a 'cleaning crop'. When turnips were grown, it was common to weed between the rows of vegetables; this substantially reduced the content of weeds in the following wheat crop and led to higher yields.

The ranking of crops is similar in Tables 7 and 8 below. The negative coefficient on clover is surprising. This may be an artefact of our assumptions about the price of clover - if we assumed a lower clover price then the estimated effect of clover would rise. Alternatively, it may genuinely be that clover was reducing output. The problem of 'clover-sickness' was not well-understood in the eighteenth century, and it may be the case that those farmers who adopted clover were sowing it too often. This actually *reduces* the yields of grain.

Table 7.

	Mean (bu per acre)	SE of Mean (bu per acre)	SD (bu per acre)	Increase in True Wheat Yield (%)
Measured Wheat Yield	23.03	1.31	7.29	
True Wheat Yield	17.00	0.62	3.45	
Wheat Yield Due To Barley	-1.89	0.50	2.47	-11.12
Wheat Yield Due To Oats	1.15	0.54	2.47	6.76
Wheat Yield Due To Beans	-1.08	0.95	3.41	-6.35
Wheat Yield Due To Clover	-0.80	0.49	1.90	-4.71
Wheat Yield Due To Peas	3.47	0.74	2.22	20.41
Wheat Yield Due To Turnips	9.17	0.66	3.24	53.94

Table 8.

	Mean (bu per acre)	SE of Mean (bu per acre)	SD (bu per acre)	Increase in True Wheat Yield (%)
Measured Wheat Yield	23.02	0.49	6.16	
True Wheat Yield	17.42	0.32	4.11	
Wheat Yield Due To Barley	0.47	0.32	3.37	2.69
Wheat Yield Due To Oats	1.75	0.28	3.16	10.05
Wheat Yield Due To Beans	0.70	0.62	3.80	4.02
Wheat Yield Due To Clover	-1.35	0.37	2.05	-7.75
Wheat Yield Due To Peas	2.72	0.52	3.75	15.61
Wheat Yield Due To Turnips	7.80	0.37	3.01	44.78

We can see in Table 8 that the failure to include the bare fallow in our calculations has biased upwards our estimates of the externalities of the other crops. Due to the bare fallow, the yield of grain is higher than predicted by the prices and yields of the other crops. Consequently, our calculations have misallocated the benefits of the bare fallow to the other crops - hence the impact of barley and oats appear to be positive. This demonstrates very clearly the importance of having *all* the relevant information.

6. Conclusions.

In this paper we have developed two new methods of aggregating different agricultural outputs into a single crop (such as wheat). This aggregation is a prerequisite for comparing output across production units and for estimating

agricultural production functions. The new techniques are called New Method and the Adjusted Method. Whilst both of these methods have their drawbacks, they are both superior to the Traditional Method (which simply weights all the crops by their price ratio with wheat).

The New Method is very simple and parsimonious - we only need to know the total acreage and the yield of one crop (preferably wheat). This is a great practical advantage in empirical work, where price information is scarce - usually non-existent for fallows - and agents may not be fully integrated into the local market. The New Method is nonetheless not entirely satisfactory because it is biased upwards and creates heteroskedasticity - so it is therefore only suitable for calculations performed in natural logarithms, such as production function analysis. However, we have shown that the Traditional Method is also biased and creates heteroskedasticity - so the New Method is strictly preferred to the Traditional Method.

The Adjusted Method is not inherently biased because it uses price data to adjust downwards the estimate of the wheat yield. The introduction of price data makes the calculation of total output much more precise but much more difficult. However, the Adjusted Method also reveals a great deal of new information about the contribution of different crops to the wheat yield. It is therefore to be preferred to both the New Method and the Traditional Method whenever the relevant data are available.

Employing the Adjusted Method on the dataset compiled by Arthur Young in 1770, we find that the impact of most crops in the rotation is broadly as we would expect - but the magnitudes are quite surprising. Cereal crops such as barley substantially reduced the yield of wheat (14 per cent) whereas root crops such as peas substantially increased the yield of wheat (19 per cent). The positive impact of new fallow crops seems to have been very great - turnips raised wheat yields by 50 per cent. This quantifies and supports the hypothesis that new crop technology was an important innovation leading to higher output during the Agricultural Revolution.

Appendix I.

Here we prove that the Traditional Method and the New Method are equivalent. We use the following definitions.

P^W=price of wheat
P^{Be}=price of beans
A^{Be}=acreage of beans
Y^W=yield per acre of wheat
Q^{Be}=quantity of beans
Q^{TM}=quantity of wheat by the Traditional Method
Q^{NM}=quantity of wheat by the New Method

We begin with the Traditional Method. Calculate the value of the bean crop by multiplying the quantity of beans by its price. Divide this amount of money by the price of wheat to find the quantity of wheat which is equal in value to the crop of beans.

$$\frac{P^{Be}.Q^{Be}}{P^W} = Q^{TM} \qquad \text{(1 repeated)}$$

The alternative method is to simply multiply the acreage of beans by the yield per acre of wheat to find the wheat equivalent output:

$$A^{Be}.Y^W = Q^{NM} \qquad \text{(3 repeated)}$$

We now prove that equation (1) equals equation (3). Note first that the quantity of beans is only the acreage of beans multiplied by the yield of beans per acre:

$$Q^{Be} = A^{Be}.Y^{Be} \qquad (13)$$

Substituting equation (13) back into equation (1) (the Traditional Method) gives:

$$\frac{P^B.A^{Be}.Y^{Be}}{P^W} = Q^{TM} \qquad (14)$$

Dividing equations (3) and (14) by $A^{Be}.Y^{Be}$ gives:

$$\frac{P^{Be}}{P^W} = \frac{Q^{TM}}{A^{Be}.Y^{Be}} \qquad (15)$$

and,

$$\frac{Y^W}{Y^{Be}} = \frac{Q^{NM}}{A^{Be}.Y^{Be}} \qquad (16)$$

But P^{Be}/ P^W is the ratio of prices, and Y^W/Y^{Be} is the Marginal Rate of Transformation (that is, the rate at which we can transform wheat into beans by switching resources from one crop to the other). The neo-classical production condition requires that the Marginal Rate of Transformation always equals the price ratio in equilibrium:[15]

Lemma 1.

$$\frac{Y^W}{Y^{Be}} = \frac{P^{Be}}{P^W} \qquad \text{(2 repeated)}$$

[15] Layard R and A Walters, *Microeconomic Theory* (McGraw-Hill, 1978).

UNIVERSITY OF OXFORD

Discussion Papers in Economic and Social History

1 Joachim Voth and Tim Leunig, *Did Smallpox Reduce Height? Stature and the Standard of Living in London, 1770-1873*

2 Liam Brunt, *Water into Wine - New Methods of Calculating Farm Output and New Insights into Rising Crop Yields during the Agricultural Revolution*