

UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2012

**Change Points and Temporal Dependence in
Reconstructions of Annual Temperature: Did
Europe Experience a Little Ice Age?**

Morgan Kelly and Cormac Ó Gráda, University College Dublin

WP12/10

March 2012

**UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD DUBLIN 4**

Change Points and Temporal Dependence in Reconstructions of Annual Temperature: Did Europe Experience a Little Ice Age?

Morgan Kelly and Cormac Ó Gráda*

March 30, 2012

Abstract

We analyze the timing and extent of northern European temperature falls during the Little Ice Age, using standard temperature reconstructions. However, we can find little evidence of long swings or structural breaks in European weather before the twentieth century. Instead, European weather between the fifteenth and nineteenth centuries resembles uncorrelated draws from a distribution with a constant mean (although there are decades of markedly lower summer temperature); with the same behaviour holding more tentatively back to the twelfth century. Our results suggest that the existing consensus about a Little Ice Age in Europe may stem from a Slutsky effect, where the standard climatological practice of smoothing data before analysis gives the spurious appearance of irregular oscillations.

*School of Economics, University College Dublin. This research was undertaken as part of the HI-POD (Historical Patterns of Development and Underdevelopment: Origins and Persistence of the Great Divergence) Project supported by the European Commission's 7th Framework Programme for Research.

A central pillar of popular scepticism about anthropogenic global warming is the Little Ice Age: if temperatures over the last few centuries could fall sufficiently for Swiss glaciers to advance, the Thames in London to freeze, and the Norse colonies in Greenland to disappear, then surely the current rise in temperatures is part of the same natural cyclicity? The goal of this paper is to estimate the magnitude and timing of climatic deteriorations during the Little Ice Age in Europe by using a variety of standard summer and winter temperature reconstructions: for Central Europe since 1500 (Dobrovolný et al., 2009); the Netherlands from 1301 (van Engelen, Buisman and IJnsen, 2001); Switzerland from 1525 (Pfister, 1992); and England from 1660 (Manley, 1974).

The consensus among climatologists is that the Northern Hemisphere above the tropics experienced sustained episodes of reduced temperatures between the fifteenth and nineteenth centuries, with particularly marked falls in Europe (Mann 2002, Matthews and Briffa 2005, Mann et al. 2009). However, we find little evidence of change points or temporal dependence in the series we examine, contrary to the existing consensus of a Little Ice Age.

Starting with standard classical tests for change points in mean temperature—Bai and Perron (1998) breakpoints and Venkatraman and Olshen (2007) binary segmentation, which we show can detect changes of one standard deviation lasting a generation, and 0.5 standard deviations lasting a century—we find that all winter series break around 1900, as does Central European summer temperature. However, there is no indication of any shift in mean temperature, apart from a brief fall in Switzerland during the 1810s, prior to this. Using the Bayesian change point analysis of Barry and Hartigan (1993), which has greater ability to detect short deviations, we find that while winters before 1900 are stable, there are occasional decades of markedly reduced summer temperature (the 1590s in Central Europe, the 1690s in England, and the 1810s everywhere), but again no sustained falls in mean temperature of the sort one would expect during a Little Ice Age.

Looking at temporal dependence, we find no sign of trends in the data before the twentieth century, and little evidence of first order autocorrelation in summer temperature, although Central European and Swiss winters display modest mean reversion, with autocorrelation coefficients of the order -0.1 to -0.2 .

The fact that European temperature series do not for the most part show significant or large first order autocorrelation leaves open the possibility that they may exhibit higher order or non-linear dependence. We therefore test whether the data show conditional mean independence: given the past history of a stationary series $\{Y_t\}$ with expectation μ , the best forecast of its current value is its unconditional mean: $E(Y_t|Y_{t-1}, Y_{t-2}, \dots) = \mu$.

There are three principal categories of tests: portmanteau tests that look at the sum of autoregressive coefficients in the data; variance ratio tests that look at how the variance of a series grows as the number of observations rises; and spectral tests that look for departures from a straight line spectrum. We apply recent versions of each to temperature reconstructions, and find few departures from conditional mean independence, with these driven by pre-1700 observations. We find similar behaviour for English and Swiss precipitation reconstructions. The weak correlation of annual Northern European temperature series is in marked contrast to the strong autocorrelation in the CRU Northern Hemisphere temperature series since 1850 which has first order autocorrelation 0.6 and significant partial autocorrelations out to lag four (McShane and Wyner, 2011), highlighting once again the importance of spatial variation in climatic patterns.

Before the start of instrumental records around 1700, the reconstructions used here are all based on documentary sources such as weather diaries and ships logs for the sixteenth and seventeenth centuries; and harvest dates, and records of when river tolls could not be collected because of drought or freezing temperatures for earlier periods. (While tree rings might seem to be a more obvious proxy, they are only reliable records of weather in cold or arid

places where trees are under continual climatic stress, something that is not the case in most of Europe). This raises an obvious question of how reliable are these subjective reconstructions, particularly for earlier centuries. We examine the validity of the Dutch reconstruction by seeing how it explains extensive English records of wheat yields and prices from the thirteenth to the fifteenth centuries, and find that it performs well, suggesting that it is a reliable approximation to historical temperature conditions.

While our series, apart from Dutch temperature from 1301, are for weather after 1500, we also have more tentative estimates of German weather back to 1000 AD by [Glaser and Riemann \(2009\)](#) who assign seasons to three categories (good, average, or bad). We find that probabilities of a good or bad summer or winter are relatively unchanged between the twelfth and nineteenth centuries, as are the probabilities of successive good or bad seasons. Tests of conditional mean independence are problematic because they are not designed for such multinomial data, but simulations show that the automatic portmanteau test of [Escanciano and Lobato \(2009a\)](#) performs well. This test does not reject conditional mean independence for winter weather between AD 1000 and 1500, while showing weak dependence in summer weather.

To see if the volatility of annual temperatures rose during the Little Ice Age we looked for change points in squared temperature series, but only detect a change in one: the volatility of Central European summer temperatures falls markedly in the 1720s, around the time that reliable instrumental records start with Fahrenheit's invention of the modern mercury thermometer. We apply a variety of tests for generalized autoregressive conditional heteroskedasticity (GARCH) but find that, with the exception of Central European summers after 1728 which shows modest persistence in volatility, the variance of temperature appears constant. Looking at Intra-Class Correlation (ICC) we find that variance between decades is small compared with variance within decades: variance between decades typically accounts

for roughly 4–10 per cent of total variance for winter temperature, and 1–2 per cent for summer temperature.

In summary, then, annual temperature reconstructions for northern Europe do not appear to exhibit temporal dependence or structural breaks consistent with the occurrence of a Little Ice Age. Naturally, our results have nothing to say about the occurrence of a Little Ice Age in other parts of the Northern Hemisphere.

That our findings run counter to the existing consensus of a European Little Ice Age may reflect the fact that our analysis is based on unsmoothed data. This is in contrast to the current practice in climatology of smoothing data using a moving average or other filter prior to plotting it. When data are uncorrelated, as annual European weather series appear to be, smoothing can introduce the appearance of irregular oscillations: a Slutsky effect.¹ The intuition behind the Slutsky effect is straightforward: just as tossing a fair coin leads to long sequences with an excess of heads or tails, so random sequences in general will occasionally throw up some unusually high or low values in close succession that will distort a smoothing filter.

The Slutsky effect is illustrated in Figure 1 which gives smoothed values of 500 standard normal variables, using moving averages of 10, 25, and 50; and

¹As a referee observed, there are two different definitions of the Slutsky effect in common use. First there is the formal sense, going back to Slutsky (1937), that applying a filter to a random series will generate *regular* cycles corresponding to peaks in the transfer function of the filter. For an m period moving average, for example, the transfer function is $f(\omega) = \left(\frac{1}{m^2}\right) (1 - \cos m\omega) / (1 - \cos \omega)$ which, for $m = 25$, has its largest peak, after zero, around 17.5 years: too short, clearly, to generate Little Ice Age behaviour. The second sense, that we use here, is the colloquial one that applying a moving average to a random series will generate the appearance of *irregular* oscillations: this is the definition given, for instance, at <http://mathworld.wolfram.com/Slutzky-YuleEffect.html>. In climatology, Burroughs (2003, 24) briefly discusses the Slutsky effect, in the second sense, in an early chapter on statistical background and gives a diagram illustrating how applying a moving average to a series of random numbers will give the appearance of irregular cycles, but does not subsequently investigate whether it can be the source of perceived climate cycles.

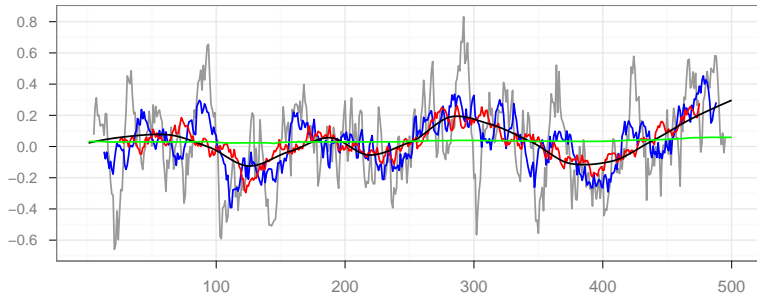


Figure 1: Slutsky effect: 500 standard normal random numbers smoothed with a 10 period moving average (grey line); 25 period moving average (blue line); 50 period moving average (red line); and loess smoother with span of one third (black line). The green line gives the posterior mean estimated by a Barry-Hartigan change point procedure.

R's loess filter with smoothing span of one third.² It can be seen that notable downward trends appear to occur around observation 100, and particularly between observations 300 and 400, which is followed by a marked upward trend. By contrast, the posterior mean estimated by a [Barry and Hartigan \(1993\)](#) change point procedure, which we will see below is particularly useful for detecting short changes in the mean value of series, shows no variation.

The Slutsky effect is illustrated in [Figure 2](#) for Low Country summer temperature since 1301. The top panel follows the standard climatological practice of smoothing the data, in this case with a 25 year moving average. These smoothed data appear to show a cooling trend from the mid-fifteenth to the early nineteenth centuries, with markedly cold episodes in the late sixteenth, late seventeenth, and early nineteenth centuries, consistent with a Little Ice Age.

However, when we look instead at the unsmoothed data, in the middle panel of [Figure 2](#), the impression is one of randomness without structural

²The variables were generated in R with seed set to 123. The loess smoother behaved almost identically to the Butterworth low pass filter with threshold of 0.025, except at the boundaries where the latter showed characteristic attenuation towards zero.

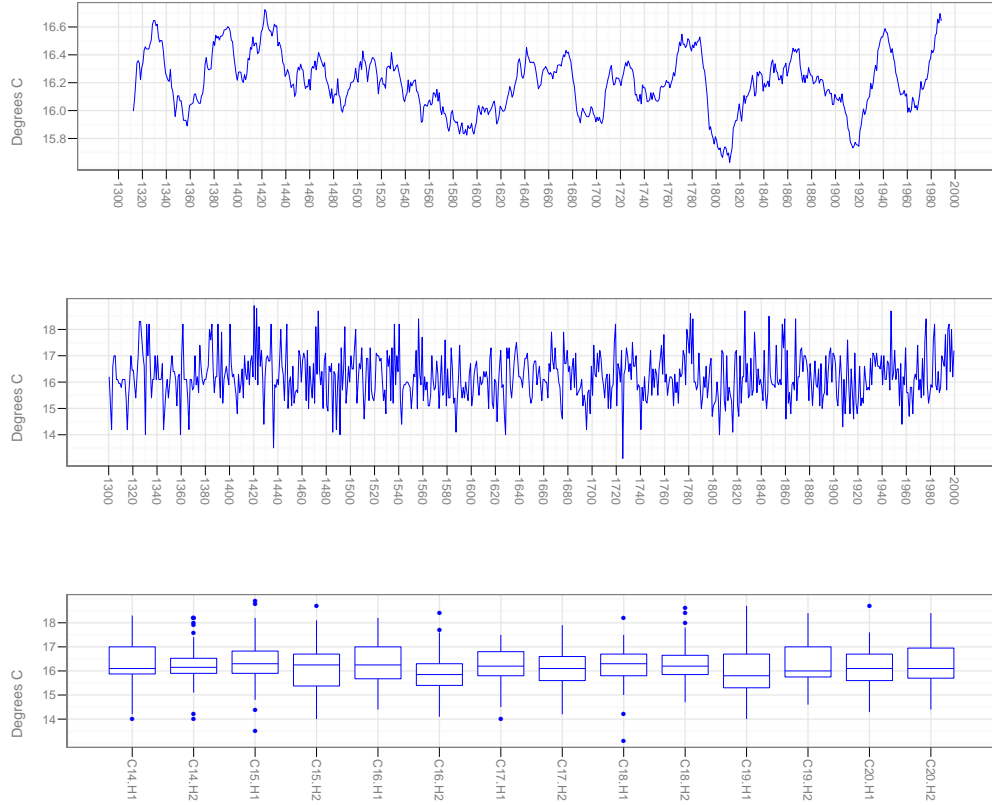


Figure 2: Low Countries summer temperature, 1301–2000. The top panel shows annual temperature smoothed by a 25 year moving average; the middle panel shows the raw series; the bottom panel shows a boxplot of the distribution of temperature by half century.

breaks, cycles or trends, something that we confirm formally below. The bottom panel shows a boxplot of the distribution of temperature by half century, which suggests that median summer temperature has fluctuated by a fraction of a degree between 1301 and 2000.

Glaciers may be seen as a physical embodiment of a Slutsky effect: their extent represents a moving average process of temperature and precipitation over preceding years, and can show considerable variation through time

even though the annual weather processes that drive them are independent draws from a fixed distribution. For example, while annual Swiss winter temperature and precipitation are close to random until the late nineteenth century, Swiss glaciers fluctuate notably, expanding from the mid-fifteenth century until 1650, contracting until 1750 and then expanding again until 1850 (Matthews and Briffa, 2005, 18–19).

By casting doubt on the occurrence of a Little Ice Age in Europe, our findings further strengthen the case for anthropogenic global warming. The global warming debate centres on whether observed rises in global temperatures over the last century are part of normal cyclical fluctuations in planetary climate, or whether they represent the outcome of human activities. By showing that, apart from some short, localized drops in summer temperature, there was little marked change in European climate between the late nineteenth century and at least the late middle ages, the rises in temperature during the twentieth century become all the more anomalous.

The rest of the paper is as follows. The traditional view of the Little Ice Age is outlined in Section 1, along with descriptions of the data sources and a comparison of Dutch weather estimates with recorded English wheat prices and yields from the thirteenth to the fifteenth centuries. Section 2 applies classical and Bayesian change point tests to the temperature reconstructions, and finds little evidence of sustained changes before the twentieth century, although there are decades of markedly lower summer temperature. Section 3 looks for temporal dependence in annual temperature using tests of conditional mean independence, while Section 4 looks at more conjectural reconstructions of German weather back to AD 1000 and finds a similar pattern of stability to the other series. Section 5 looks for breaks or autoregression in the variance of temperature, while Section 6 finds that English and Swiss precipitation behave similarly to temperature. Appendices look at instrumental records from European cities since 1700, and at the reconstruction of

average temperature across all of Europe since 1500 by [Luterbacher, Dietrich and Wanner \(2004\)](#).

1 The Little Ice Age.

Originally coined in 1939 by Matthes to describe the increased extent of glaciers over the last 4,000 years, the term ‘Little Ice Age’ now usually refers instead to a climatic shift towards colder weather occurring during the second millennium. While most climatologists dismiss the idea of the Little Ice Age as a global event, there is a consensus that much of the Northern Hemisphere above the tropics experienced several centuries of reduced mean summer temperatures, although there is some variation over dates with [Mann \(2002\)](#) suggesting the period between the fifteenth and nineteenth centuries, [Matthews and Briffa \(2005\)](#) 1570–1900, and [Mann et al. \(2009\)](#) between 1400 and 1700.

A combination of resonant images invoked by [Lamb \(1995\)](#) has linked the Little Ice Age firmly to Northern Europe. These include the collapse of Greenland’s Viking colony and the end of grape-growing in southern England in the fourteenth century; the Dutch winter landscape paintings of Pieter Bruegel (1525-69) and Hendrik Avercamp (1585-1634); the periodic ‘ice fairs’ on London’s Thames, ending in 1814; and, as the Little Ice Age waned, the contraction of Europe’s Nordic and Alpine glaciers.

1.1 Data Sources.

In this paper we analyse weather reconstructions for Europe based on documentary sources. An immediate question is why more systematic proxies such as tree rings cannot be used instead. There are two reasons. First, tree rings only reflect annual weather conditions in cold or arid areas. Secondly, [McShane and Wyner \(2011\)](#) demonstrate that currently used proxies such as tree rings, lake sediments, and ice cores have low explanatory power for

recorded Northern Hemisphere temperature since 1850, and the fit of series going back to AD 1000 is particularly weak.

Instead, the abundance of sources allows European temperatures to be reconstructed from documentary sources for the period before instrumental records around 1700. For the sixteenth and seventeenth centuries weather diaries and ships logs exist in considerable numbers. For earlier centuries, information about weather conditions is available from recorded harvest dates for grains, hay, and grapes; and, in particular, records of river tolls and water mills: how long each year were rivers unnavigable or mills unusable because waterways were frozen in winter or dried up in summer: a useful survey of these documentary sources is given by [Brázdil, Pfister and Luterbacher \(2005\)](#).

The two pioneering documentary reconstructions of European weather are Swiss temperature and precipitation from 1525 ([Pfister, 1992](#)); and the Netherlands from 1301 ([van Engelen, Buisman and IJnsen, 2001](#)),³ and we analyse both here. The current definitive reconstruction is the Central Europe reconstruction since 1500 by [Dobrovolný et al. \(2009\)](#), which includes authors of most of the previous major European weather reconstructions, and attempts to improve calibration of documentary records against instrumental records by trying to correct instrumental records for urban heat island effects, and the impact of switching the location of thermometers from north-facing walls to modern louvered boxes. The Central England series of [Manley \(1974\)](#) is based entirely on instrumental records, albeit with some heroic data splicing before 1700, and is added to look at weather in a more oceanic zone of Europe.

³Although these series start in AD 800, there are increasing numbers of missing observations as we go back past 1301 and the authors are less confident of their accuracy, putting them in wider bands that they denote by Roman rather than Arabic numerals. In running times series tests, missing observations during the 14th and early 15th centuries (30 for winter, 11 for summer) were set at the median value of the entire series.

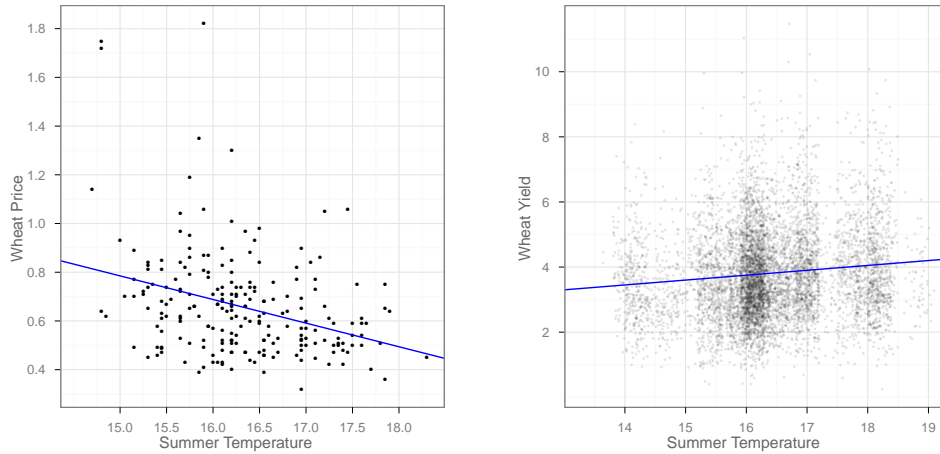


Figure 3: (a) Log English wheat price versus average Netherlands temperature over previous two years, 1213–1500. (b) Wheat yield per seed versus Netherlands summer temperature on 231 English manors, 1270–1450.

Documentary estimates of German temperature have been extended back to 1000 AD by [Glaser and Riemann \(2009\)](#) who label years as good, average, or bad. Because these data are multinomial we analyse them separately in Section 4 below.

1.2 Reliability of Documentary Reconstructions.

How believable are these documentary reconstructions? With the exception of the Dutch series, which provides detailed accounts (in Dutch) of its sources, most studies give little detailed information on sources used and methodologies used to translate documentary records into temperatures. However we can still validate these series by seeing how they correlate with records of agricultural activity not used in their reconstruction.

The most detailed and extensive records of agricultural output in Europe before the establishment of research stations at the end of the nineteenth century, are the accounts kept by English manors between the thirteenth and fifteenth centuries, which have been tabulated by [Campbell \(2007\)](#). The

	Intercept	Summer	Lag Summer	Winter	Lag Winter	RMSE	R^2	N
Price	5.771** (0.454)	-0.050** (0.019)	-0.048** (0.020)	-0.003 (0.012)	-0.010 (0.012)	0.261	0.076	164
Yield	0.504** (0.077)	0.046** (0.005)		0.001 (0.003)		0.369	0.016	6037

Regression of annual English wheat prices and yields (both in logs) on current and lagged Dutch summer and winter temperature. Standard errors in parentheses. ** denotes a coefficient significant at 1 per cent.

Table 1: Regression of annual English wheat prices (1211–1450) and yields (1270–1450) on estimated Dutch temperatures.

left hand panel of Figure 3 plots Dutch summer temperatures against the annual ratio of wheat harvested to wheat sown on 144 manors from the start of accurate records in 1270 (the earliest accounts start in 1211 but some of these early records claim anomalously high yields that are not reflected in low wheat prices: including these records did not affect the results materially); and end in 1450 by when this pattern of seigneurial agriculture carried out by coerced labour had virtually disappeared. Points are jittered to separate overlapping ones. It can be seen that yields move roughly in line with the Dutch temperature estimates of [van Engelen, Buisman and IJnsen \(2001\)](#) although explanatory power in Table 1 is low: a one degree rise in summer temperature increases the average yield ratio by 5 per cent while winter temperature has no impact. Estimating the regression with mixed effects to allow intercept and slopes to vary across manors did not show large variation across manors or change the reported estimates markedly. As a further test of the validity of these results we looked at the impact of temperature on the yields of barley and oats which are known to be more weather resistant than wheat, and found that summer temperature had a smaller effect on barley and none on oats.

English medieval agriculture was highly commercialized, and price series for wheat exist back to 1211 (Clark, 2004).⁴ Wheat could be stored for a year after harvesting, so prices reflect the previous two years' harvests: one poor harvest had a limited impact, but successive harvest failures (such as occurred during the Great Famine from 1316–17, when wheat prices rose to nearly three times their average level—shown by the two points in the northwest corner of the second panel of Figure 3) were lethal. We show elsewhere (Kelly and Ó Gráda, 2010) that death rates at all levels, from unfree tenants to the high nobility, rose sharply after poor harvests which caused epidemic disease to spread across society.

We analyse wheat prices from 1211 until 1500: a period during which the general price level was stable before the Price Revolution of the sixteenth century. Regressing log price on current and lagged summer and winter temperatures in Table 1, it can be seen that a one degree rise in summer temperature reduced prices by 5 per cent in the current and following year, while, again as we say with yields, winter temperature has no discernible impact. In summary then, the ability of the Dutch summer temperature series to predict medieval English wheat yields and prices suggests that the reconstruction is a reliable one.

2 Change Points in Temperature Since the Middle Ages.

To examine how weather deteriorated during the Little Ice Age we analyze several widely used annual summer and winter temperature reconstructions up to 2000 for Western Europe: Central Europe from 1500; Low Countries from 1301; Switzerland from 1525; and England from 1660.⁵ The Central

⁴We move Clark's price series back by one year to align them with calendar years rather than harvest years.

⁵We look at the Luterbacher, Dietrich and Wanner (2004) reconstruction of average temperature across all of Europe since 1500 in the Appendix below.

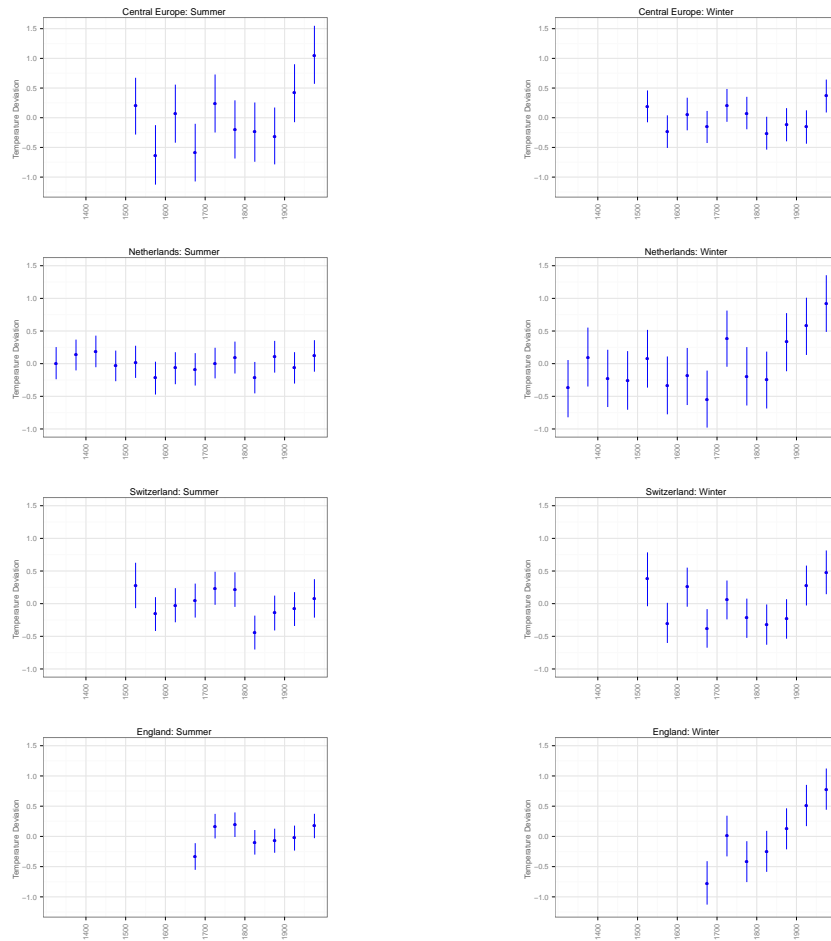


Figure 4: Mean summer and winter temperature by half century with 95 per cent credible intervals.

Europe series is expressed as a deviation in degrees from the 1961 to 1990 average; the Swiss series is measured on a continuous scale from plus to minus three; while the Dutch and English series are expressed in degrees Celsius. We subtract the mean of the Dutch and English series prior to analysis.

In this section we look at how stable mean temperatures have been over time by applying classical and Bayesian change point analysis. However, we start with a simple one-way ANOVA to examine winter and summer

temperature by half-century for each series. Temperature in year i during half-century j is assumed to be normally distributed $y_{ij} \sim N(\alpha_j, \sigma_w^2)$, while mean temperature during half-century j $\alpha_j \sim N(\mu, \sigma_b^2)$. To identify coefficients, we constrain the mean temperatures over half-centuries to sum to zero $\sum_j \alpha_j = 0$, and impose standard non-informative priors: that $\mu \sim N(0, 10\,000)$, $\sigma_b \sim U[0, 20]$, $\sigma_w \sim U[0, 20]$. This was estimated by MCMC in JAGS with 10 000 iterations, the first 2 500 being discarded. Trace plots indicate rapid convergence on the posterior distribution, and the Gelman-Rubin diagnostic supports convergence.

Figure 4 shows little variation in summer temperature with most observations lying within 0.25 degrees C of the series mean. For winter series and Central European summers the rise around the late nineteenth century is evident.

2.1 Classical Change Point Tests.

We now look at the stability of each series: can we find structural breaks in mean temperature corresponding to different phases of climate? We look at Bayesian tests below, but start with classical tests: the [Bai and Perron \(1998\)](#) procedure, implemented by [Zeileis et al. \(2002\)](#), which uses least squares to find the optimal location of k breakpoints, and then uses a Bayes Information Criterion to choose among k s; and the [Venkatraman and Olshen \(2007\)](#) circular modification of [Sen and Srivastava \(1975\)](#) binary segmentation which looks for the largest change in the partial sums of observations.⁶

For these tests to be informative, we must know their power: are they capable of detecting shifts in mean temperature of the sort that would have occurred during the Little Ice Age? We will examine the power of these

⁶The older CUSUM test [Zeileis et al. \(2002\)](#) performed poorly in simulations, only detecting half as many breaks in short series as Bai-Perron, and we do not report its results here. Because we are using seasonal averages, and only have annual data, the [Dierckx and Teugels \(2010\)](#) test for changes in the parameter of the Pareto distribution generating extreme values is not applicable.

tests to find changes of means in series which are independent draws from a normal distribution: we will see in Section 3 below that this assumption of no temporal dependence in weather series appears valid.

In 1,000 simulations where 150 observations of mean zero are followed by 150 with a mean of 0.5, BP detected the break in 90 per cent of cases, and VO in 64 per cent; while for 100 observations in each group the success rates are 71 per cent and 43 per cent. By contrast, the Bayesian change point analysis of [Barry and Hartigan \(1993\)](#) that we use below, while performs better than classical tests in detecting short breaks, finds only around one quarter of 0.5 standard deviation changes halfway through a series of length 200 or 300.⁷

Looking at a series of 150 observations where the middle 50 are 1 standard deviation higher, BP detects 96 per cent and VO 95; for a rise of 0.75 the percentages detected are 71 and 65; while for a rise of 0.5 the detection rates were 28 and 22 per cent. For a rise in 33 observations in the middle of a series of 100; for a one standard deviation increase BP detected 84 per cent of cases and Segment 77; for a rise of 0.75 the rates are 52 and 40; while for a rise of 0.5 the detection rates are 21 and 13. In other words, binary segmentation and BP can detect changes of 1 standard deviation in annual temperature (roughly 1 degree Celsius for summer temperatures in Northern Europe) that last a generation, and fairly reliably detect 0.5 standard deviation changes that last a century.

Table 2 reports the change points detected in our four series of summer and winter temperature using BP and VO. In all cases we find a break in winter temperatures around the start of the twentieth century, but no indication of any change before that. For summer temperature, England and the Netherlands show no change points. Central Europe shows a rise in the late

⁷For a change in the posterior mean to be detectable by eye, it requires a posterior probability of a change point of at least 0.15, which, looking at the maximum posterior probability for 10 observations on either side of the break, occurs in around 25 per cent of simulations.

	Start	Winter		Summer	
		BP	VO	BP	VO
C. Europe	1501	1909	1909	0	1982
Netherlands	1301	1861	1897	0	0
Switzerland	1525	1910	1911	0	1813, 1818
England	1660	1910	1911	0	0

Change points in mean winter and summer temperatures identified by Bai-Perron breakpoints, and Venkatraman-Olshen binary segmentation.

Table 2: Change points in mean winter and summer temperatures.

twentieth century, while Switzerland records a shift between 1813 and 1818: we return to this below.

In summary, classical change point tests suggest that if sustained falls in temperature did occur in Europe during the Little Ice Age their magnitude was below half a standard deviation.

If we use smoothed instead of raw data, the number of breakpoints predictably increases. For example, if data are smoothed with a loess smoother with span of one third, Bai-Perron now identifies breaks in 1576, 1651, 1803 and 1926 in Central European summers; and 1527, 1632, and 1896 in Dutch summers.

2.2 Bayesian Change Points.

The breakpoint and segmentation methods are known to perform well in detecting long-lasting changes in series, but do less well at finding shorter breaks. We therefore consider the Bayesian change point analysis of [Barry and Hartigan \(1993\)](#) implemented through the MCMC approximation of [Erdman and Emerson \(2007\)](#). Figure 5 shows the estimated mean of each series inside a 95 per cent credible interval, with the posterior probability of a

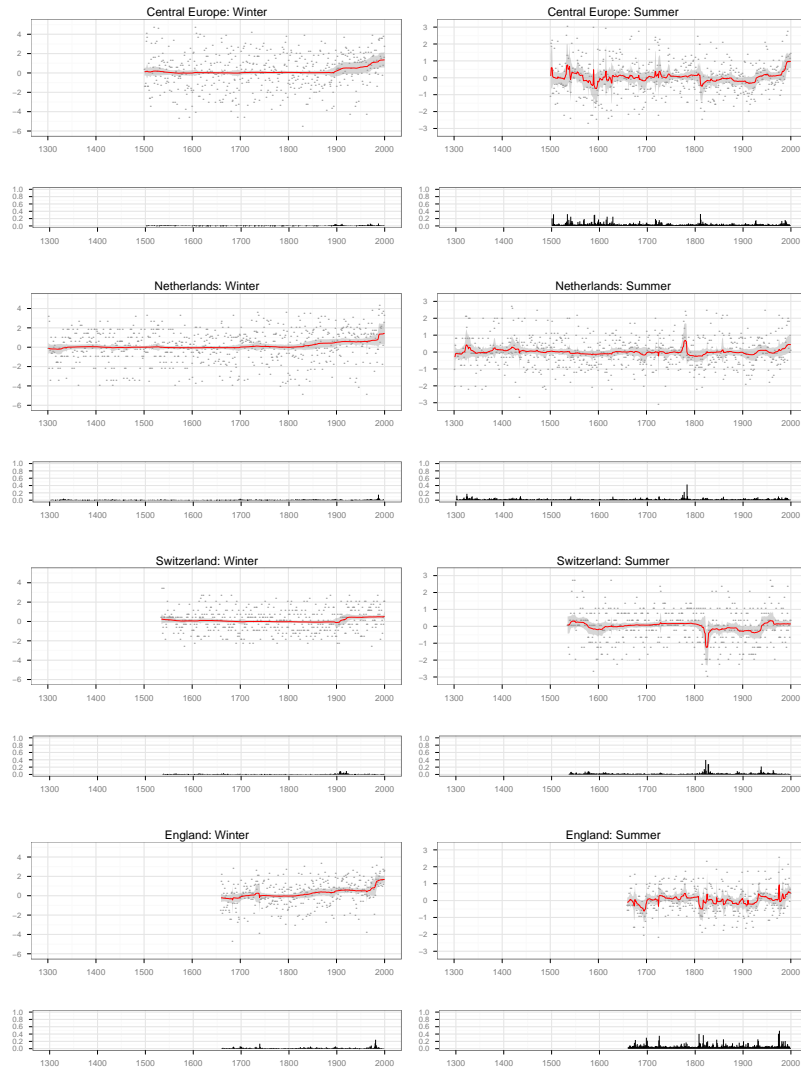


Figure 5: Posterior mean and probability of breakpoint for Barry and Har-tigan (1993) product partition model.

change point plotted below, all estimates being carried out using the default values of [Erdman and Emerson \(2007\)](#).

It is evident that winter temperatures are stable until the twentieth century when they rise markedly, particularly for England. For summers, the

older reconstructions for Dutch and Swiss temperature show little variation before the twentieth century, apart from a rise in the late eighteenth century for the Netherlands, and a drop in the 1810s for Switzerland.

The English instrumental series and the Central European series, although they do not show sustained changes before 1900, do show considerable volatility around their mean, with episodes of sustained falls in summer temperature lasting around a decade. The most notable of these, that also appears in many of the city series in Figure 7 below, occurs in the 1810s when temperatures in England were below average every year between 1809 and 1817, and in Central Europe between 1812 and 1818. Similarly, temperatures in Central Europe were below average every year from 1591 to 1598; and in England from 1687 to 1698. While summer temperatures do not show the prolonged changes that one would expect during a Little Ice Age, there are decades of notably worse weather.

3 Conditional Mean Independence.

Looking now at temporal dependence in weather, Table 3 give the results of a first order autoregression with trend $y_t = \alpha + \beta y_{t-1} + \gamma t$ for each weather series. We shall see below that this specification is adequate: there is little indication of higher order dependencies or non-linearities. Regressions where the [Bai and Perron \(1998\)](#) procedure identifies a break-point are split at that break, and the results given separately for each sub-sample.

For every series the regression R^2 is below 0.05 and in most cases below 0.02. The size of autocorrelation is small in every case, and only a few series show statistically significant autoregression at conventional levels. Central European and Dutch summers show significant correlation but with a coefficient of only 0.1: a one degree rise in temperature one summer increases average temperature next summer by an almost imperceptible tenth of a degree. The significance of the Dutch series is caused by reconstructions from

	Start	Intercept	Lag Temp.	Trend	RMSE	R^2
Summer						
C. Europe	1501	-0.084 (0.107)	0.114* (0.049)	-0.001 (0.000)	1.071	0.017
	1928	-0.756** (0.19)	0.11 (0.104)	0.016** (0.004)	0.797	0.258
Netherlands	1302	0.023 (0.069)	0.102** (0.038)	0.000 (0.000)	0.904	0.011
Switzerland	1526	0.035 (0.093)	0.069 (0.046)	0.000 (0.000)	0.995	0.007
England	1661	-0.087 (0.088)	0.099 (0.054)	0.001 (0.000)	0.804	0.015
Winter						
C. Europe	1501	-1.163** (0.195)	-0.111* (0.049)	0.000 (0.001)	1.89	0.013
	1910	-0.639 (0.377)	0.143 (0.107)	0.011 (0.007)	1.701	0.058
Netherlands	1302	-0.214 (0.14)	-0.055 (0.042)	0.000 (0.000)	1.647	0.003
	1862	0.19 (0.295)	0.137 (0.086)	0.005 (0.004)	1.701	0.037
Switzerland	1526	-0.279* (0.117)	-0.145** (0.051)	-0.001* (0.001)	1.136	0.029
	1910	0.135 (0.286)	0.11 (0.116)	0.000 (0.006)	1.224	0.012
England	1661	-0.57** (0.181)	-0.053 (0.066)	0.002 (0.001)	1.344	0.015
	1896	0.17 (0.25)	0.122 (0.099)	0.007 (0.004)	1.241	0.051

Regression of annual temperature on lagged temperature and a trend. Standard errors in parentheses. ** denotes a coefficient significant at 1 per cent. All series end in 2000, except Swiss which end in 1989. Start dates correspond to structural breaks identified by a Bai-Perron (1998) procedure.

Table 3: Regression of annual temperature reconstructions on lagged temperature and trend.

the fourteenth century, and disappears from the fifteenth century onwards; while significance for the Central Europe series disappears after 1700. Similarly, the small negative autocorrelation in the Central Europe and Swiss winter series disappears when non-instrumental observations before 1700 are excluded.

While these temperature reconstructions do not, with some exceptions for early periods, display first order autoregression, there remains the possibility that the series show some other form of dependence, either higher order linear or non-linear. To investigate this possibility we analyse our weather series for what statisticians call conditional mean independence and financial econometricians, who developed these tests to see if changes in asset prices are unpredictable, call martingale differences.

Specifically, for a stationary series $\{Y_t\}$, let $I_t = \{Y_t, Y_{t-1}, \dots\}$ denote the information set at time t . Under martingale differences $E(Y_t | I_{t-1}) = \mu$ or, equivalently, $E[(Y_t - \mu) \omega(I_{t-1})] = 0$ where $\omega(\cdot)$ is a weighting function (Escanciano and Lobato, 2009b).

For linear weights $\omega(I_{t-1}) = Y_{t-i}$, martingale differences imply zero covariances $\gamma_i = E[(Y_t - \mu)(Y_{t-i} - \mu)] = 0$ for $i > 0$ or, equivalently, correlations $\rho_i = \gamma_i / \gamma_0 = 0$. This leads to the standard Ljung-Box portmanteau test based on the sum of the first p squared correlations $T \sum_{i=1}^p \tilde{\rho}_i^2$ where $\tilde{\rho}_i = \rho_i \sqrt{(T+2)/(T-i)}$. We apply the modification of Escanciano and Lobato (2009a) where correlations are divided by sample autocovariances of the squared series to provide robustness against heteroskedasticity, and p is chosen by a data dependent procedure where a penalty term is subtracted that switches between Akaike and Bayes Information Criteria.

The other linear test we apply is a variance ratio test, based on the idea of Lo and MacKinlay (1989) that, for uncorrelated series, estimated variance should rise in proportion to the length of the series: $AVR = 1 + 2 \sum_{i=1}^{p-1} (1 - i/p) \rho_i$. We apply the Kim (2009) modification where p is chosen by a data dependent procedure, and the distribution of the test statistic is derived by applying a wild bootstrap, where each term Y_t of the original series is multiplied by a random variable with zero mean and unit variance.

The third class of tests for temporal dependence in time series are tests for the departure of the series spectrum from linearity. In econometrics these originate with Durlauf (1991), and we report the generalized spectral

	Summer			Winter		
	Q	VR	Spec.	Q	VR	Spec
1701–1900						
C. Europe	0.23	0.06	0.43	0.29	0.00	0.78
Netherlands	0.13	0.05	0.1	0.3	0.24	0.86
Switzerland	0.53	0.43	0.25	0.17	0.12	0.29
England	0.42	0.42	0.51	0.46	0.6	0.62
1701–2000						
C. Europe	0.01	0.02	0.01	0.32	0.32	0.83
Netherlands	0.05	0.01	0.10	0.63	0.64	0.62
Switzerland	0.01	0.26	0.19	0.71	0.74	0.47
England	0.22	0.14	0.32	0.00	0.04	0.04
pre-1701						
C. Europe	0.03	0.01	0.01	0.07	0.00	0.09
Netherlands	0.24	0.00	0.29	0.61	0.92	0.85
Switzerland	0.02	0.25	0.03	0.45	0.36	0.2
England	0.16	0.15	0.4	0.06	0.04	0.05
pre-1901						
C. Europe	0.03	0.02	0.02	0.04	0.00	0.23
Netherlands	0.01	0.00	0.00	0.38	0.41	0.58
Switzerland	0.19	0.13	0.54	0.01	0.65	0.01
England	0.00	0.03	0.09	0.62	0.79	0.48

p values for tests of conditional independence of means of temperature series until 2000. Q is a robustified portmanteau test with automatic lag selection. VR gives the wild bootstrap test results for an automatic variance ratio test. Spec is a generalized spectral test.

Table 4: Tests for conditional independence in means.

test of [Escanciano and Velasco \(2006\)](#), which corresponds to an exponential weighting function ω .

Looking at the small sample properties of these tests for samples with 100, 300 and 500 observations, [Charles, Darné and Kim \(2011\)](#) find that the reported size of all tests is approximately correct, and that against models of linear dependence the automated variance ratio test shows highest power, while against a variety of non-linear processes, the generalized spectrum test works best. The temperature series here do not appear to exhibit

non-linearity: applying a Terasvirta-Lin-Granger (1993) test for non-linearity in means using Trapletti and Hornik (2010) led to p values in excess of 0.09 for all series, with values above 0.5 in most cases.

Table 4 reports the p values of these three tests for each temperature series, calculated using the default values of Kim (2010). The first block reports uses each series from 1701 to 1900; the second is for all years after 1701; the third has each series from its start (1500 for Central Europe, 1301 for the Netherlands, 1525 for Switzerland, and 1660 for England) until 1700; while the final block gives results for the pre-1901 period.

It can be seen that for 1701–1900 (a period whose start lies in conventional definitions of the Little Ice Age) the only test that rejects conditional mean independence at conventional levels is the VR test for European winters. This seems to result from excessive sensitivity of the VR test: this series has a first order autoregressive coefficient of -0.06 with p -value of 0.36; and if the residuals from this first order autoregression are tested, the VR test returns a p -value of 0.82 (with the automatic portmanteau and generalised spectrum giving similar values), indicating that the first-order specification is adequate. Adding in twentieth century observations in the second block, the only series to show systematic departures from conditional mean independence are Central European summers and English winters, both of which rise notably after 1900.

The excess sensitivity of the VR test also appears in the pre-1701 Central Europe winter data: the coefficient of a first order autoregression is -0.14 with p -value of 0.05; and applying the VR test to residuals gives a p -value of 0.97. For the summer data, the coefficient of a first order autoregression is 0.12 with p -value of 0.09; and applying the VR test to residuals gives a p -value of 0.82. We find the same thing with pre-1701 Dutch temperatures: the coefficient of a first order autoregression is -0.04 with p -value of 0.38, and applying the VR test to residuals gives a p -value of 0.87.

4 German Temperatures over the Past Millennium.

Reconstructions of German temperature by [Glaser and Riemann \(2009\)](#) allow us to extend our analysis, tentatively, back to 1000 AD. For years until 1500 [Glaser and Riemann \(2009\)](#) assign each season a minus one for bad, zero for average, and one for good. After 1500, months are placed on an integer scale from minus three to three. To splice the two series, we take January and July values after 1500, and set all minus values to minus one, and all plus values to one. Between 1101 and 1900 (the first century of observations contains a lot of missing data), 40 per cent of winters are labelled bad, and thirty per cent are good; while 30 per cent of summers are bad and 35 per cent good.

We start by looking at the proportion of good or bad years in each century. We assume that y_j , the number of good (or, alternatively, bad) years in century j , has binomial distribution $y_j \sim \text{bin}(n = 100, p_j)$ where the logit of p_j is normally distributed $\log(p_j/1 - p_j) \sim N(\mu, \sigma^2)$. We impose the priors that $\mu \sim N(0, 10\,000)$, $\sigma \sim U[0, 20]$. This was estimated by MCMC in JAGS with 10 000 iterations, the first 2 500 being discarded, with the Gelman-Rubin diagnostic indicating convergence.

We also looked at the probability of a good year, conditional on the previous year being good; and of a bad year, conditional on the previous year being bad; on the grounds that successive good years may have been more common in the Medieval Warm Period, and successive bad ones in the Little Ice Age.

Figure 6 plots these probabilities, with 95 per cent credible intervals, for summers on the left and winters on the right. The eleventh and twentieth centuries at either end stand out from the rest. The twentieth century has a higher probability of good winters and successive good winters than earlier centuries, consistent with global warming. At the other end, the eleventh century has a lower probability of good or bad years, because years with

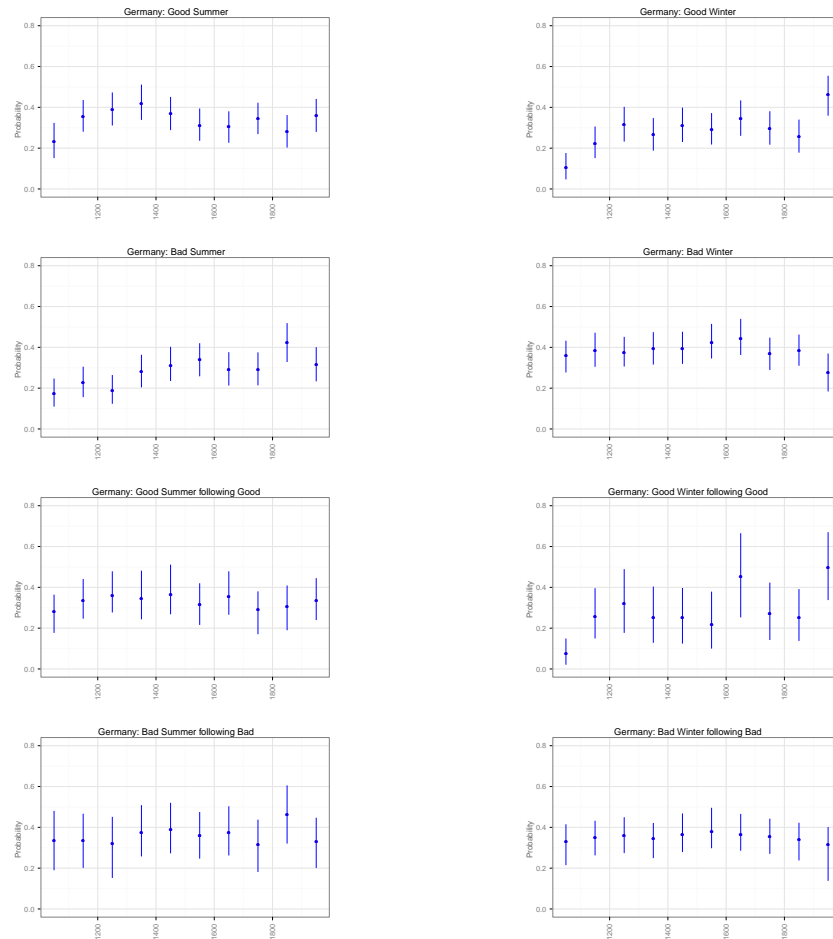


Figure 6: Probabilities of good, bad, good following good, and bad following bad summers and winters by century for Germany, 1001–2000 AD, with 95 per cent credible intervals.

missing values are assigned an average value by [Glaser and Riemann \(2009\)](#). In between, the probability of good or bad winters appears fairly constant, as does the probability that a bad winter will be followed by a bad one. The probability of a good winter conditional on the previous winter being good is also fairly constant, except for the seventeenth century which is nearly 20

per cent higher than surrounding centuries despite being in the depth of the Little Ice Age, although the credible intervals overlap.

For summers, the probability of a good summer, or successive good or bad summers is fairly constant. However the probability of bad summers is somewhat lower in the twelfth and thirteenth centuries, and higher in the nineteenth; but the credible intervals again overlap with other centuries.

Next we consider conditional mean independence. An immediate problem is that the tests used earlier assume normal data, whereas these data are multinomial. Carrying out Monte Carlo simulations on independent sequences of length 400 where minus one, zero and one occur with equal probability showed that two of the tests performed poorly. In over 99 per cent of simulations the automatic variance ratio test returned a p -value of 0.74, and 0.37 in the remaining cases. Similarly, the generalized spectrum test returned a p -value of 0.58 in 99 per cent of simulations, and 0.28 in the remainder. However, the automated portmanteau test proved robust to multinomial data, although the nominal p -values are slightly low. In 10 000 simulations, the first percentile was 0.003; the fifth was 0.035; and the tenth was 0.085.

For Germany from 1101 to 1900, the p -value for winter is 0.26 and for summer 0.00. From 1101 to 1500 the winter and summer values are 0.49 and 0.03 (these do not change substantially if the eleventh century is included); while for 1501 to 1900 they are 0.46 and 0.02. These p -value are based on applying the automated portmanteau test to 10 000 simulations of length 400 where minus one occurs with probability 0.4 and plus one with probability 0.3 for winter simulations; and minus one with probability 0.3; and plus one with probability 0.35 for summer simulations. The data indicate conditional mean independence in winter weather, and suggest some slight dependence in summer weather.

	Summer					Winter				
	$Q_1(5)$	$Q_1(10)$	$Q_3(10)$	$Q_6(20)$	HS	$Q_1(5)$	$Q_1(10)$	$Q_3(10)$	$Q_6(20)$	HS
C. Europe	1.18	1.6	8.49	6.14	3.05	0.34	0.7	0.21	0.55	0.34
Netherlands	0.45	0.54	0.31	0.63	0.42	0.19	0.4	0.49	0.36	0.93
Switzerland	1	0.91	5.52	3.25	0.89	0.53	0.54	0.62	0.83	0.35
England	0.3	0.33	0.3	0.74	0.29	0.24	0.3	0.31	0.12	0.24

Each number is the value of the test statistic relative to its 5 per cent critical value. $Q_1(5)$ and $Q_1(10)$ are McLeod-Li portmanteau tests truncated at 5 and 10 lags; $Q_3(10)$ and $Q_6(20)$ are ? volatility tests truncated at 10 and 20 lags, and HS are [Harvey and Streibel \(1998\)](#) tests.

Table 5: Tests of conditional heteroskedasticity in summer and winter temperatures.

5 Temperature Variance.

While there appears to be little sustained change in mean European temperatures before the late nineteenth century, there is the possibility that variance may have changed through time, so that earlier centuries may have experienced greater volatility of temperatures.

Looking at the square of a series of normal variables with mean zero where the variance standard deviation rises from 1 to 1.25 halfway through, the [Bai and Perron \(1998\)](#) procedure detects a break 39 per cent of the time for series of length 300, 36 per cent of the time for series of length 200, and 20 per cent of the time for series of length 100. Because the data are not normal, the [Venkatraman and Olshen \(2007\)](#) binary segmentation procedure detects fewer than 3 per cent of changes. Where the standard deviation rises to 1.5, the respective success rates of Bai-Perron are 100, 87, and 54. In other words, the break point procedure can fairly reliably detect a 50 per cent rise in standard deviation in the middle of a series of length 200.

Applying breakpoints to the squared weather data, the only series to show a break is Central European summers, in 1728, shortly after the start of reliable instrumental records with Fahrenheit’s 1724 invention of the mercury thermometer. Before this variance is 1.49, and falls to 0.80 afterwards.

Tests for autoregressive heteroskedasticity look for temporal dependence in squared residuals. The standard McLeod-Li test applies a Ljung-Box test but has low power to detect weak autocorrelation and we therefore also consider the [Rodriguez and Ruiz \(2005\)](#) $Q_i(M) = T \sum_{k=1}^{M-i} (\sum_{l=0}^i \tilde{\rho}_{k+l})^2$ and [Harvey and Streibel \(1998\)](#) tests $HS = T^{-1} \sum_{i=1}^{T-1} \rho_i i$ which [Rodriguez and Ruiz \(2005\)](#) show to perform considerably better in detecting a variety of patterns of dependence.

Table 5 reports the results of applying these tests to each temperature series. The first two columns report McLeod-Li tests truncated at 5 and 10 lags; the next two give Rodriguez-Ruiz tests truncated at 10 and 20 lags (with parameter i set to their recommended value of one third of the lag minus one); and the final one the Harvey-Streibel statistic. Each statistic is reported relative to its 5 per cent critical value, so values over one denote values significant over 5 per cent. Changing the lags of the McLeod-Li and Rodriguez-Ruiz tests did not alter the pattern of significance. Given the break in many series in the late nineteenth century, each series is truncated in 1900. We apply the tests directly to each squared series.

It can be seen that the only series showing systematically significant temporal dependence in variance is Central European summers. If we split this series in 1728, when we know that variance falls, the value of the Harvey-Streibel statistic falls to 0.65 for earlier observations, but remains at 3.21 for later observations, with other statistics showing the same pattern of significance. The Rodriguez-Ruiz statistic also detects dependence in Swiss summers.

To estimate the parameters of the variance of Central European summers between 1729 and 1900 we assume a GARCH(1,1) process where the deviation of temperature from its mean $y_t = \epsilon_t$ where ϵ_t has t distribution with ν degrees of freedom, and variance σ_t^2 that follows $\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \beta \sigma_{t-1}^2$. We estimated this using the BayesGARCH package of [Ardia and Hoogerheide \(2010\)](#), using the package's default diffuse priors. Convergence to the

	Summer			Winter		
	ICC	σ_w	σ_b	ICC	σ_w	σ_b
C. Europe	0.009 [0,0.036]	1.911 [1.779,2.048]	0.148 [0.014,0.365]	0.038 [0.001,0.102]	1.07 [0.991,1.152]	0.198 [0.03,0.355]
Netherlands	0.018 [0,0.064]	1.639 [1.551,1.738]	0.191 [0.018,0.419]	0.037 [0.001,0.094]	0.895 [0.846,0.951]	0.164 [0.032,0.283]
Switzerland	0.007 [0,0.042]	1.136 [1.056,1.22]	0.074 [0.006,0.237]	0.056 [0.006,0.138]	0.964 [0.9,1.035]	0.224 [0.078,0.383]
England	0.025 [0,0.099]	1.352 [1.239,1.477]	0.184 [0.012,0.446]	0.083 [0.018,0.207]	0.773 [0.706,0.849]	0.225 [0.108,0.388]

Intraclass correlation, between decade standard deviation, and within decade standard deviation of annual temperature series. 95 per cent credible intervals in brackets.

Table 6: Within and between decade standard deviation of annual temperature series before 1900.

posterior was slow for ν , and we used a burn-in of 25 000 iterations, followed by 75 000 iterations. The estimates were $\alpha_0 = 0.396$, with 95 per cent credible interval $[0.095, 0.787]$, $\alpha_1 = 0.147 [0.017, 0.357]$, $\beta = 0.408 [0.021, 0.805]$ suggesting modest persistence in volatility, although it must be remembered that the series is only a fraction of the thousand observations usually felt to be the minimum for reliable inference of GARCH parameters.

Given that variance appears fairly constant for most series, it is worthwhile to compare the variance of temperature between periods with the variance within periods. We use the same specification, priors, and number of iterations for annual temperature as in Figure 4, but use decades rather than half-centuries as our period of analysis (the results are almost identical if we use half-centuries), and end each series in 1900.

Table 6 reports the estimated standard deviation within decades σ_w and between decades σ_b , and also the intra-class correlation: the percentage of the variance of the series accounted for by between-class variance: $ICC \equiv \sigma_b^2 / (\sigma_b^2 + \sigma_w^2)$. It can be seen that the standard deviation of temperature between decades is low: of the order of one quarter of a degree Celsius. Similarly, the intra-class correlation is low, with between-decade variance

	Start	ρ	SE(ρ)	ICC	Q	VR	Spec
Switzerland							
Summer	1525	0.127	0.059	0.056	0.04	0.01	0.05
	1812	0.000	0.076	0.045	1.00	0.75	0.26
Winter	1525	-0.006	0.052	0.002	0.92	0.87	0.07
	1899	0.051	0.105	0.095	0.63	0.10	0.12
England							
Summer	1766	0.093	0.092	0.042	0.29	0.07	0.45
	1883	-0.018	0.096	0.038	0.85	0.72	0.78
Winter	1766	0.043	0.103	0.11	0.66	0.10	0.74
	1864	-0.061	0.09	0.034	0.45	0.88	0.66

Summary statistics for precipitation. ρ is first order autoregressive coefficient, and SE(ρ) its associated standard error. ICC is the intra-class correlation between decades. Q is the p value of a robustified portman-teau test with automatic lag selection. VR gives the wild bootstrap test results for an automatic variance ratio test. Spec is a generalized spectral test.

Table 7: Temporal dependence in annual precipitation in Switzerland and England.

accounting for one to three per cent of variance for summer temperature, and four to eight per cent for winter.

6 Precipitation.

While our focus has been on temperature, we also have precipitation estimates for Switzerland from 1525, and from England from 1760. Using a [Bai and Perron \(1998\)](#) procedure to detect breaks, English summer precipitation falls from an average of 450 mm before 1883 to 410 mm after, while mean winter precipitation rises from 450mm to 500 mm after 1864. For Switzerland, where precipitation is measured on a scale from 3 (very wet) to -3 (very dry), mean summer precipitation falls from 0.22 to -0.07 after 1812, while winter precipitation rises from -0.24 to 0.25 after 1899.

Table 7 shows the first order autocorrelation and standard error for each series. It can be seen that only Swiss summer precipitation before 1812

shows autocorrelation significant at 5 per cent, but the magnitude, 0.13, is small. The third column shows intra-class correlation over decades (where each series is rounded to the nearest 10 years) which ranges from 3 to 10 per cent. The last three columns report the tests for conditional independence in means described earlier. It can be seen that, with the exception of Swiss summer precipitation, no series gives evidence of temporal dependence.

7 Conclusions.

Our intention was to estimate the extent and timing of climate changes during the Little Ice Age. To our surprise, standard documentary reconstructions give little indication of sustained structural breaks in weather series before the late nineteenth century, although there are occasional decades of notably poor summer weather. If Europe experienced a Little Ice Age, the weather reconstructions analyzed suggest that temperature falls were of the order of less than half of one standard deviation.

Our findings are a reminder that some of the changes claimed by [Lamb \(1995\)](#) to be consequences of the Little Ice Age may have other possible causes. The freezing of the Thames—which for most people is the most salient fact about the Little Ice Age—was due to Old London Bridge which effectively acted as a dam, creating a large pool of still water which froze twelve times between 1660 and 1815. Tidal stretches of the river have not frozen since the bridge was replaced in 1831, even during 1963 which is the third coldest winter (after 1684 and 1740) in the Central England temperature series that starts in 1660.

For Greenland's Vikings, competition for resources with the indigenous Inuit, the decline of Norwegian trade in the face of an increasingly powerful German Hanseatic League, the greater availability of African ivory as a cheaper substitute for walrus ivory, overgrazing, plague, and marauding pirates probably all played some role in its demise ([Brown, 2000](#)); and even if

weather did worsen, the more fundamental question remains of why Greenland society failed to adapt ([McGovern, 1981](#)). The disappearance of England's few vineyards is associated with increasing wine imports after Bordeaux passed to the English crown in 1152, suggesting that comparative advantage may have played a larger role than climate.

Similarly, the decline of wheat and rye cultivation in Norway from the thirteenth century may owe more to lower German cereal prices than temperature change ([Miskimin, 1975](#), 59). Moreover, with worsening climate we would expect wheat yields to fall relative to the more weather-robust spring grains barley and oats whereas [Apostolides et al. \(2008, Tables 1A, 1B\)](#) find that between the early fifteenth and late seventeenth century, wheat yields show no trend relative to oats, and rise steadily relative to barley.

Finally, demography supports our reservations about a European Little Ice Age. We would expect northern Europe to have shown weak population growth as the Little Ice Age forced back the margin of cultivation. In fact, while the population of Europe in 1820 was roughly 2.4 times what it had been in 1500, in Norway the population was about 3.2 times as large, in Switzerland 3.5 times, in Finland 3.9 times, and in Sweden 4.7 times as large as in 1500 ([Maddison, 2009](#)).

In summary, this paper makes two points: one methodological, one historical. First, smoothing random or near random data is problematic, but the most reliable and informative results, both in terms of avoiding spurious oscillations and detecting real breaks, are given by the [Barry and Hartigan \(1993\)](#) procedure. Secondly, although most of us have strong priors that Europe experienced bouts of markedly worse weather during the Little Ice Age, such episodes are not apparent in standard temperature reconstructions.

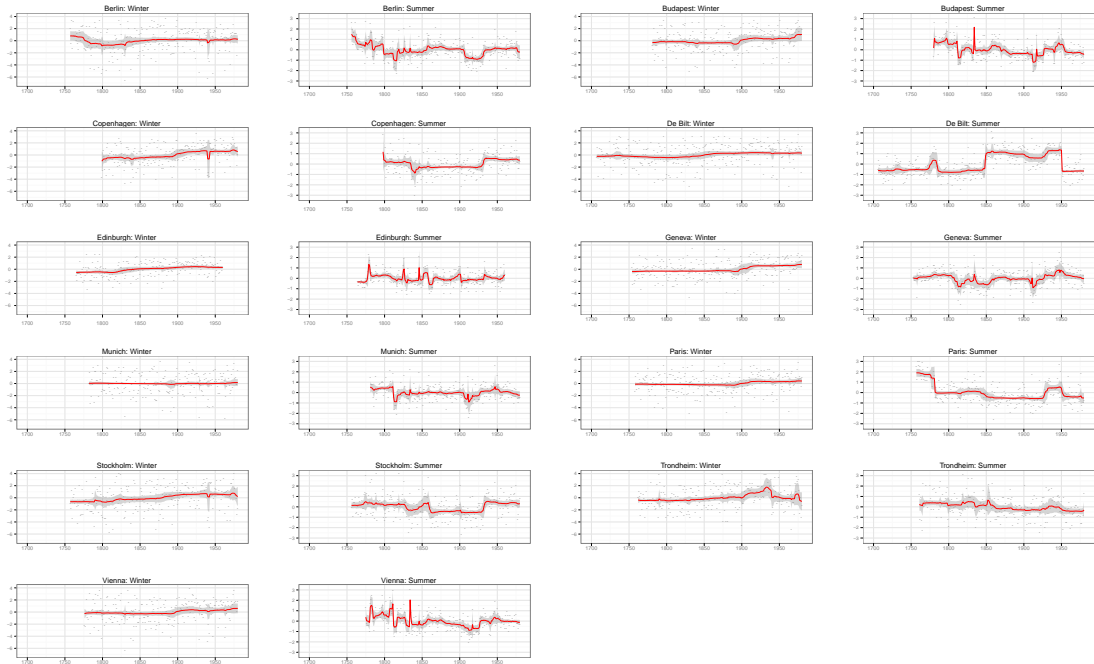


Figure 7: Posterior mean of city summer and winter temperatures since 1700.

Appendix: European Cities since 1700.

Looking at documentary reconstructions of European temperature we have found a common pattern of considerable short run instability in summer temperature but little indication of larger structural breaks or temporal dependence before the late nineteenth century. To validate these results we can compare them with instrumental records from European cities, that begin shortly after 1700 and end in 1980.⁸

Figure 7 plots winter and summer temperature along with Barry-Hartigan estimates of means for each city. It can be seen that winter temperatures

⁸European cities summer and winter temperatures are averages of June to August, and December to January temperatures from the jonesnh.dat file in “An Updated Global Grid Point Surface Air Temperature Anomaly Data Set: 1851-1990” available at <http://cdiac.ornl.gov/ftp/ndp020/>.

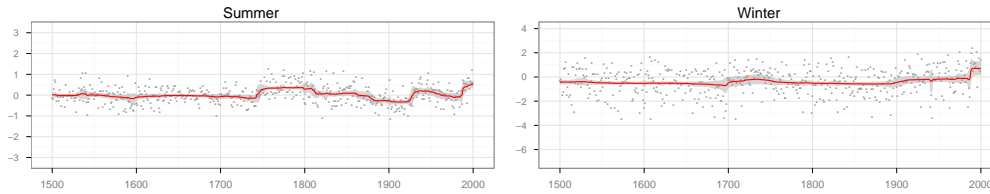


Figure 8: Barry-Hartigan posterior means for Luterbacher et al (2004) reconstructions of European summer and winter temperature.

are stable or slightly upward trending, with a small but notable dip in many places during the 1940s. Summer temperatures are a good deal more volatile. Some of the larger breaks, such as De Bilt (Netherlands) between 1850 and 1950, and Paris during the 1930s and 1940s do not appear in nearby cities and appear to be the results of changes in recording methodology; but there are common falls such as the 1810s in Berlin, Munich, Budapest and Geneva, and the 1910s in much of central Europe.

Appendix: European Average Temperature Estimates of [Luterbacher, Dietrich and Wanner \(2004\)](#).

Plots of city temperatures emphasize the extreme spatial variability of weather across northern and central Europe. An attempt to construct an average annual temperature series for all of Europe was carried out by [Luterbacher, Dietrich and Wanner \(2004\)](#), using instrumental data after 1659; and the estimates analysed earlier, or earlier estimates by the same authors, along with tree ring and ice core data for earlier centuries. These estimates are plotted with Barry-Hartigan posterior means in [Figure 8](#). It is immediately evident that the annual variation is considerably lower than in the series for individual places analysed earlier. Winter shows the same pattern of stabil-

ity, but there is a large rise in temperature in the late eighteenth century that does not appear in any other of the European reconstructions but can be seen in some of the city records. Another marked difference is the marked autocorrelation of the summer reconstructions: the first order autoregression coefficient for temperatures before 1900 is only -0.05 for winter with p -value of 0.25, but 0.23 for summer, significant at 1 per cent.

Appendix: Data Sources.

- Central European temperature reconstructions from 1500 by [Dobrovolný et al. \(2009\)](#) are available at <ftp://ftp.ncdc.noaa.gov/pub/data/paleo/historical/europe/dobrovolny2010temperature.xls>
- The Low Countries temperature series of [van Engelen, Buisman and IJnsen \(2001\)](#) are available at www.knmi.nl/kd/metadata/nederland_wi_zo.html.
- Swiss summer and winter temperature and precipitation from [Pfister \(1992\)](#) are available at <ftp://ftp.ncdc.noaa.gov/pub/data/paleo/historical/switzerland/clindef.txt>
- Monthly mean Central England temperature from 1659 are from <http://hadobs.metoffice.com/hadcet/cetml1659on.dat>, and the monthly England and Wales precipitation series from 1766 are from http://hadobs.metoffice.com/hadukp/data/monthly/HadEWP_monthly_qc.txt.
- German temperature reconstructions from 1000 AD by [Glaser and Riemann \(2009\)](#) are available at <ftp://ftp.ncdc.noaa.gov/pub/data/paleo/historical/europe/glaser2009temperature/glaser2009temperature.xls>

References

- Apostolides, Alexander, Stephen Broadberry, Bruce Campbell, Mark Overton and Bas van Leeuwen. 2008. English Gross Domestic Product, 1300–1700: Some Preliminary Estimates. Working paper University of Warwick. [32](#)
- Ardia, David and Lennart F. Hoogerheide. 2010. “Bayesian Estimation of the GARCH(1,1) Model with Student-t Innovations.” *The R Journal* 2(2):41–47. [28](#)
- Bai, Jushan and Pierre Perron. 1998. “Estimating and Testing Linear Models With Multiple Structural Changes.” *Econometrica* 66:47–78. [2](#), [15](#), [19](#), [27](#), [30](#)
- Barry, Daniel and J A Hartigan. 1993. “A Bayesian Analysis for Change Point Problems.” *Journal of the American Statistical Association* 88(421):309. [2](#), [6](#), [16](#), [17](#), [32](#)
- Brázdil, Rudolf, Christian Pfister and Jürg Luterbacher. 2005. “Historical Climatology In Europe – The State Of The Art.” *Climatic Change* 70:363–430. [10](#)
- Brown, Dale M. 2000. “The Fate of Greenland’s Vikings.” *Archaeology* . Available online: <http://www.archaeology.org/online/features/greenland>. [31](#)
- Burroughs, William James. 2003. *Weather Cycles: Real or Imaginary?* Second ed. Cambridge: Cambridge University Press. [5](#)
- Campbell, Bruce M. S. 2007. *Three Centuries of English Crops Yields, 1211–1491*. WWW Document. URL <http://www.cropyields.ac.uk>. [11](#)

- Charles, Amélie, Olivier Darné and Jae H. Kim. 2011. “Small Sample Properties of Alternative Tests for Martingale Difference Hypothesis.” *Economics Letters* 110(2):151–154. [22](#)
- Clark, Gregory. 2004. “The Price History of English Agriculture, 1209–1914.” *Research in Economic History* 22:41–123. [13](#)
- Dierckx, G. and J. L. Teugels. 2010. “Change point analysis of extreme values.” *Environmetrics* 21:661–686. [15](#)
- Dobrovolný, Petr, Anders Moberg, Rudolf Brázdil, Christian Pfister, Rüdiger Glaser, Rob Wilson, Aryan Engelen, Danuta Limanówka, Andrea Kiss, Monika Halíčková, Jarmila Macková, Dirk Riemann, Jürg Luterbacher and Reinhard Böhm. 2009. “Monthly, Seasonal and Annual Temperature Reconstructions for Central Europe Derived from Documentary Evidence and Instrumental Records since AD 1500.” *Climatic Change* 101(1-2):69–107. [2](#), [10](#), [35](#)
- Durlauf, S. 1991. “Spectral Based Testing of the Martingale Hypothesis.” *Journal of Econometrics* 50(3):355–376. [21](#)
- Erdman, Chandra and John W. Emerson. 2007. “bcp: An R Package for Performing a Bayesian Analysis of Change Point Problems.” *Journal of Statistical Software* 23(3):1–13. [17](#), [18](#)
- Escanciano, J and C Velasco. 2006. “Generalized Spectral Tests for the Martingale Difference Hypothesis.” *Journal of Econometrics* 134(1):151–185. [22](#)
- Escanciano, J. Carlos and Ignacio N. Lobato. 2009a. “An Automatic Portmanteau Test for Serial Correlation.” *Journal of Econometrics* 151(2):140–149. [4](#), [21](#)
- Escanciano, J. Carlos and Ignacio N. Lobato. 2009b. Testing the Martingale Hypothesis. In *Palgrave Handbook of Econometrics: Volume 2: Applied*

- Econometrics*, ed. Terence C. Mills and Kerry Patterson. London: Palgrave Macmillan. 21
- Glaser, Rüdiger and Dirk Riemann. 2009. “A Thousand Year Record of Climate Variation for Central Europe at a Monthly Resolution.” *Journal of Quaternary Science* 24:437–449. 4, 11, 24, 25, 35
- Harvey, Andrew and Mariane Streibel. 1998. “Testing for a Slowly Changing Level with Special Reference to Stochastic Volatility.” *Journal of Econometrics* 87:167–189. 27, 28
- Kelly, Morgan and Cormac Ó Gráda. 2010. Living Standards and Mortality since the Middle Ages. Working paper School of Economics, University College Dublin. 13
- Kim, Jae H. 2009. “Automatic Variance Ratio Test under Conditional Heteroskedasticity.” *Finance Research Letters* 6(3):179–185. 21
- Kim, Jae H. 2010. *vrtest: Variance Ratio tests and other tests for Martingale Difference Hypothesis*.
URL: <http://cran.r-project.org/web/packages/vrtest/index.html> 23
- Lamb, H. H. 1995. *Climate, History and the Modern World*. Second ed. London: Routledge. 9, 31
- Lo, Andrew W. and A. Craig MacKinlay. 1989. “The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation.” *Journal of Econometrics* 40:203–238. 21
- Luterbacher, Jürg, Daniel Dietrich and Heinz Wanner. 2004. “European Seasonal and Annual Temperature Variability, Trends, and Extremes Since 1500.” *Science* 303:1499–1503. 9, 13, 34

- Maddison, Angus. 2009. *Statistics on World Population, GDP and Per Capita GDP, 1-2006 AD*. WWW document URL: <http://www.ggdc.net/maddison/>. 32
- Manley, Gordon. 1974. “Central England Temperatures: Monthly Means 1659 to 1973.” *Quarterly Journal of the Royal Meteorological Society* 100:389–405. 2, 10
- Mann, Michael E. 2002. Little Ice Age. In *Encyclopedia of Global Environmental Change*, ed. Michael C. McCracken and John S. Perry. Wiley. 2, 9
- Mann, Michael E., Zhihua Zhang, Scott Rutherford, Raymond S. Bradley, Malcolm K. Hughes, Drew Shindell, Caspar Ammann, Greg Faluvegi and Fenbiao Ni. 2009. “Global Signatures and Dynamical Origins of the Little Ice Age and Medieval Climate Anomaly.” *Science* 326:1256–1260. 2, 9
- Matthews, John A. and Keith R. Briffa. 2005. “The ‘Little Ice Age’: Re-evaluation of an evolving concept.” *Geografiska Annaler* 87 A:17–36. 2, 8, 9
- McGovern, T. H. 1981. The Economics of Extinction in Norse Greenland. In *Climate and History*, ed. T. M. L. Wrigley, M. J. Ingram and G. Farmer. Cambridge: Cambridge University Press pp. 404–433. 32
- McShane, Blakeley B. and Abraham J. Wyner. 2011. “A Statistical Analysis of Multiple Temperature Proxies: Are Reconstructions of Surface Temperatures Over the Last 1000 Years Reliable?” *Annals of Applied Statistics* . Forthcoming. 3, 9
- Miskimin, Harry A. 1975. *The Economy of Early Renaissance Europe, 1300-1460*. Cambridge: Cambridge University Press. 32

- Pfister, C. 1992. Monthly Temperature and Precipitation in Central Europe 1525-1979. In *Climate Since A.D. 1500*, ed. R.S. Bradley and P.D Jones. London: Routledge. [2](#), [10](#), [35](#)
- Rodriguez, J. and E. Ruiz. 2005. “A Powerful Test for Conditional Heteroscedasticity for Financial Time Series with Highly Persistent Volatilities.” *Statistica Sinica* 15:505–525. [28](#)
- Sen, Ashish and Muni S. Srivastava. 1975. “On Tests for Detecting Change in Mean.” *The Annals of Statistics* 3(1):pp. 98–108. [15](#)
- Slutsky, Eugen E. 1937. “The Summation of Random Causes as the Source of Cyclic Processes.” *Econometrica* 5:105–146. [5](#)
- Teraesvirta, T, C. F. Lin and C. W. J. Granger. 1993. “Power of the Neural Network Linearity Test.” *Journal of Time Series Analysis* 14:209–220. [23](#)
- Trapletti, Adrian and Kurt Hornik. 2010. *tseries: Time Series Analysis and Computational Finance. R Package*. [23](#)
- van Engelen, A.F.V., J. Buisman and F. IJnsen. 2001. A Millennium of Weather, Winds and Water in the Low Countries. In *History and Climate: Memories of the Future?*, ed. P. D. Jones, A. E. J. Ogilvie, T. D. Davies and K. R. Briffa. Boston: Kluwer Academic. [2](#), [10](#), [12](#), [35](#)
- Venkatraman, E. S. and Adam B. Olshen. 2007. “A Faster Circular Binary Segmentation Algorithm for the Analysis of Array CGH Data.” *Bioinformatics* 23:657–663. [2](#), [15](#), [27](#)
- Zeileis, Achim, Friedrich Leisch, Kurt Hornik and Christian Kleiber. 2002. “strucchange: An R Package for Testing for Structural Change in Linear Regression Models.” *Journal of Statistical Software* 7(2):1–38. [15](#)

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS

- [WP11/12](#) Cormac Ó Gráda: 'Five Crises' June 2011
- [WP11/13](#) Alan Fernihough: 'Human Capital and the Quantity-Quality Trade-Off during the Demographic Transition: New Evidence from Ireland' July 2011
- [WP11/14](#) Olivier Bargain, Kristian Orsini and Andreas Peichl: 'Labor Supply Elasticities in Europe and the US' July 2011
- [WP11/15](#) Christian Bauer, Ronald B Davies and Andreas Haufler: 'Economic Integration and the Optimal Corporate Tax Structure with Heterogeneous Firms' August 2011
- [WP11/16](#) Robert Gillanders: 'The Effects of Foreign Aid in Sub-Saharan Africa' August 2011
- [WP11/17](#) Morgan Kelly: 'A Note on the Size Distribution of Irish Mortgages' August 2011
- [WP11/18](#) Vincent Hogan, Patrick Massey and Shane Massey: 'Late Conversion: The Impact of Professionalism on European Rugby Union' September 2011
- [WP11/19](#) Wen Fan: 'Estimating the Return to College in Britain Using Regression and Propensity Score Matching' September 2011
- [WP11/20](#) Ronald B Davies and Amélie Guillin: 'How Far Away is an Intangible? Services FDI and Distance' September 2011
- [WP11/21](#) Bruce Blonigen and Matthew T Cole: 'Optimal Tariffs with FDI: The Evidence' September 2011
- [WP11/22](#) Alan Fernihough: 'Simple Logit and Probit Marginal Effects in R' October 2011
- [WP11/23](#) Ronald B Davies and Krishna Chaitanya Vadlamannati: 'A Race to the Bottom in Labour Standards? An Empirical Investigation' November 2011
- [WP11/24](#) Wen Fan: 'School Tenure and Student Achievement' November 2011
- [WP11/25](#) Mark E McGovern: 'Still Unequal at Birth - Birth Weight, Socioeconomic Status and Outcomes at Age 9' November 2011
- [WP11/26](#) Robert Gillanders: 'The Mental Health Cost of Corruption: Evidence from Sub-Saharan Africa' November 2011
- [WP11/27](#) Brendan Walsh: 'Well-being and Economic Conditions in Ireland' December 2011
- [WP11/28](#) Cormac Ó Gráda: 'Fetal Origins, Childhood Development, and Famine: A Bibliography and Literature Review' December 2011
- [WP12/01](#) Mark E McGovern: 'A Practical Introduction to Stata' January 2012
- [WP12/02](#) Colm McCarthy: 'Ireland's European Crisis: Staying Solvent in the Eurozone' January 2012
- [WP12/03](#) Colm McCarthy: 'Improving the Eurosystem for Old and New Members' January 2012
- [WP12/04](#) Ivan Pastine and Tuvana Pastine: 'All-Pay Contests with Constraints' February 2012
- [WP12/05](#) David Madden: 'Methods for Studying Dominance and Inequality in Population Health' February 2012
- [WP12/06](#) Karl Whelan: 'ELA, Promissory Notes and All That: The Fiscal Costs of Anglo Irish Bank' February 2012
- [WP12/07](#) Olivier Bargain, Eliane El Badaoui, Prudence Kwenda, Eric Strobl and Frank Walsh: 'The Formal Sector Wage Premium and Firm Size for Self-employed Workers' March 2012
- [WP12/08](#) Brendan Walsh: 'The Influence of Macroeconomic Conditions and Institutional Quality on National Levels of Life Satisfaction' March 2012
- [WP12/09](#) Ronald B Davies and Rodolphe Desbordes: 'Greenfield FDI and Skill Upgrading' March 2012