Appendix 1.

1. Precipitation Reconstruction

In order to guarantee the reliability of our reconstructed precipitation, we used 21 documentary-based single-proxy hydro-climate reconstructions (annual resolution) were synthesized from 13 published references, of which the data length is greater than 500 years to guarantee the reliability of reconstruction. By adopting the documentary-based, we can obtain relatively pure information on precipitation change. These documentary-based single-proxy hydro-climate reconstructions cover all the regions in China as shown in Table A1.

In addition, the study aims to uncover the relationship between climate change and nomadic migration in the long term. Therefore, the data length of proxies should be greater than 500 years. However, these 21 data series still have different data length. Following current practices in the paleo-climate reconstruction, a same interval of mean average on the overlapping periods is set in order to make all the series with different data length comparable in time (Mann *et al.* 2008; Marcott *et al.* 2013). In our study, the overlapping period is 1200-1900 AD. Therefore, based on this overlapping period of 1200-1900 AD, we normalized all the data series for reconstruction.

In China, there are seven natural divisions which are identified by similar moisture condition, soil and vegetation which depend on climate, landscape, geological history within their borders (Zhao 1986). Here is the Seven natural divisions of China (Zhao 1986):

1. East Asian Monsoon China
 1.1 Northeastern China (NE)
 1.2 Northern China (N)
 1.3 Central and Southern China (CS)
 1.4 Southern China (S)
2. Northwestern Arid China
 2.1 Inner Mongolia (IM)
 2.2 Northwestern China (NW)
3. Qinghai-Tibetan Frigid Plateau
 Qinghai-Tibetan Plateau (QT)

Table A1 Data source for precipitation anomaly reconstruction

	Data type	Time span	Spatial coverage of Natural Divisions	Resolution	Reference
1	Historical flood & drought records	AD 1120–1980	CS	Decadal	Chen (1987)
2	Historical records of lake surface area variations	1000 BC–AD 1989	CS+IM+N+NW+QT	Decadal	Fang (1993)
3	Historical flood & drought records	AD 230–1919	CS+IM+N+NW	Half-decadal	Gong & Hameed (1991)
4	Historical flood & drought records	AD 1000–1950	CS	Annual	Jiang *et al.* (2005)
5	Historical flood & drought records	AD 952–1997	IM+N+CS+S	Annual	Qian et al. (2003)
6	Historical flood & drought records	AD 960-2000	N+QT	Decadal	Tan *et al.* (2008)
7	Historical flood & drought records	AD 1–2000	N+QT	Decadal	Tan et al. (2010)
8	Historical flood & drought records	AD 1–1979	CS+N	Decadal	Yan et al.(1993)
9	Historical records of Yellow River floods	206 BC–AD 1938	N	Decadal	Yellow River Water Resources Committee (1982)
10	Historical dust–storm records	AD 300–1939	CS+IM+N+NE+NW	Decadal	Zhang (1984)
11	Historical flood & drought records	AD 1–1996	CS+N	Decadal	Zhang et al. (1997b)
12	Historical flood & drought records	AD 960–1992	CS+N	Annual	Zhang et al. (1997a)
13	Historical flood & drought records	AD 500–2000	CS+IM+N+NE	Annual	Zheng et al. (2006)

Then, according to Table A1, the adopted documentary-based single-proxies have covered all seven natural divisions. Therefore, the reconstructed precipitation could well represent the China-wide precipitation. The spatial coverage of selected proxies in the study is reliable to show the precipitation change on the national scale in China. As mentioned, our aim is to study the relationship between climate change and nomadic migration in the long term. Therefore, all these normalized data series with spatially coverage of whole China is arithmetically averaged. In this way, we believe that the reconstruction could show general implications of past precipitation change on a long term scale.

For the temperature reconstruction series in the study, a reconstructed temperature series Yang et al's (2002) from multi-proxies (e.g., ice cores, lake sediment, and so on) throughout China is adopted. Temperature is reconstructed based on decadal resolution and covers AD 1–2000. The temperature reconstruction covers whole China, which has the same spatial scale to our reconstructed precipitation. Furthermore, Yang et al's temperature anomaly has been adopted as a reliable climatic indicator to study the relationship between climate change and social responses in historical China (Zhang *et al.* 2005; Lee & Zhang 2013).

2. Granger Causality Analysis

Background

In 2003, Sir Clive William John Granger was awarded with the Nobel Memorial Prize in Economic Sciences, in recognition to his contributions to the discoveries in the analysis of time series data which fundamentally changes economic study on the data of time. GCA is one of his brilliant achievements, which was emphasized by him when he was giving the prize speech (Granger 2003). Although GCA stems from economics, it has also been used commonly in the empirical studies of psychology (Sobel 1995), politics (Umez 1993), sociology (Walker & Jackson 2007; Lin & Ali 2009), biology (Fujita *et al.* 2010), medicine (Florin *et al.* 2008), and so on.

Theoretical calculation process

Granger's notion of causality is that '… Y_t is causing X_t if we are better able to predict X_t using all available information than if the information apart from Y_t had been used.'(Granger 1969). The GCA proposes a two-variable causal model with two stationary time series, X_t and Y_t, with zero means:

(S1) $$X_t = \sum_{j=1}^{m} a_j X_{t-j} + \sum_{j=1}^{m} b_j Y_{t-j} + \varepsilon_t .$$

where, a and b are the coefficients of the time series. j is the data of the time series at time point j, and m is the length of the time series which is set based on lag. ε is the residue and t is the time step.

The Augmented Dickey–Fuller (ADF) test approach controls higher-order correlation by adding lagged difference terms of the dependent variable Y to the right-hand side of the regression, which can be written as the following equation (Agung 2009):

(S2) $$DY_t = \mu + \delta Y_{t-1} + \beta_1 DY_{t-1} + \beta_2 DY_{t-2} + ... + \beta_p DY_{t-p} + \varepsilon_t ,$$

where

(S3) $$DY_t = Y_t - Y_{t-1} .$$

The null hypothesis of the series $\{Y_t\}$ has a unit root, that is, H_0: δ=0.

where, β, δ is the coefficient of the time series. p is the data of the time series at time point p. D means the differencing. ε is the residue and t is the time step.

The maximum lag in ADF Test can be chosen based on the following statistical formula (Hayashi 2000), which is also adopted by EViews as a default:

(S4) $$Lag_{Max} = \text{int}\left[12 \times (\frac{T}{100})^{0.25}\right] .$$

where T is sample size and *int* means integer.

After applying the ADF test on the stationarity status of each data series, the lag length should be selected for the GCA. In the model, the lags of X_t and Y_t are set equally. Given that the t or F statistic is only a function, it depends on the correlation between the two variables and the set of conditioning variables. If the lag is set the same, then the same conditioning variables can be included (Kirchgässner & Wolters 2007). Akaike's information criterion (AIC) is adopted to determine the appropriate lag length (Akaike 1974) as the statistical criteria.:

$$\text{(S5)} \qquad AIC(m) = 2m - 2\ln(L) \qquad 1 \le m \le m_{\max}$$

where L is the maximum likelihood achievable by the model, m is the number of parameters of the model, and N is the number of data points used in the fit. m_{max} is the maximum lag, which can be chosen based on statistical Formula A4. We obtain the m for the AIC lag in the ADF test and then apply m to set the lag in Formula S1 for the GCA.

3. ADF Test Results and Lag Selections

Table A2 Time lag in Granger causality analysis on "climate change → nomadic migration → conflicts between pastoralists and agriculturalists"

Null hypothesis	Lag
Precipitation does not (Granger) cause nomadic migration.	1#
Temperature does not (Granger) cause nomadic migration.	1#
Nomadic migration does not (Granger) cause conflict number.	21△
Conflict number does not (Granger) cause nomadic migration.	25△

Note: # means the instantaneous lag for GCA. △ means the AIC lag for GCA.

Table A3 Augmented Dickey–Fuller test in Granger causality analysis of "climate change → nomadic migration → conflicts between pastoralists and agriculturalists"

Series	Prob.	Difference level
Precipitation	0.000#	No difference
Temperature	0.001#	No difference
Nomadic migration	0.000#	No difference
Conflict number	0.000△	No difference
Nomadic migration	0.000△	No difference

Note: # means the instantaneous lag for GCA. △ means the AIC lag for GCA.

Table A4 Time lag in Granger causality analysis of precipitation and northward/southward migration

Null hypothesis	Lag
Precipitation does not (Granger) cause northward migration. #	1#
Precipitation does not (Granger) cause southward migration. #	1#

Note: # means the instantaneous lag for GCA.

Table A5 Augmented Dickey–Fuller test in Granger causality analysis of precipitation and northward/southward migration

Series	Prob.	Difference level
Precipitation	0.000#	No difference
Northward migration	0.000#	No difference
Southward migration	0.000#	No difference

Note: # means the instantaneous lag for GCA.

Table A6 Augmented Dickey–Fuller test in Granger causality analysis of precipitation and conflicts between pastoralists and agriculturalists

Series	Prob.	Difference level
Precipitation	0.002	No difference
Conflict number	0.000	No difference

Table A7 Time lag in Granger causality analysis of precipitation and northward/southward migration

Null hypothesis	Lag
Precipitation does not (Granger) cause conflict number.	21

References

Agung I.G.N. (2009). *Time Series Data Analysis Using Eviews*. John Wiley & Sons (Asia) Pte Ltd, Singapore.

Akaike H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.

Chen J.Q. (1987). An approach to the data processing of historical climate materials on the basis of floods and droughts of Tai Hu Basin. *Acta Geographica Sinica*, 42, 231-42.

Fang J. (1993). Lake Evolution during the Last 3000 Years in China and Its Implications for Environmental Change. *Quaternary Research*, 39, 175-185.

Florin E., Reck C., Burghaus L., Lehrke R., Gross J., Sturm V., Fink G.R. & Timmermann L. (2008). Ten Hertz thalamus stimulation increases tremor activity in the subthalamic nucleus in a patient with Parkinson's disease. *Clinical Neurophysiology*, 119, 2098-2103.

Fujita A., Severino P., Sato J. & Miyano S. (2010). Granger Causality in Systems Biology: Modeling Gene Networks in Time Series Microarray Data Using Vector Autoregressive Models. In: *Advances in Bioinformatics and Computational Biology* (eds. Ferreira C, Miyano S & Stadler P). Springer Berlin / Heidelberg, pp. 13-24.

Gong G. & Hameed S. (1991). The variation of moisture conditions in China during the last 2000 years. *International Journal of Climatology*, 11, 271-283.

Granger C.W.J. (1969). Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*, 37, 424-438.

Granger C.W.J. (2003). Time Series Analysis, Cointegration, and Applications. *The American Economic Review*, 94, 421-425.

Jiang T., Zhang Q., Blender R. & Fraedrich K. (2005). Yangtze Delta floods and droughts of the last millennium:abrupt changes and long term memory. *Theoretical and Applied Climatology*, 82.

Kirchgässner G. & Wolters J. (2007). *Introduction to Modern Time Series Analysis*. Springer, Verlag Berlin Heidelberg.

Lee H.F. & Zhang D.D. (2013). A tale of two population crises in recent Chinese history. *Climatic Change*, 116, 285-308.

Lin E.S. & Ali H.E. (2009). Military Spending and Inequality: Panel Granger Causality Test. *Journal of Peace Research*, 46, 671-685.

Mann M.E., Zhang Z., Hughes M.K., Bradley R.S., Miller S.K., Rutherford S. & Ni F. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proceedings of the National Academy of Sciences*, 105, 13252-13257.

Marcott S.A., Shakun J.D., Clark P.U. & Mix A.C. (2013). A Reconstruction of Regional and Global Temperature for the Past 11,300 Years. *Science*, 339, 1198-1201.

Qian W., Hu Q., Zhu Y. & Lee D.K. (2003). Centennial-scale dry-wet variations in East Asia. *Climate Dynamics*, 21, 77-89.

Sobel M.E. (1995). Causal Inference in the Social and Behavioral Sciences. In: *Handbook of statistical modeling for the social and behavioral sciences* (eds. Arminger G, Clogg CC & Sobel ME). Plenum Press New York, p. 14.

Tan L., Cai Y. & An Z. (2010). Precipitation variations of Longxi over the last 2000 years and the possible driving forces. *Journal of Arid Land Resources and Environment*, 24, 99-116.

Tan L., Cai Y., Yi L., An Z. & Ai L. (2008). Precipitation variations of Longxi, northeast margin of Tibetan Plateau since AD 960 and their relationship with solar activity. *Climate of the Past*, 4,

19-28.
Umez B. (1993). Has social mobilization caused political instability in Africa? A granger-causality test. *The Review of Black Political Economy*, 22, 33-54.
Walker D.M. & Jackson J.D. (2007). Do Casinos Cause Economic Growth? *American Journal of Economics and Sociology*, 66, 593–607.
Yan Z., Li Z. & Wang X. (1993). An analysis of decade-to-century-scale climatic jumps in history. *Scientia Atmospherica Sinica*, 17, 663-672.
Yang B., Braeuning A., Johnson K.R. & Yafeng S. (2002). General characteristics of temperature variation in China during the last two millennia. *Geophysical Research Letters*, 29, 1324.
Yellow River Water Resources Committee (1982). *Prospective Summarization of Yellow River Water Resources*. Yellow River Water Resources Press, Beijing.
Zhang D. (1984). Preliminary analysis of dust storm and its associated weather in the historical periods of China. *Scientia Sinica*, 3, 278-288.
Zhang D., Jim C., Lin C., He Y. & Lee F. (2005). Climate change, social unrest and dynastic transition in ancient China. *Chinese Science Bulletin*, 50, 137-144.
Zhang D., Liu C. & Jiang J. (1997a). Reconstruction of six regional dry/wet series and their abrupt changes during the last 1000 years in East China. *Quaternary Sciences*, 1, 1-10.
Zhang P., Ge Q., Zhang S. & Liu X. (1997b). The modes and abrupt changes of climate in China during recent 2000 years. *Quaternary Sciences*, 12-20.
Zhao S. (1986). *Physical Geography of China*. Science Press, John Wiley & Sons, Beijing.
Zheng J., Wang W., Ge Q., Man Z. & Zhang P. (2006). Precipitation variability and extreme events in Eastern China during the past 1500 years. *Terrestrial Atmospheric and Oceanic Sciences*, 17, 579-92.