Statistical Examination of William Gillies' Still Life Paintings

1. Introduction

The analyses reported here were prompted by Andrew F. McPherson who has made a study of the Scottish painter William Gillies, 1898–1973.

Gillies is known to have created 308 still life paintings during the period 1925 to 1972. Following direct inspection or reference to documents, McPherson created descriptions for 206 of these paintings including categorisation of the objects contained therein. More specifically, he has adjudged some objects, such as a black vase, a poppy or of pottery made by his sister Emma, to be memorial objects.

We have made use of the numerical data generated on those 206 'observed' paintings and the dates assigned to the 308 known still life paintings.

The main thrust of the analyses has been to follow-up on conjectures made by McPherson on the number (termed 'intensity') of still life paintings recorded across the period 1925 to 1972. He paid special attention to Gillies' inclusion in the 206 'observed' still life paintings of memorial objects ('memorialisation'). Both 'intensity' and 'memorialisation' were conjectured to have relationship to events during the period.

In the statistical modelling reported here, Intensity and Memorialisation are regarded as the dependent variables susceptible to effect by time and events. In general, statistical modelling assumes that it is variation in the dependent variable which includes some element of random error which is modelled in relation to a range of further variables including the date ascribed to the painting. In this problem, however, there is also acknowledged to be variability (error) in this date. The exact date for each painting is not known, although it is understood that Gillies could decide to paint a still life and often finish it quickly, sometimes doing both on the same day. Based on his study of available documentation, McPherson derived three estimates for the Year of the painting. His 'best estimate' is generally used, supplemented by an 'earliest date' and a 'latest date'.[i]

Two approaches were deployed when examining the data. One was to use LOWESS, a 'locally weighted regression' smoothing technique.[ii] LOWESS provided useful insights into the variability over time for both the intensity of William Gillies' painting and for his use of memorial objects. More specifically, the LOWESS plots displayed strong indication of non-random variation for each of 'intensity' and 'memorialisation'. However, there was no such indication in the LOWESS plots of any relationship between the two.

The second approach was to set up several hypotheses for 'exploratory testing'. Caution is required as the conjectures are being tested on the same data from which they were generated. However, this can be regarded as a strong test when conjectures from observation of supposed relationships subsequently fail a statistical test.[iii] The method used was polynomial logit regression, with differential weighting for dates and backward elimination of any terms found to be insignificant.

The outcome of testing suggested by the LOWESS smoothing is as follows:

- All of the null hypotheses relating to intensity over time were rejected at the 0.05 significance level. There is evidence of non-random structure which lends support for all hypothesised relationships.

- None of the null hypotheses relating to memorialisation over time were rejected at the 0.05 significance level. There is no statistical support for these relationships.
- Similarly, there is no statistical evidence for any such supposed relationship between memorialisation and intensity; none of the null hypotheses were rejected at the 0.05 significance level.

Opportunity was also taken to test conjectures previously made by McPherson, based on his prior expert study of the paintings:

- His conjectures about the existence of a relationship of painting intensity across time, including a step change at the year 1955, all gained support; the null hypothesis that there was no relationship across time was rejected using polynomial logit regression.
- However, as stated, none of the conjectures concerning use of memorial objects over time were supported in the tests using polynomial regression.
- Nor was there any evidence from statistical testing for any systematic relationship between the use of memorial object and painting intensity.

2. **Examining Data**

The main purpose in using LOWESS was exploratory, enabling visual examination of possible relationships; the technique involves re-estimating the value of each data point using information taken from nearby data points. LOWESS has the added advantage of smoothing across the imprecision in the assigned date of each painting.

Two sets of analysis are presented using LOWESS smoothing:

- The first is of 'intensity', that is of the number of known still life paintings attributed to a given year. There are 308 data points in that analysis, corresponding to the total number of known still life paintings.
- The second addresses the use of memorial objects, 'memorialisation'. That analysis has 206 data points, corresponding to the paintings with assessed content.

1.1 Intensity

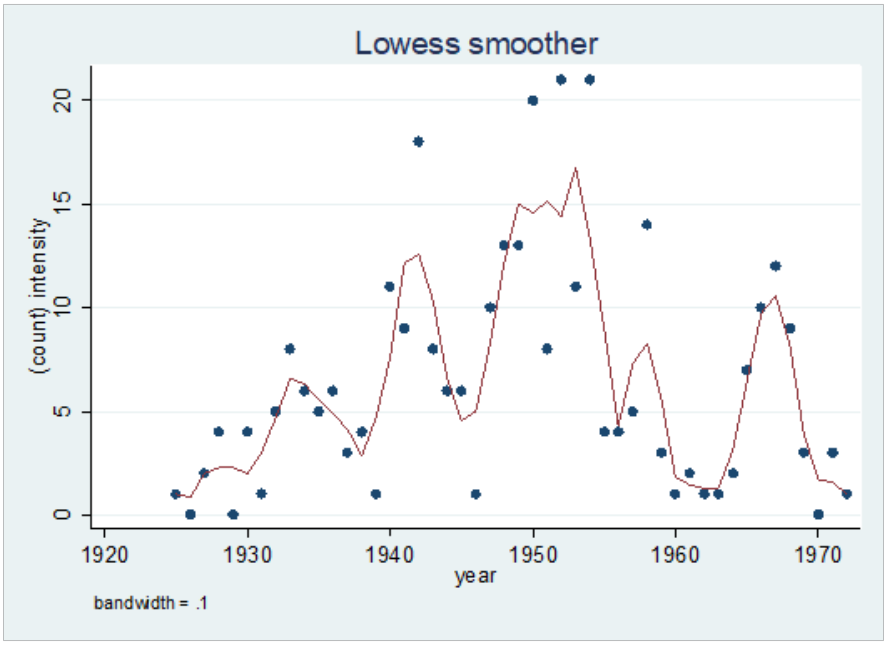
William Gillies is known to have created about 4 to 6 still life paintings per year; the average using the median as the measure was 5. The yearly totals are presented in the Appendix.

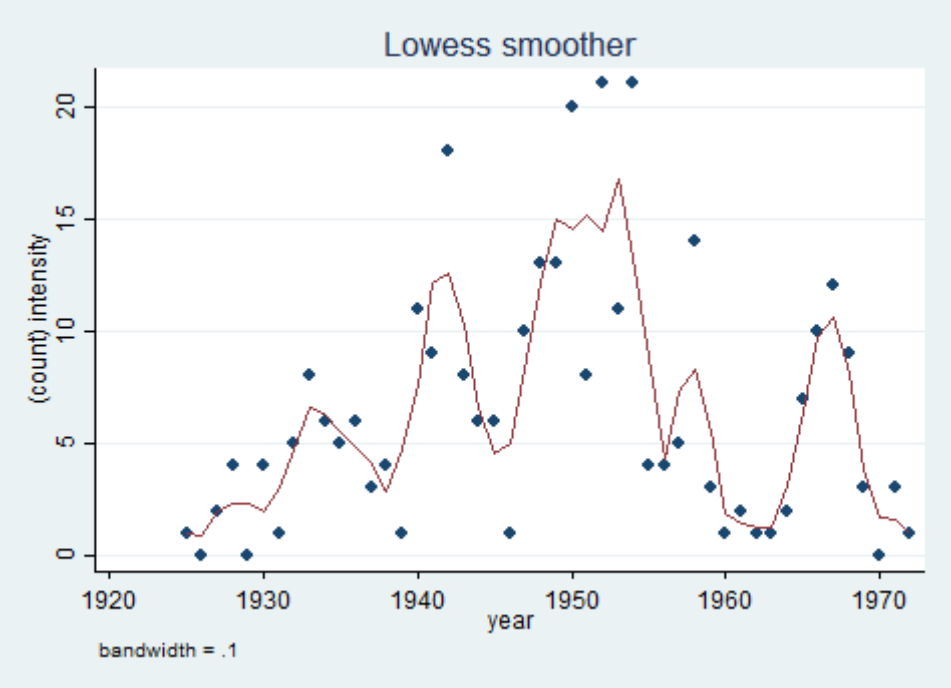
There was considerable fluctuation in the yearly number of paintings of still life across the 48 years from 1925 to 1972. The number of still life paintings in any given year varied widely, sometimes none in a year and sometimes as many as 21.

The first graph below shows the smoothed line fitted by LOWESS using a low bandwidth setting (0.1). This displays as several peaks and troughs.

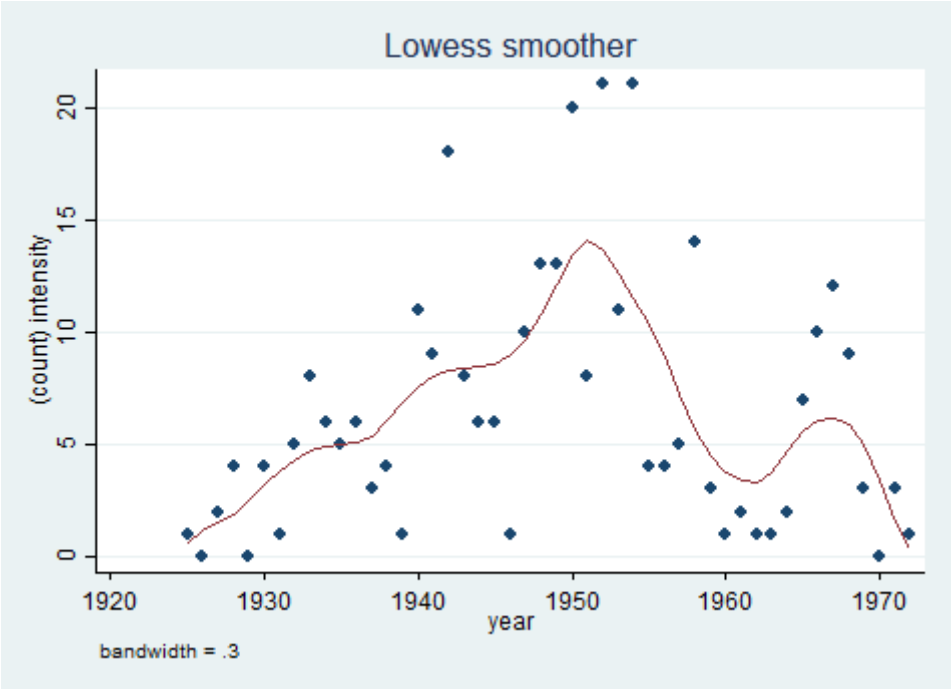
Visual examination suggests five periods of high intensity of painting. These begin with increasing activity during the period up until the 1950s, with three phases. The first phase peaked in c.1934, declining until 1939, the start of WW2. After a short period of less activity, there was a second phase of intensity during the War which peaked in c.1942 (with a fitted value of 12 still life paintings) before returning to the median average of 5 still life paintings in 1947. The third and largest peak was in the first part of the 1950s. There is the suggestion of a minor (fourth) peak around 1958, in what was otherwise a low level of activity during the early 1960s, before a surge in and around 1967. In overall terms there is a strong

suggestion of an increase from the earliest period (circa 1925) to a peak at around 1952 followed by a decline thereafter, albeit with the variations mentioned above.



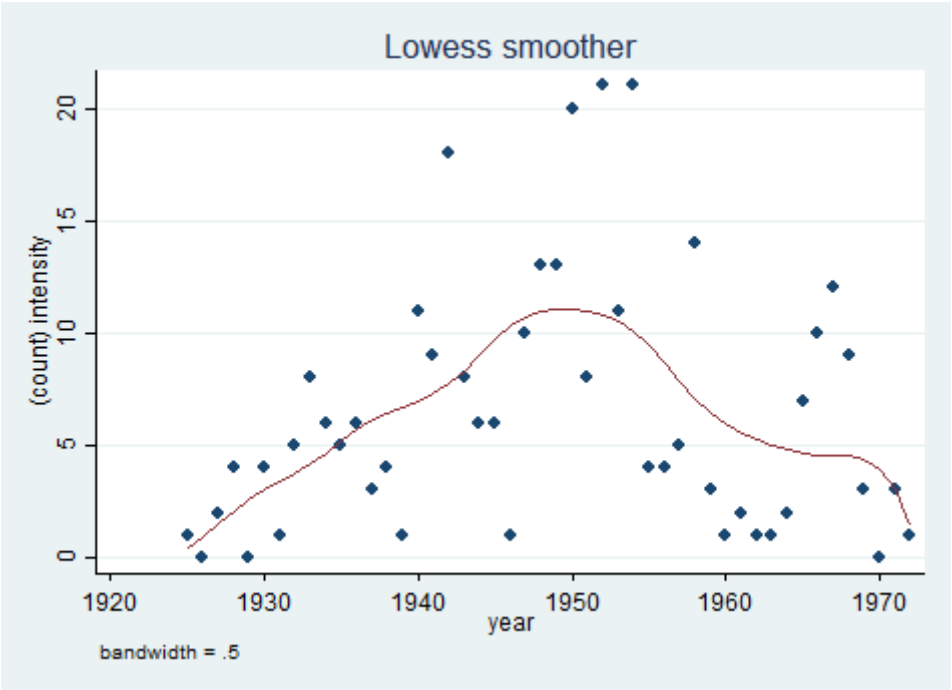


Adjusting the bandwidth to 0.3 in LOWESS produces a greater degree of smoothing, the resulting line less susceptible to the outlying number of paintings per year.



Visually, this fitted line brings out the steady rise in painting activity from 1925 to a peak at c.1952. Intensity then declines with only a second, much lower, peak in the later 1960s.

Setting a higher bandwidth (0.5) for LOWESS increases smoothing even further. The result is unimodal: a steady increase in intensity peaks c.1952, then declines almost as steadily.

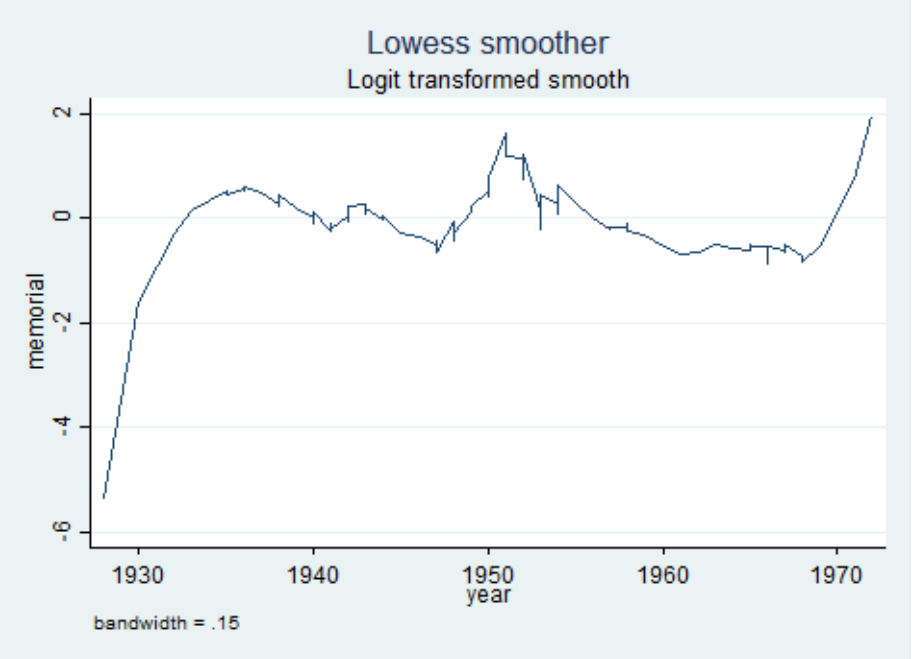


Interest after curve fitting is then in the outliers, especially those in the early 1950s and the late 1960s. Some of the explanation for those outliers might lie with the imprecision in the dating of paintings although this is partly dealt with by the smoothing process.

1.2 Memorial Objects

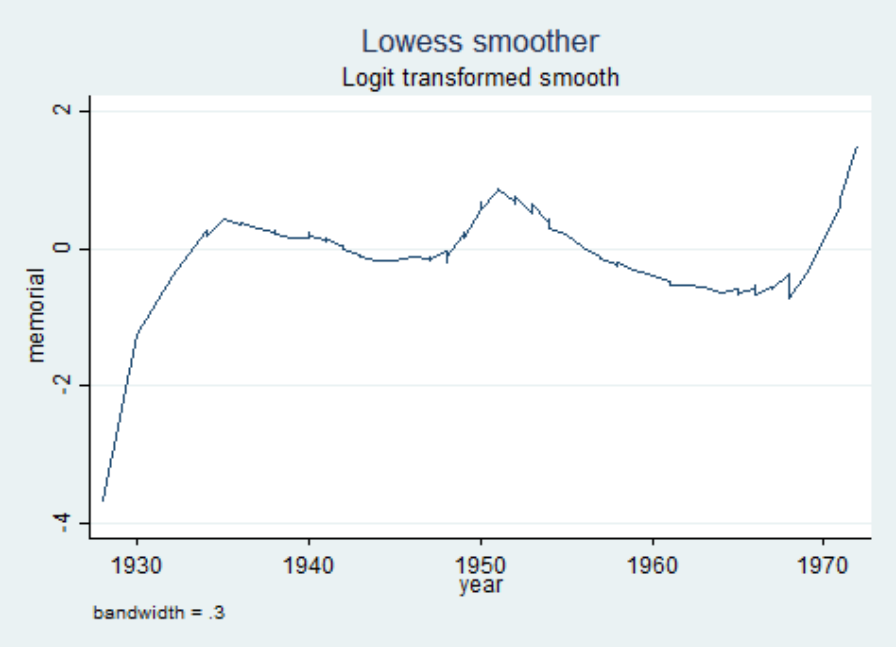
There are 206 data points in this analysis, one for each of the paintings for which there is knowledge of its content. The indicator for the absence/presence (0, 1) of memorial objects assigned to each painting is transformed using the logit link.[iv]

Visual examination of the first LOWESS plot, again using a low bandwidth setting (0.15) to expose local variability, suggests three, possibly four, peaks in the use of memorial objects.



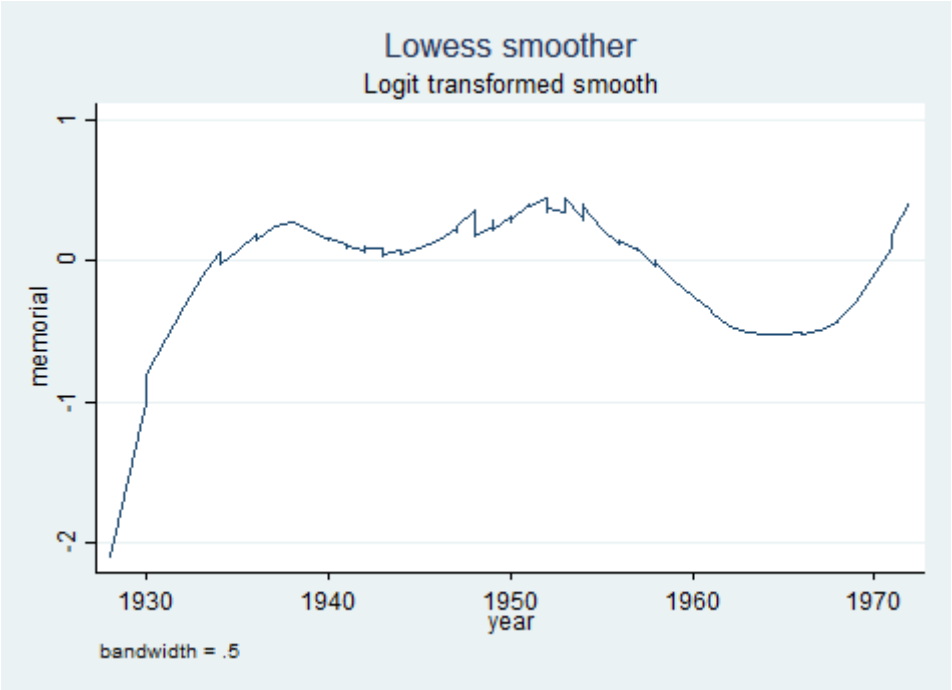
The first peak, at c.1936, comes after an initial period of rapid increase, followed by decline to c.1947, with slight inflection around 1942 to 1943. There is then a sustained increase from c.1948, reaching a peak in c.1952. That is followed by sudden decline and then gradual reduction until a final rapid increase, post 1968, to 1972 and William Gillies' death.

The next plot uses a higher bandwidth setting (0.3) with the same data.



This results in additional smoothing although the two peaks, at around c.1936 and c.1952, remain. Each peak is followed by a period of decline, the second more pronounced. The rise in the use of memorial objects in Gillies' last years remains evident.

Further smoothing, to give greater emphasis to a general underlying trend, is achieved by using the larger bandwidth setting (0.5). The peaks at around 1935 and 1952 are much less evident, with the decline setting in c.1954. Rapid growth in usage around each of the two end periods, prior to 1937 and after 1968, is clearly visible.

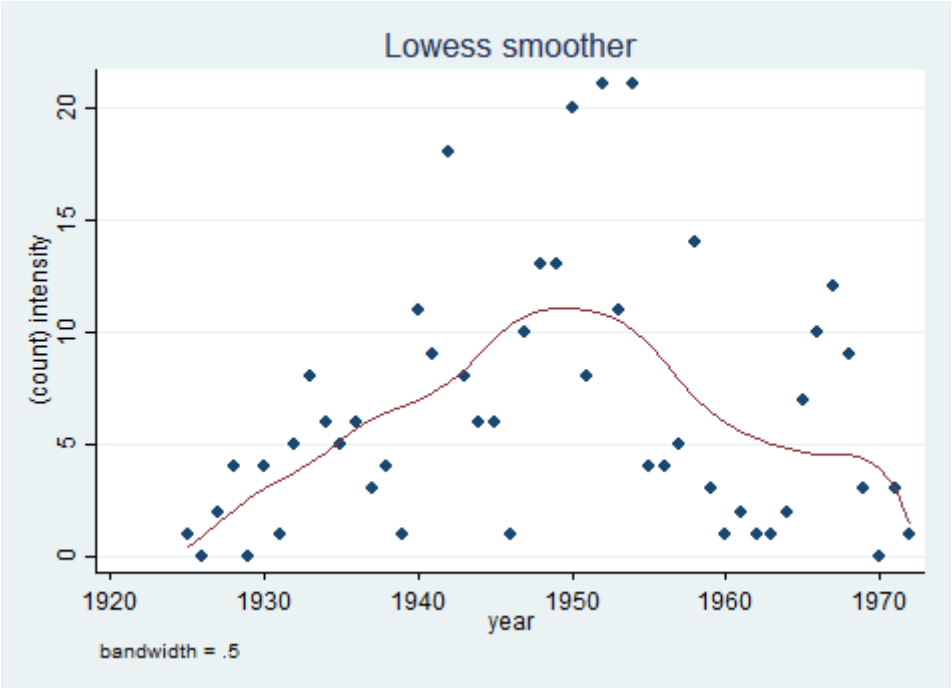


2. Testing of exploratory hypotheses

The hypotheses we set up here are motivated by our examination of the locally fitted LOWESS plots shown above, particularly those using the higher bandwidth setting (0.5). Of course, we are aware that deriving hypotheses and then testing them on the same data risks falsely supposing statistical significance. However, the converse, namely the failure to record statistical significance, can be interpreted as a firm rejection of favoured hypotheses: the supposed relationships judged therefore to be a consequence of chance factors.

The series of these 'exploratory hypotheses' was examined using polynomial regression. The models take account of the imprecision in the dates of paintings by using a weighted procedure: essentially this applies less weight to the paintings for which dating is assessed as less reliable.[v]

3.1 Painting activity (Intensity) across time



Based on the inspection of the LOWESS plot shown above, we formulated the following six 'exploratory hypotheses' for testing:[vi]

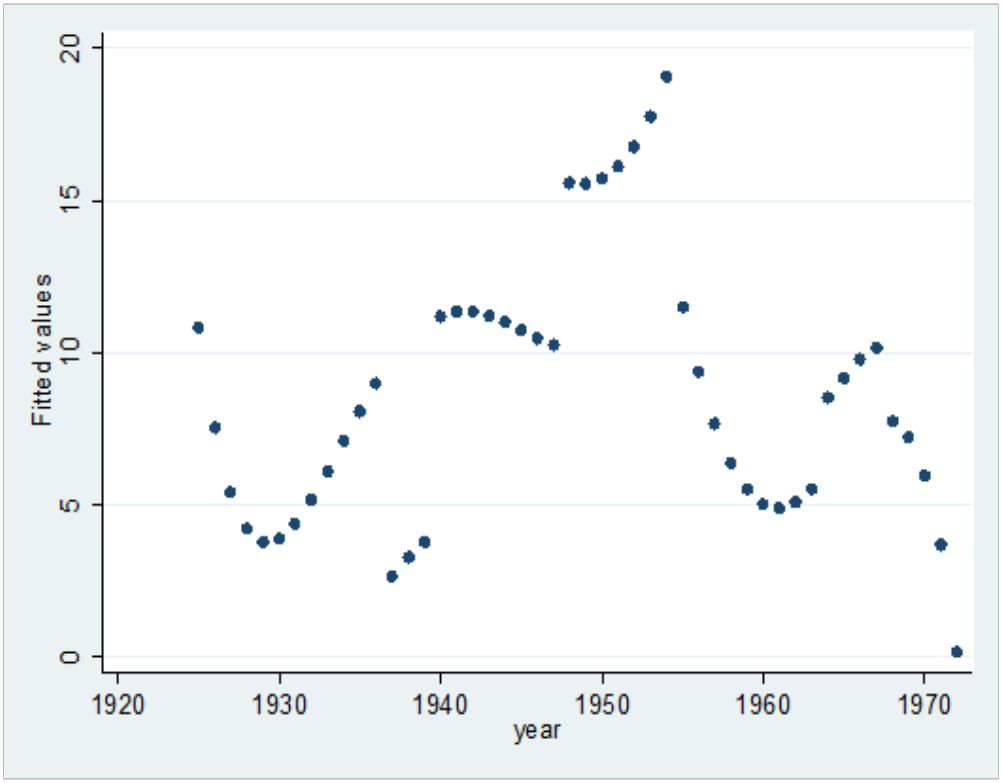
- H1 Intensity increased over the period from 1925 to 1950
- H2 Intensity decreased over the period from 1951 to 1972
- H3 Strength of relationship with time was less in the second period than in the first
- H4 Intensity increased in period 1948 to 1954 (inc.) compared with all other years
- H5 Intensity reduced in period 1937 to 1939 (inc.) following Emma's death
- H6 Intensity increased in period 1964 to 1967 (inc.), when compared with all other years

We next turn to test one hypothesis which was posited by McPherson prior to our analysis of the data. This is the closest we come to a 'confirmatory test', even given that he had become very familiar with the data he himself had collected and examined. This hypothesis can be regarded as having two parts:

- H7 (a) The level of Intensity of Gillies' painting was non-random across time
- H7 (b) The relationship of Intensity to time changed in 1955

Testing was carried out using backward elimination, having started with a regression model intended to mimic the data as closely as possible within the polynomial constraints:[vii] that included polynomial coefficients up to the degree 5 for the date of the paintings. The result was that the quintic (degree 5) term remained significant, as did all other lower order terms excepting the cubic term, which was included, nevertheless.

Shown below is the plot of the fitted values from the fully saturated model of that polynomial regression, including both the change in intercept term (H7a) and the change in slope term (H7b).



Each of the hypotheses relating to intensity of painting activity, set out as H1 to H6, was found to be statistically significant, as shown below:

Hypothesis	Coefficient	Standard Error	t value	Test outcome
H1	0.58	0.05	12.34	Significant
H2	-0.65	0.09	-7.60	Significant
H3	-1.22	0.10	-12.58	Significant
H4	5.46	1.69	3.23	Significant
H5	-7.12	0.90	-7.87	Significant
H6	2.45	1.12	2.18	Significant

In particular, the plot clearly illustrates how the intensity of painting reduced during the years 1937 to 1939 following Emma's death (H5); this is reflected in the size of the coefficient (-7.12) expressing that negative effect. This is large relative to the estimated standard error (0.90) and regarded as statistically significant, with sizeable t value of -7.87. The large (5.50) positive increase in the immediate post WW2 years is also very evident.

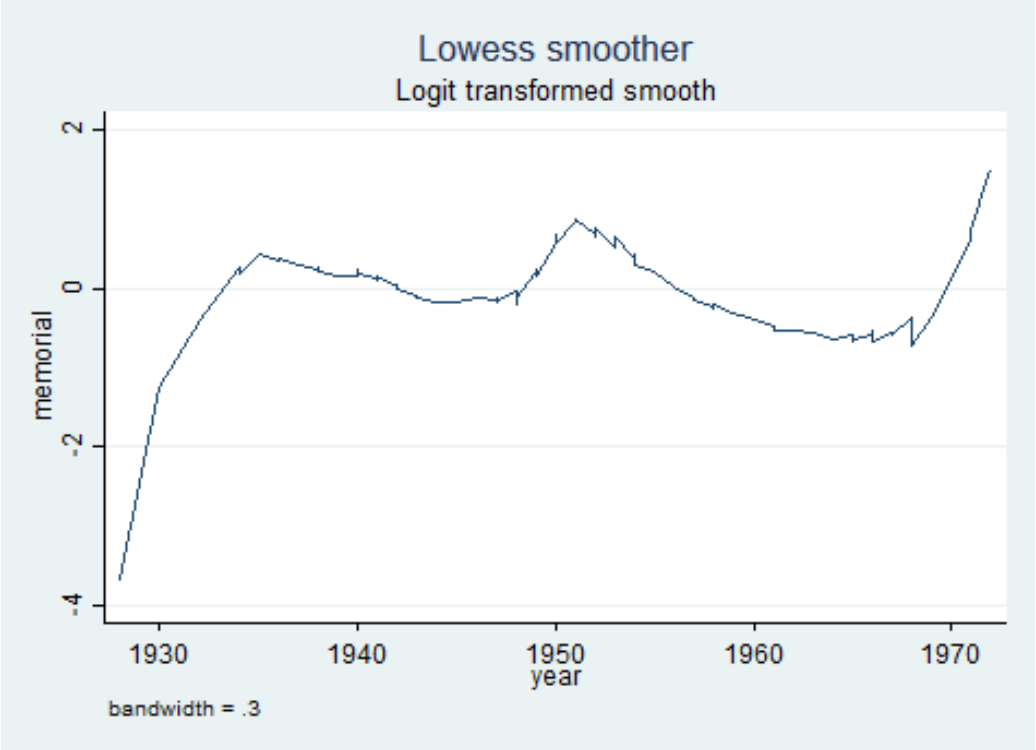
The outcome of testing hypotheses relating to painting intensity after 1955 is as follows.

Hypothesis	Term	Coefficient	Standard Error	t value	Test
H7a	Change in fitted value at 1955	-4.28	1.46	-2.94	Significant
H7b	Pivot: change in slope at 1955	-3.92	3.87	-1.01	Not significant

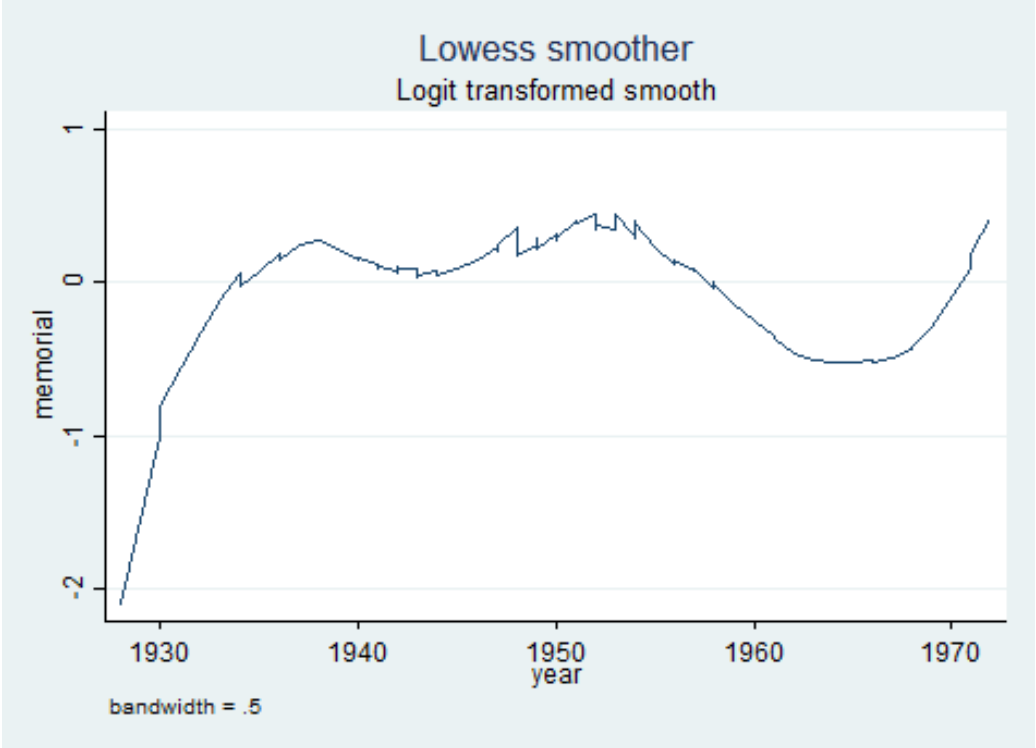
The sizeable reduction in the fitted values post-1955, corresponding to the large (-4.28) estimated coefficient, is found to be significant (H7a). However, that was not true for the coefficient of the pivot indicator (-3.92) representing the change apparent in the slope; this was because there was an almost equally large estimated standard error.

3.2 Use of Memorial Objects across time

The LOWESS smoother displays a second prominent peak after 1948 with bandwidth set to 0.3.



This is flattened but not eliminated with Increased bandwidth (0.5).



Based on inspection of the LOWESS plot above we formulated the following hypotheses for Gillies' use of memorial objects.[viiiix]

H8a An increase during 1948 to 1954 (inc.) compared with all other years

H8b An increase during 1949 to 1953 (inc.) compared with all other years[x]

H9 Decreased use during the years 1954 to 1962 (inc.)

H10 Reduction following Emma's death during 1937 to 39, compared with all other years

The absence/presence (0, 1) of memorial objects is transformed using the logit link. This provides a dependent variable measured on an interval scale which can be regressed on explanatory variables with error terms assessable as to the normality assumptions.

The outcome of the polynomial regression is as follows:

Hypothesis	coefficient	standard error	t value	Test outcome
H8a	0.52	0.54	0.95	Not significant
H8b	0.85	0.45	1.89	Significant[xi]
H9	-0.05	0.15	-0.32	Not significant
H10	1.06	1.17	0.91	Not significant

There is statistical support for one of the supposed relationships, namely the increase in the use of memorial objects between 1949 and 1953 inclusive (H8b). This one directional hypothesis is statistically significant at the 0.05 level using a one tailed test.

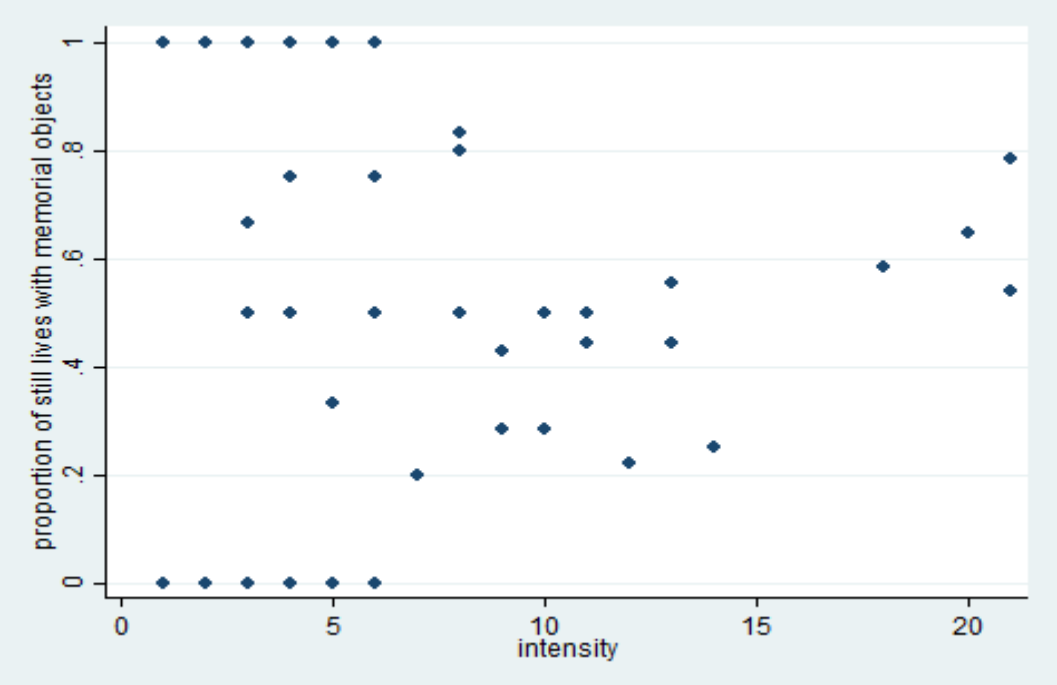
The observed increase in the use of memorial objects after 1948 to 1954 (H8a) can be attributed to chance factors, as can that observed decrease after 1954 (H9), albeit that the power of the test of H9 is limited by the small sample size.

There is no statistical evidence for a change in memorialisation in the years immediately following Emma's death (H10).

3.3 Use of Memorial Objects and Intensity of still life painting: examining the relationship.

Finally, we investigated the possible existence of a relationship between the use of memorial objects in paintings and the intensity of painting.

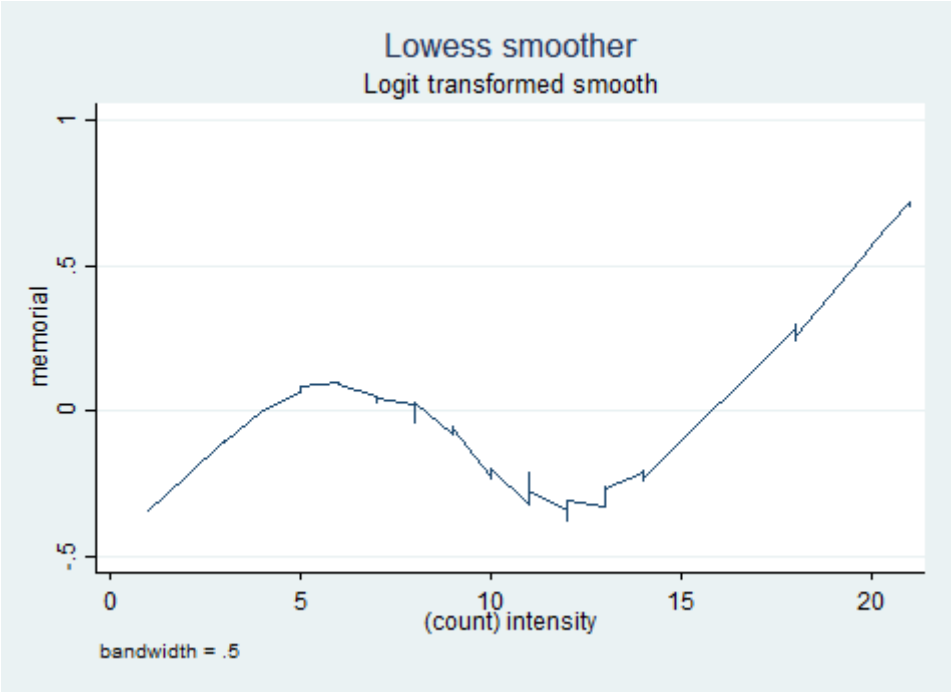
The base data is set out as yearly aggregates in tabular form in an Appendix. These are used in the scatterplot displayed below. This shows the proportion of paintings in a given year found to have one or more memorial objects plotted (in the vertical y axis) against the number of still life paintings in that year.[xii]



There is no evident relationship: the estimated correlation between the use of memorial objects in paintings and the intensity of painting of 0.123 is regarded as insignificant, having standard error of 0.438.

The appropriate approach, as was done in the earlier sections, is to conduct analysis at the painting level. Here we use the 206 paintings for which there was information gleaned about use of memorial objects. Information about intensity (that is, the number of paintings is a given year) is then deployed as a variable assigned to each painting according to that year.

Prior to our analyses, suggested by his close study of Gillies' still life paintings, McPherson had conjectured that there was a relationship between memorialisation and intensity. However, the LOWESS plot of memorial objects and intensity using all the 206 'observed paintings', displayed below using the bandwidth set at 0.5, shows that there is not a simple linear relationship.



The next stage was to formulate four 'exploratory hypotheses':

H11 That an overall relationship exists
H12 That relationship exists during the period 1948 and 1954 (inc.)

Two null hypotheses due to McPherson
H13 That there is a change in the relationship at year 1955
H14 The there is a pivot change in the intercept at year 1955

Once again, the absence/presence (0, 1) of memorial objects is transformed using the logit link to provide a dependent variable measured on an interval scale which can be regressed on explanatory variables with error terms assessable according to the normality assumptions.

The four exploratory hypotheses relating to a possible relationship between Gillies's use of memorial objects (memorialisation) in his still life paintings and the annual intensity of his painting were then tested using indicator variables in an analysis by polynomial regression:

Hypothesis	coefficient	standard error	t value	test outcome
H11	0.04	0.02	1.46	Not significant
H12	0.02	0.05	0.49	Not significant
H13	-0.09	0.07	-1.23	Not significant
H14	0.05	0.73	0.06	Not significant

None of these hypotheses were supported by the analysis of the 'observed painting' data.[xiii]

Table 1: Known Still Life Paintings and Occurrence of Memorial Objects, by Year

(1) Year	(2) Known Still Life Paintings	(3) Proportion Observed	(4) Paintings Observed	(5) 1+ Memorial Objects	(6) Total Mem. Objects
1925	1	0.00	0	0	0
1926	0	-	-	-	-
1927	2	0.00	0	0	0
1928	4	0.50	2	0	0
1929	0	-	-	-	-
1930	4	0.50	2	0	0
1931	1	0.00	0	0	0
1932	5	0.20	1	0	0
1933	8	0.63	5	4	4
1934	6	0.67	4	3	4
1935	5	0.60	3	1	1
1936	6	0.67	4	2	2
1937	3	0.33	1	1	2
1938	4	1.00	4	3	3
1939	1	1.00	1	1	1
1940	11	0.73	8	4	4
1941	9	0.78	7	2	2
1942	18	0.67	12	7	11
1943	8	0.75	6	3	3
1944	6	0.33	2	2	4
1945	6	0.17	1	0	0
1946	1	1.00	1	0	0
1947	10	0.70	7	2	3
1948	13	0.69	9	4	4
1949	13	0.69	9	5	10
1950	20	0.85	17	11	20
1951	8	0.75	6	5	7
1952	21	0.67	14	11	15
1953	11	0.82	9	4	8
1954	21	0.57	13	7	11
1955	4	0.25	1	1	2
1956	4	0.25	2	1	2
1957	5	0.40	2	2	4
1958	14	0.57	8	2	2
1959	3	0.67	2	1	1
1960	1	1.00	1	0	0
1961	2	1.00	2	2	4
1962	1	1.00	1	0	0
1963	1	1.00	1	0	0
1964	2	0.50	1	0	0
1965	7	0.71	5	1	2
1966	10	1.00	10	5	7
1967	12	0.75	9	2	2
1968	9	0.78	7	3	4
1969	3	0.67	2	0	0
1970	0	-	-	-	-
1971	3	1.00	3	2	3
1972	1	1.00	1	1	1
	308		206	105	

End Notes

[i] Use of dates as whole years inevitably still adds a degree of error: two different paintings which might have been done within weeks of one another, during December and January (say) but might been assigned to different years.

[ii] This approach was first proposed by William Cleveland as early as 1979. Initially regarded as too computationally intensive, this is now being adopted more widely, largely for exploratory purposes. See Cleveland_JASA_1979.pdf (washington.edu). A more readable version for layperson is obtained in https://www.itl.nist.gov/div898/handbook/pmd/section1/pmd144.htm

[iii] This is couched in term of rejection or otherwise of a null hypothesis (classical statistical procedure). Hypotheses in main text are worded in the reverse fashion. Hypotheses H1-H6, H8-H10 are one directional, the remainder being two directional.

[iv] "The logit transformation of the binary measure of the absence/presence (0, 1) of memorial object(s) in a work creates a dependent variable measured on an interval scale which can be regressed on explanatory variables with error terms that can be assessed against the assumption of the normal distribution." Technically logit(p)=log(p/(1-p)).

[v] Weights are a threefold categorisation of a variable constructed from the difference between the latest assessed date and the earliest assessed date. A weight of 1 was given to paintings for which the difference in in attributed years between 'first;' and 'last' dates was three or more, a weight of 2 for a difference of two and a weight of 3 when the difference was 1 or less. All weighting was standardised in the analysis (default) to result in correct standard errors.

[vi] The models deployed for tests H1-H3 included year as a linear term only, without the pivot in 1955, but it did include the statistically significant dummy variables corresponding to H4-H6.

[vii] The changes in intercept and slope found when testing H7a and H7b were represented in the models used to test hypotheses H4 to H6 this being consistent with the aim is to fit the underlying relationship to year to as close an extent as possible without running into any issues with overfitting.

[viii] As noted earlier in the main text, tests for Hypotheses H8a, H10 were carried out by polynomial regression using backward elimination, having started with a model with year at degree 6, i.e., not just linear, quadratic, cubic terms etc. and finally fitting a model of degree 5 in year. The indicator for a pivot at 1955 was found to be non-significant for memorial objects (not separately reported) and so was eliminated and the regression rerun with only dummy variables relevant to Hypotheses H8a, H10 included. The coefficients for two dummy variables (corresponding to Emma's death, post-war) were then re-assessed within that model. With both included, none were significant and the model was then rerun with each of the (Emma's death, post-war) dummy variables introduced singly. Neither of the dummy variables was found to be statistically significant and so reporting was for the model with both included. For H9 the model included, simply, year of painting as a linear term only.

[ix] As for the analyses of intensity an iterative approach to the modelling was adopted. Firstly models were fitted without the dummy variables corresponding the two hypotheses under test and a test was made for the presence of a pivot at 1955. No evidence was found for this. Then a test of the degree of the polynomial in year was assessed and terms up to and including degree 5 were found to be required. Then dummy variables corresponding to hypotheses H8a and H10 were fitted, neither being found statistically significant. The final models for H8a and H10 fitted each dummy variable on their own. The final model for H8b used that of H8a, replacing the dummy variable with the new appropriate one.

[x] The two possible ranges of years were chosen due to the theoretical uncertainty of the supposed effects of trauma following the World War 2 period.

[xi] When the hypothesis is directional, not just any change but a change in a particular direction, it is appropriate to use a one tailed test.

[xii] The proportion is computed by dividing column 5 ('with 1+ Memorial Objects') in Table 1 by column 4 ('Paintings Observed').

[xiii] Note that the use of weighting (to reflect the imprecision in the dates of paintings), assessed on the testing of H11, has little effect on the outcome of these tests: this was tested on the test of H11 in which the unweighted modelling produced a coefficient of 0.030 with standard error 0.024, almost identical to the weighted analysis.