nature human behaviour

ARTICLES
https://doi.org/10.1038/s41562-019-0750-z

Historical analysis of national subjective wellbeing using millions of digitized books

Thomas T. Hills[1,2*], Eugenio Proto[3,4,6], Daniel Sgroi[3,5] and Chanuki Illushka Seresinhe[2]

In addition to improving quality of life, higher subjective wellbeing leads to fewer health problems and higher productivity, making subjective wellbeing a focal issue among researchers and governments. However, it is difficult to estimate how happy people were during previous centuries. Here we show that a method based on the quantitative analysis of natural language published over the past 200 years captures reliable patterns in historical subjective wellbeing. Using sentiment analysis on the basis of psychological valence norms, we compute a national valence index for the United Kingdom, the United States, Germany and Italy, indicating relative happiness in response to national and international wars and in comparison to historical trends in longevity and gross domestic product. We validate our method using Eurobarometer survey data from the 1970s and demonstrate robustness using words with stable historical meanings, diverse corpora (newspapers, magazines and books) and additional word norms. By providing a window on quantitative historical psychology, this approach could inform policy and economic history.

Investigations of subjective wellbeing span the social-science disciplines[1–5]. Subjective wellbeing has also become the focus of governments and international organizations who see it as an important target for government policy alongside the more traditional focus on national income. For example, the United Nations released the first World Happiness Report in 2011 alongside the Organisation for Economic Co-operation and Development (OECD) launch of the Better Life Index. Unfortunately, compared with data about national income—which have been collected since the 1930s in many nations—analyses of subjective wellbeing suffers from a considerable shortfall in the availability of long-run data. Historical approaches have computed national income statistics as far back as 1820, courtesy of the Maddison Project[6], and we have centuries of additional data for some nations[7]. By contrast, consistent measures of subjective wellbeing have been collected only since the 1970s.

Our goal here is to present and validate a reliable historical measure of national subjective wellbeing going back 200 years, enabling direct comparisons with GDP and other long-run data, such as longevity, internal conflict and democratization. To do this, we derived a National Valence Index (NVI) from the words used in historical texts. In addition to other corpora that are described below, our main index was drawn from the Google Books corpus[8], which is a collection of word frequency data for over 8 million books, providing a digitized historical record of more than 6% of all of the books that have been physically published[9]. We use the words published in these books to compute subjective wellbeing at a given time by using affective word norms to derive sentiment from text. Affective word norms are ratings provided by groups of individuals who examine a list of words and rate them on their valence, indicating how good or bad individual words make them feel. Using these ratings, we worked through millions of books, enumerating the complete published list of Google Books by year and by language. Here we present these data for four countries: United States, United Kingdom, Germany and Italy.

Our approach has previously been shown to enable the prediction of economic, political and cultural trends, including recovering large-scale opinions about political candidates[10], predicting stock market trends[11], understanding diurnal and seasonal mood variation[12], detecting the social spread of collective emotions[13], predicting depression in medical patients[14] and understanding the impact of events with the potential for large-scale societal effects, such as death of celebrities, earthquakes and economic bailouts[15,16]. Our specific approach is directly supported by a study of 17 million blog posts[17], which found that a simple calculation on the basis of the weighted affective ratings of words was highly accurate (70%) at predicting the mood of blogs as provided by the bloggers themselves. Words with positive valence are therefore taken to indicate positive connotations for the subjective wellbeing of the user, and those with negative valence are taken to have an equivalent negative connotation. This might not be true for any individual chosen at random or for any individual word in context, but the power of large data is that idiosyncratic noise is averaged out when averaging over many authors and words.

Results

Comparison of the NVI with survey-based measures of wellbeing. To validate the NVI, we first compared it with existing survey-based measures of subjective wellbeing. The measure of life satisfaction that we took as the ground truth was the average per year and per country data obtained from the Eurobarometer survey, which was conducted by the European Commission. The question answered was 'On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?', coded on a four-point scale from 'very satisfied' to 'not at all satisfied'. To facilitate the reader's intuition, we reversed the code so that a lower number corresponded to less satisfaction. The Eurobarometer survey is the oldest survey available that is representative of the countries that we use. The first wave covers each year dating back to 1973. It contains data from the United Kingdom (104,068 interviews),

[1]Department of Psychology and CAGE, University of Warwick, Coventry, UK. [2]The Alan Turing Institute, British Library, London, UK. [3]IZA Institute of Labor Economics, Bonn, Germany. [4]CESifo, Munich, Germany. [5]Department of Economics and CAGE, University of Warwick, IZA, Coventry, UK. [6]Present address: Adam Smith Business School, University of Glasgow, Glasgow, UK. *e-mail: t.t.hills@warwick.ac.uk

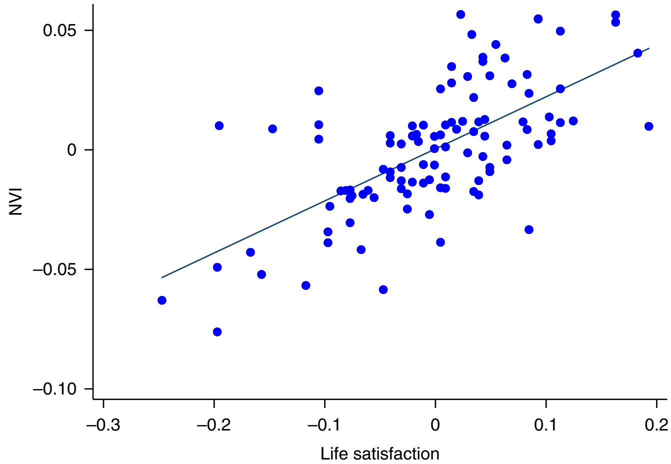


Fig. 1 | Correlation of the NVI and aggregate life satisfaction data from the Eurobarometer survey. The NVI (our measure of subjective wellbeing derived from digitized text) is compared with aggregate life satisfaction (obtained from the Eurobarometer survey-based measure) for the United Kingdom, Germany and Italy (the three countries for which both measures exist) from 1973 to 2009 (the period over which both measures are available). Both variables (the NVI and Eurobarometer life satisfaction measures) are expressed in the form of residuals after controlling for country fixed effects so that values represent variations around the averages for each of the three countries.

Table 1 | The NVI predicts aggregate life satisfaction

	Year fixed effects	Country-specific trends
NVI (β (s.e.))	2.8551*** (0.2867)	1.6596** (0.2246)
GDP	Yes	Yes
Country-specific trend	No	Yes
Year fixed effects	Yes	No
r^2	0.730	0.588
n	104	104

The NVI is a statistically significant predictor in an ordinary least squares estimate with country fixed effects of aggregate life satisfaction. The dependent variable is average life satisfaction per country and year, obtained from the Eurobarometer survey-based measure. The period covered is 1973 to 2009, the period over which both measures exist. The countries considered are Germany, Italy and the United Kingdom, the three countries for which both datasets exist. GDP per capita (expressed in terms of purchasing power parity) was obtained from the PWT 8.0 dataset. Column 1 includes year fixed effects (to help to deal with spurious correlations over time) and column 2 includes country-specific trends (to help to deal with spurious correlations across countries). Robust standard errors clustered at country levels are given in brackets. $^{**}P<0.05$, $^{***}P<0.01$. Full statistical information for this table is provided in the Supplementary Information.

Germany (102,795 interviews, with only West Germany covered before 1990) and Italy (103,789 interviews). Figure 1 shows the relationship between the NVI and aggregate life satisfaction derived from Eurobarometer data for the corresponding country. The data are presented in the form of residuals after controlling for country fixed effects and have a positive Pearson's correlation of $r(102)=0.53$, $P<0.001$, 95% confidence interval (CI) $=0.37$–0.66. Although the United States is not included in the Eurobarometer survey, there are fragmented life satisfaction data available from the World Database of Happiness that are positively correlated with the NVI; these data are provided in Supplementary Fig. 5.

The analysis presented in Table 1 shows that the positive relationship is robust to the introduction of GDP, using GDP per capita data obtained from the Penn dataset (version PWT 8.0)[18], in which data are in 2005 international dollars and are adjusted for purchasing power parity. Moreover, the inclusion of year fixed effects (Table 1, column 1), to control for the possibility of biases generated by shocks common to all of the countries in the dataset, and country-specific trends (Table 1, column 2), do not qualitatively change the main result. In all cases, the coefficient on the NVI is positive and significant. NVI is also a better predictor of Eurobarometer life satisfaction than either GDP (Pearson's $r(102)=0.36$, $P<0.001$, 95% CI $=0.17$–0.52) or life expectancy (Pearson's $r(102)=0.15$, $P=0.12$, 95% CI $=-0.04$–0.34); each of these variables was computed after controlling for country fixed effects. Supplementary Table 2 provides an alternative approach and robustness check by comparing the annual change in Eurobarometer life satisfaction and the annual change in NVI and GDP.

As a further validation, consider the following nonparametric exercise: if our measure is valid, then the average valence of all of the words (obtained from the Google Books corpus for each of the three countries) that have a frequency correlated significantly and positively with life satisfaction (obtained from the Eurobarometer data) should be significantly higher than the average valence of the words that have a frequency correlated significantly and negatively. We find that this is indeed the case. The analysis is described in more detail in the Supplementary Information and the results are provided in Supplementary Fig. 3.

Historical analysis. In Fig. 2, we show the NVI for the United Kingdom, the United States, Germany and Italy from 1820 to 2009, which is the last year that is currently available from the Google Books corpus. The red vertical lines represent key political events in each country as indicated in the figure caption. The relationship of these political events with NVI supports a contemporary historical understanding of the impacts of these events on subjective wellbeing. Internal conflicts, such the American Civil War, the 1848 Year of Revolution in Europe and the two World Wars, coincide with decreases in the NVI for the countries affected. The peak in the US data in the 1920s followed by a downward trend after the Wall Street crash in 1929 supports the view that the crash followed a period of over-optimism in response to sustained economic prosperity. As our data are drawn from published text, it may be subject to censorship. Although we use the Polity IV democracy variable in Table 2 in an effort to control for this, no control can be perfect; the case of Germany in the 1940s, when negative portrayals of the Nazi regime were censored, is a case in point. Our prior knowledge of this censorship suggests an overstatement of the NVI during that time in Germany. The democracy variable is provided in Supplementary Table 3.

In Table 2, we compare the NVI to the two welfare indicators for which the longest series of data are available—GDP and life expectancy at birth—with both showing a positive relationship with the NVI (Table 2, columns 1, 2 and 3). We also investigated the effect of internal conflicts, which show a negative relationship that is consistent with the data in Fig. 2 and as we would expect (Table 2, column 4). To account for potential lags between changes in the key variables and the appearance of their influence in published text, we empirically determined the lags for each variable on the basis of their influence on the NVI (details are provided in section 2 of the Supplementary Information and Supplementary Tables 4–6).

As is common in the economics literature, for historical GDP, we used data from the Maddison Project (http://www.ggdc.net/maddison/maddison-project/home.htm, 2013 version) in which data are in 1990 international dollars. The results presented in Table 1 do not change qualitatively if we use data from after 1972 from the Maddison Project instead of from the Penn dataset. The other main explanatory variables are the historical data on life expectancy at birth and on internal conflict—which indicates each year of major conflicts that directly affect the domestic population, such as internal unrest or invasions—both of which were obtained from the OECD and include data from 1820 onwards[19].

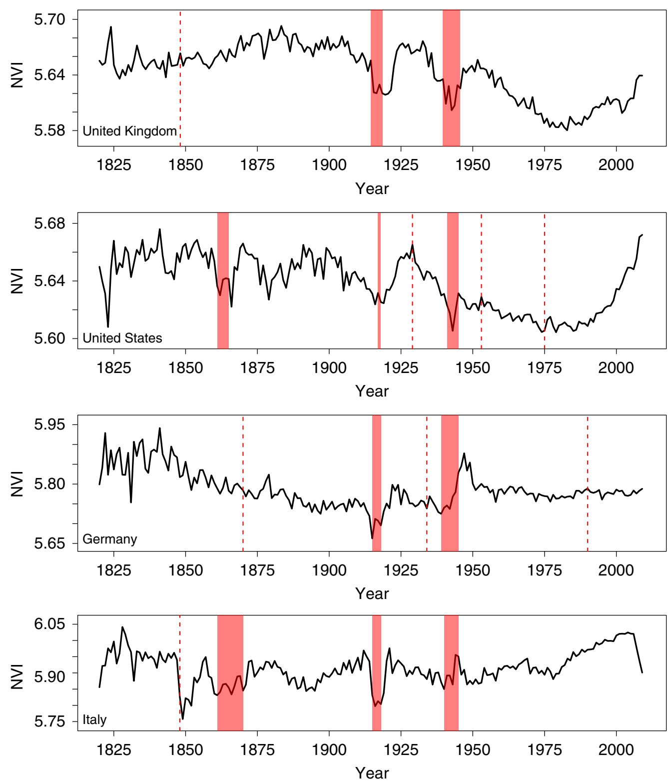


Fig. 2 | NVI through the period 1820–2009. The NVI from 1820 to 2009. Various important events are highlighted in red (for periods of time) or with a dashed vertical red line for events that correspond to a single year. For all countries, the red shaded lines include World War I (approximately 1914–1918) and World War II (approximately 1938–1945). In the three European countries, the line in 1848 indicates the Year of Revolution. In the United States, there is an additional shaded area that represents the Civil War (1861–1865) and vertical red lines that represent the Wall Street Crash (1929), the end of the Korean War (1953) and the fall of Saigon (1975). For Germany, the vertical red lines represent the end of Franco–Prussian War and reunification (1870), Hitler's ascendency to power (1934) and the reunification (1990). In Italy, there is an additional shaded area that represents the unification (1861–1870).

Other variables that we used as controls include educational inequality (measured as a Gini index, which we used as a proxy for the inclusivity of the demand for books within society) and the index of democracy (from the Polity IV project of the OECD)[19]. Finally, we introduced the share of words in the corpora for which we have the valence measures. The data are further summarized in Supplementary Table 1.

The NVI is probably affected by the market for literature and, more generally, by the evolution of literature and language (Supplementary Information). Over the long run, as the target for a typical published book moved from the wealthy elite to the general public, the content of these books changes. Moreover, patterns in literary style changed considerably in the early part of the nineteenth century with the advent of literary realism and social commentary. To help to deal with problems of this sort, we included control variables that were specifically chosen to correct for year-on-year trends. This is reflected in two alternative econometric specifications presented in Table 2 that correspond to two different hypotheses on the evolution of literature and language. One model, which controlled for year fixed effects (Table 2, columns 1, 2 and 3), assumes that the market for books and language itself evolved in a similar way across the different countries that we considered and controls for this change. The other model (Table 2, column 4), which introduces country-specific trends, assumes that the evolution of the market for books and of language itself affects written texts of different languages differently. Therefore, by including country-specific trends, we corrected any source of bias to the extent that it generates roughly linear trends. Our results show that these two models generate similar findings. Note also that as year fixed effects are potentially correlated with the years in which internal conflicts took place owing to the likelihood of spillover effects from such conflicts from one country to another, we cannot include both year fixed effects and a measure of major conflict in the same regression. We therefore introduced internal conflicts only in column 4, in which the model with country-specific trends is presented.

Looking more closely at the results presented in Table 2, we note that in column 1 and 2, the effects of GDP per capita and life expectancy are both positive and significant, respectively. In column 3, in which we introduced both simultaneously, the effect of life expectancy becomes smaller and non-significant, which probably reflects the high level of collinearity between the two variables. The internal conflict variable in column 4 is negative and significant. We also performed an analysis of possible stochastic trends (including appropriate augmented Dicky–Fuller tests) that might affect the regressions presented in Supplementary Tables 1 and 2, further supporting the results presented here.

A key contribution of the NVI is the ability to quantify historical indicators of psychological wellbeing. For example, Table 2 enabled us to compute that one extra year of life expectancy is worth as much as 4.3% annual growth in GDP per capita. Letting $\Delta\log[\mathrm{GDP}(t-5)]$ and Δlife expectancy$(t-1)$ indicate a change in GDP and life expectancy, respective, then from column 3 of Table 2, $0.0698\times\Delta\log[\mathrm{GDP}(t-5)]=0.0030\times\Delta\text{life expectancy}(t-1)$, such that when Δlife expectancy$(t-1)=1$, $\Delta\log[\mathrm{GDP}(t-5)]=0.043$. One fewer year of internal conflict is worth as much as 30% annual growth in GDP per capita. From column 4 of Table 2, $0.0550\times\Delta\log[\mathrm{GDP}(t-5)]=-0.0184\times\Delta\text{years of conflict}$, so that $\Delta\log[\mathrm{GDP}(t-5)]=0.33$ for each year of conflict. More generally, the results in this section largely follow our intuitions about the probable impact of historical changes on subjective wellbeing, while also providing a quantitative measure of their relative impact as a basis for future inquiry.

Discussion

Using conventional regression analysis and nonparametric methods, we show that the NVI is highly consistent with existing wellbeing measures going back to 1973. This indicates that, on average, the valence enshrined in literature matches the mood of the population as represented in published books. We further validate our measure by showing a relationship with variables that are known to have a relationship with wellbeing, such as conflict, life expectancy and GDP, going back to 1820.

The NVI highlights a number of interesting patterns. For example, there is a rise in subjective wellbeing in Italy and Germany since the 1900s matched by a comparative decline in the United Kingdom and the United States. However, since the 1970s, all four nations—with the possible exception of Germany—have seen a steady rise in subjective wellbeing. Internal and external conflicts represent considerable shocks to subjective wellbeing, but people tend to bounce back following these shocks even if they do not always return to their precise prior levels. These observations currently stand as hypotheses; however, the NVI enables these types of observations by presenting psychological history in a format that is available for explanation. An extended overview of how the NVI has changed in response to major historical events is provided in the Supplementary Information.

It is worth commenting on the relationship between the NVI and historical GDP in light of the controversy surrounding the link between national income and national happiness, referred to as the

Table 2 | Historical determinants of the NVI from 1820 to 2009

	Year fixed effects	Year fixed effects	Year fixed effects	Country-specific trends
log[GDP($t-5$)] (β (s.e.))	0.0826*** (0.0090)		0.0698*** (0.0106)	0.0550** (0.0130)
Life expectancy ($t-1$) (β (s.e.))		0.0048** (0.0013)	0.0030 (0.0014)	0.0016 (0.0013)
Internal conflict ($t-1$) (β (s.e.))				−0.0184** (0.0040)
Words covered (t)	Yes	Yes	Yes	Yes
Democracy (t)	Yes	Yes	Yes	Yes
Educational inequality (t)	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	No
Country-specific trends	No	No	No	Yes
r^2	0.752	0.705	0.774	0.571
n	412	412	412	412

Ordinary least squares regression with country fixed effects of the NVI. The countries included are Germany, Italy, the United Kingdom and the United States. The explanatory variables include GDP per capita (in 1990 international dollars, obtained from the Maddison Project), words covered (the percentage of all of the words that are included in the text-derived valence measure) and a variety of measures provided by the OECD, including life expectancy from birth, internal conflict (such as civil wars, revolutions and internal unrest), democracy and educational inequality (which offers a control for literacy). Lags are indicated in relation to time, t. The estimation controls for year fixed effects in columns 1–3 (to help to deal with spurious correlations over time) and country-specific trends in column 4 (to help to deal with spurious correlations across countries). Internal conflict is not included when year fixed effects are controlled because year fixed effects are heavily correlated with the years in which internal conflicts took place and there are likely to be spillover effects from such conflicts from one country to another. The lags of the regressors are empirically determined and details are provided in the Supplementary Information. Robust standard errors clustered at country levels are given in brackets. Full statistical information for this table is provided in the Supplementary Information.

Easterlin paradox[3,20,21]. According to the Easterlin paradox, happiness changes in direct response to temporary changes in income both within and between nations, but does not show long-term trends upwards with rising national income. In our analysis, we found a positive relationship between GDP and NVI as a function of localized change. However, the size of the coefficient is relatively small; a substantial increase in GDP over a short time period would be needed to generate a significant increase in the NVI. Our time series do not feature any clear trends in long-term wellbeing despite the well-known steady increase of GDP in all of the countries over the period considered here. This reinforces the point that the overall impact of GDP is relatively small and subjectively relative to historically recent events. As we did find a significant relationship between NVI and GDP, albeit quite small in size, we consider that our findings are neither inconsistent with the Easterlin paradox nor with studies that found a significant relationship between GDP and subjective wellbeing. However, it is worth noting that Easterlin's key work usually considers happiness rather than life satisfaction, whereas our findings are based primarily on life satisfaction. Importantly, life satisfaction may react more to income than measures such as happiness[22].

Language evolved considerably throughout the period considered here, and it changes according to who is writing and the markets that drive information evolution more generally[23–25]. This problem is similar, in essence, to the problem of comparing economic growth and income levels across many centuries when lifestyles have changed beyond recognition. In the Supplementary Information, we show that the results presented here are corroborated by alternative indices created from other independent corpora including 'Find My Past' data from the British Library's 'British Newspaper Project', which covers 65 million newspaper and periodical articles from the United Kingdom across 200 periodicals going back to 1710, the US English Corpus of Historical American English, which includes 400 million words from 1810–2000 and two alternative indices derived from SenticNet data, pleasantness and polarity.

Caution is needed when considering any long-run socioeconomic data. In all cases, there is a need for what historians call a 'close read' of the historical data. Nonetheless, the utility of having long-run data is hard to overstate. Consider, for example, urbanization, cultural and political dynamics, increased technological advances (such as mechanization, computerization, mobile telephony and the Internet) and countless other important changes, all of which have made intertemporal comparisons of national income challenging but have not prevented the development and widespread use of historical measures of GDP to inform the influences on and impact of our economic history[26]. By generating an economic indicator of historical subjective wellbeing, we provide a measure of quantitative psychological history to the list of important economic indicators.

Methods

Historical corpora. We used the largest available sets of affective word norms for four languages: English (British), English (American), German and Italian. To enable comparison across languages, all of our valence norms contain a subset of approximately 1,000 words adapted from the 'Affective Norms for English Words'[27] (ANEW), which are words chosen, in part, because they capture the range of emotional sentiment. The original ANEW list served as the basis for developing valence ratings for each of the other languages in our study. Here we exclusively used the mean valence rating of words. The Google Books corpus also includes additional languages. For example, French and Spanish are included in the corpus and valence is available for these two languages, but our ability to draw sensible inferences for these countries was hampered by the market for books in French and Spanish outside France and Spain. In Supplementary Fig. 1, we present a sample of the words covered in all of the languages that we considered. In the Supplementary Information, we also show that our results are supported by alternative methods for computing historical sentiment, including using only the most stable historical words (which are more resistant to changes in meaning over time), computing time-locked valences for each word and using independent valence norms from the alternative AFINN[28] word norms (Supplementary Tables 7–11, Supplementary Fig. 9).

Valence norms. For English, we used the affective rating norms[29]. These norms are a database of nearly 14,000 English words, all rated on a valence scale of 1 to 9. Each word was rated by 20 participants and the mean valence rating was used for the purpose of our study. These ratings have high reliability and represent an important resource within psychology. The 14,000 words in the database contain a subset of the 1,034 ANEW words. For German, we used the affective norms for German sentiment terms[30]—a list of 1,003 words—and German translations of the ANEW list. The valence ratings were collected on a scale of −3 to +3. The mean values were adjusted to reflect a scale of 1 to 9 in our analysis. For Italian, we used an adaptation of the ANEW norms[31], which contains 1,121 Italian words; as with the English words, the ratings were collected on a scale of 1 to 9.

The NVI. Using our historical record and word valences, for each language i we computed the NVI, $\mathrm{NVI}_{i,t}$, for each year t and language i as follows

$$\mathrm{NVI}_{i,t} = \sum_{j=1}^{n} v_{j,i} p_{j,i,t}$$

where $v_{j,i}$ is the valence for word j in language i, and $p_{j,i,t}$ is the proportion of word j in year t for the language i. The proportion is computed each year over the words in the corpus for which we have valence ratings. Although the Google Books database includes books from 1500 to 2009, the number of books included for the first three centuries is fairly sparse. We limited our analyses to the period from 1820 to 2009, for which sufficient data are available[9,32]. In addition, the complete series of data on national income collected in the Maddison project[6]—which we used to validate our measure—uses 1820 as a benchmark year.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability
The data necessary to reproduce the analyses presented in this article are provided at https://github.com/warwickpsych/NationalValenceIndex.

Code availability
The code necessary to reproduce the analyses presented in this article is provided at https://github.com/warwickpsych/NationalValenceIndex.

Received: 20 April 2018; Accepted: 30 August 2019;
Published online: 14 October 2019

References
1. Di Tella, R., MacCulloch, R. J. & Oswald, A. J. Preferences over inflation and unemployment: evidence from surveys of happiness. *Am. Econ. Rev.* **91**, 335–341 (2001).
2. Deaton, A. Income, health, and well-being around the world: evidence from the Gallup world poll. *J. Econ. Perspect.* **22**, 53–72 (2008).
3. Stevenson, B. & Wolfers, J. Economic growth and subjective well-being: reassessing the Easterlin paradox. *Brookings Pap. Econ. Ac.* **39**, 1–87 (2008).
4. Benjamin, D. J., Kimball, M. S., Heffetz, O. & Rees-Jones, A. What do you think would make you happier? What do you think you would choose? *Am. Econ. Rev.* **102**, 2083–2110 (2012).
5. Proto, E. & Rustichini, A. A reassessment of the relationship between GDP and life satisfaction. *PLoS One* **8**, e79358 (2013).
6. Bolt, J. & van Zanden, J. L. The Maddison Project: collaborative research on historical national accounts. *Econ. Hist. Rev.* **67**, 627–651 (2014).
7. Broadberry, S., Campbell, B., Klein, A., Overton, M. & Van Leeuwen, B. *British Economic Growth, 1270–1870: An Output-Based Approach* School of Economics Discussion Paper http://hdl.handle.net/10871/13984 (University of Exeter, 2012).
8. Lin, Y. et al. Syntactic annotations for the Google Books Ngram Corpus. In *Proc. ACL 2012 System Demonstrations* 169–174 (Association for Computational Linguistics, 2012).
9. Michel, J.-B. et al. Quantitative analysis of culture using millions of digitized books. *Science* **331**, 176–182 (2011).
10. Connor, B., Balasubramanyan, R., Routledge, B. R. & Smith, N. A. From tweets to polls: linking text sentiment to public opinion time series. *International Conference on Web and Social Media, North America (ICWSM)* **11**, 122–129 (AAAI, 2010).
11. Bollen, J., Mao, H. & Zeng, X. Twitter mood predicts the stock market. *J. Comput. Sci.* **2**, 1–8 (2011).
12. Golder, S. A. & Macy, M. W. Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* **333**, 1878–1881 (2011).
13. Chmiel, A. et al. Collective emotions online and their influence on community life. *PLoS One* **6**, e22207 (2011).
14. Eichstaedt, J. et al. Facebook language predicts depression in medical records. *Proc. Natl Acad. Sci. USA* **115**, 11203–11208 (2018).
15. Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A. & Danforth, C. M. Temporal patterns of happiness and information in a global social network: hedonometrics and Twitter. *PLoS One* **6**, e26752 (2011).
16. Thelwall, M., Buckley, K. & Paltoglou, G. Sentiment in Twitter events. *J. Am. Soc. Inf. Sci. Technol.* **62**, 406–418 (2011).
17. Nguyen, T., Phung, D., Adams, B., Tran, T. & Venkatesh, S. Classification and pattern discovery of mood in weblogs. *Adv. Knowl. Discov. Data Min.* **6119**, 283–290 (2010).
18. Feenstra, R. C., Inklaar, R. & Timmer, M. P. The next generation of the Penn World Table. *Am. Econ. Rev.* **105**, 3150–3182 (2015).
19. van Zanden, J. L. et al. *How Was Life? Global Well-Being Since 1820* (OECD Publishing, 2014).
20. Easterlin, R. A. Does economic growth improve the human lot? Some empirical evidence. *Nations Househ. Econ. Growth* **89**, 89–125 (1974).
21. Easterlin, R. A., McVey, L. A., Switek, M., Sawangfa, O. & Zweig, J. S. The happiness–income paradox revisited. *Proc. Natl Acad. Sci. USA* **107**, 22463–22468 (2010).
22. Kahneman, D. & Deaton, A. High income improves evaluation of life but not emotional well-being. *Proc. Natl Acad. Sci. USA* **107**, 16489–16493 (2010).
23. Hills, T. T. The dark side of information proliferation. *Perspect. Psychol. Sci.* **14**, 323–330 (2019).
24. Li, Y., Engelthaler, T., Siew, C. S. & Hills, T. T. The macroscope: a tool for examining the historical structure of language. *Behav. Res. Methods* **51**, 1864–1877 (2019).
25. Hills, T. T. & Adelman, J. S. Recent evolution of learnability in American English from 1800 to 2000. *Cognition* **143**, 87–92 (2015).
26. Jerven, M. An unlevel playing field: national income estimates and reciprocal comparison in global economic history. *J. Glob. Hist.* **7**, 107–128 (2012).
27. Bradley, M. M. & Lang, P. J. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings* Technical Report C-1 (The Center for Research in Psychophysiology, Univ. Florida, 1999).
28. Nielsen, F. Å. A new ANEW: evaluation of a word list for sentiment analysis in microblogs. in *Proceedings of the ESWC2011 Workshop on 'Making Sense of Microposts': Big things come in small packages* 93–98 http://arxiv.org/abs/1103.2903 (2011).
29. Warriner, A. B., Kuperman, V. & Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* **45**, 1191–1207 (2013).
30. Schmidtke, D. S., Schröder, T., Jacobs, A. M. & Conrad, M. ANGST: affective norms for German sentiment terms, derived from the Affective Norms for English Words. *Behav. Res. Methods* **46**, 1108–1118 (2014).
31. Montefinese, M., Ambrosini, E., Fairfield, B. & Mammarella, N. The adaptation of the Affective Norms for English Words (ANEW) for Italian. *Behav. Res. Methods* **46**, 887–903 (2014).
32. Greenfield, P. M. The changing psychology of culture from 1800 through 2000. *Psychol. Sci.* **24**, 1722–1731 (2013).

Acknowledgements
We thank colleagues for discussions on this research, especially S. Allen, S. Becker, S. Broadberry, N. Crafts, R. Duch, A. Oswald, L. Pascali, G. Ricco, D. Ronayne, J. Smith and T. Van Rens; and T. Engelthaler and L. Ying for research assistance. This work was supported by a Royal Society Wolfson Research Merit Award WM160074 (to T.T.H.), the Alan Turing Institute (to T.T.H. and C.I.S.), and The Center for Competitive Advantage in the Global Economy at the University of Warwick (to D.S. and E.P.). This research used cloud computing resources kindly provided through a Microsoft Azure for Research Award. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions
T.T.H., E.P., D.S. and C.I.S. were involved in the study design, project planning, data analysis and writing the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41562-019-0750-z.

Correspondence and requests for materials should be addressed to T.T.H.

Peer review information Primary Handling Editor: Stavroula Kousta.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

nature human behaviour

SUPPLEMENTARY INFORMATION

https://doi.org/10.1038/s41562-019-0750-z

In the format provided by the authors and unedited.

Historical analysis of national subjective wellbeing using millions of digitized books

Thomas T. Hills[1,2*], Eugenio Proto[3,4,6], Daniel Sgroi[3,5] and Chanuki Illushka Seresinhe[2]

[1]Department of Psychology and CAGE, University of Warwick, Coventry, UK. [2]The Alan Turing Institute, British Library, London, UK. [3]IZA Institute of Labor Economics, Bonn, Germany. [4]CESifo, Munich, Germany. [5]Department of Economics and CAGE, University of Warwick, IZA, Coventry, UK. [6]Present address: Adam Smith Business School, University of Glasgow, Glasgow, UK. *e-mail: t.t.hills@warwick.ac.uk

Supplementary Notes

Correlations between Words and Average Life satisfaction

In this section, we conduct a non-parametric analysis that complements the conventional regression analysis in the main text. First, we calculated the relative frequency of all words for which there is a valence measure for every year. The relative frequency is simply the number of times the word appears in each year t and country j in the Google book corpus data, divided by the average frequency of every word in the same language j and year t; then we select the words for which the level of correlation between the relative frequency and life satisfaction is significant at the usual threshold of the 0.05% level and calculate the averages of the valence across the words correlating positively and negatively.

If the valences of the words carry information about life satisfaction then the average valence of all words that correlate positively with life satisfaction should be significantly higher than the average valence of the words that correlate negatively. This is exactly what the bars of Supplementary Figure 3 suggest. Words that correlate positively (negatively) with life satisfaction also correlate positively (negatively) with valence. This indicates that valence is aligned with reported life satisfaction over the period for which both are available.

The Publishing Industry: Market Forces and Lags

In this section we analyse the possible channels of transmission from events like wars or reflected in GDP and life expectancy through to literature and then to the NVI. We also empirically determine plausible time lags.

The Publishing Industry

Unless we have reason to suspect some behavioural forces or market failure, economists would normally assume that firms aim to profit maximize. To put this in context, we can think of publishers as fulfilling two roles. First, they attend to the physical (and costly) production of books, which for the period in question almost entirely concerns the manufacture and distribution of printed texts: crucially they cannot publish every book they receive. This leads to their second role, filtering from the mass of submitted books those they wish to publish in order to maximize sales. In this way they act as an intermediary, taking the supply of (largely) unsolicited written books and selecting from them books they feel will match the demand of the reading public. Recently this role has been partly carried out by "agents" who receive unsolicited manuscripts and select from those they wish to bring to the attention of publishers. Nonetheless, the end result is that only a small minority of authors end up with a publishing contract: some estimates suggest that publishers (and more recently, agents) can receive hundreds or thousands of unsolicited manuscripts a year and might select only a handful. Alberge and colleagues[1] give two specific examples of publishers' acceptances from unsolicited fiction submissions: 3/5,000 at Jonathan Cape, and 1/400 at HarperCollins. On that basis the text of published books represents a tiny proportion of the words written by all (published and unpublished) authors. The insight from economic theory is that in order to maximize profits publishers filter in a non-random way to match their choice of which books to publish with the demand from potential readers. The positive correlation we find in Figure 1 also indicates that publishers match books typified by predominantly high valence words ("happy books") to "happy people" and books typified by predominantly low valence words ("sad books") to "sad people." Later in this section we will list some quotes from publishers and authors concerning their rationale for rejecting books submitted for publication. The aim is to provide some supporting evidence for the importance of the potential demand-side to publishers.

We first need to note that there is a strong "survivor bias" when examining rejection letters: the vast majority of books that are rejected by publishers will not see print and it is highly unlikely that rejection letters for these books will come to light. The rejection letters that survive tend to be for books which become successful. What is helpful for us is that the bias works in favour of our hypothesis: if publishers are rejecting books that later do become a success on market-based grounds, it seems likely that they are rejecting many more books that never come to print on the same grounds. What follows are a few notable examples for quite famous books which hint at the importance that publishers place on the marketable nature of books and whether books are a good match for readers: note that these authors and books were eventually printed at some later date which might mean that a book was not a good match at one point but later became a better match for the market, or of course that different publishers had different ideas about what might be a good match.

The examples included here are derived from a very much longer list that can be found in[2] and directly relate the decision to reject to demand from the reading population:

- John Gallsworthy's book "A Man of Property" from "The Forsyte Saga" was rejected on the grounds that "The author writes to please himself rather than to please the novel reading public and accordingly his novel lacks popular qualities" and that the book "would have no real sale in this country".
- Simon Brett recalled the following rejection: "I'm afraid the current state of the fiction market is too depressing for me to offer you any hope for this": this could mean that literally the market demanded depressing books but more likely it is a statement that the publisher felt that demand in the market offered no hope to Brett whose work was not a good match. Either way it supports our argument.
- Harlan Ellison recalls having a piece rejected by Playboy magazine because, while the story was "a knockout piece of writing" it did not match the philosophy of action of the "young urban male readership".
- Laurence J. Peter's book "The Peter Principle: Why Things Always Go Wrong" was rejected by McGraw-Hill in 1964 with the following words: "I can foresee no commercial possibilities for such a book and consequently can offer no encouragement".
- Stephen King remarks that he sent three chapters of a book to a publisher before he had published anything else and the rejection informed him that "We are not interested in science fiction which deals with negative utopias, they do not sell".
- Thomas Hardy's book "Tess of the D'Urbervilles" was rejected on the grounds that the readership might be concerned by "improper explicitness".
- Sherwood Anderson's book "Winesburg, Ohio" was rejected on the grounds that readers might find it "far too gloomy".
- George Moore was told about his book "Esther Waters" that it would "hardly go down here" because of certain scenes (such as childbirth) that might upset the potential readers.

- Herman Melville was told that “Moby Dick” would be “unsuitable for the Juvenille Market in [England]”.
- Laurence Wylie’s chronicle of French country life “A Village in the Vaucluse” was rejected on the grounds that “It is so far from being a book for the general reader that nothing can be done about it”.
- Barbara Pym was told after submitting her novel “An Unsuitable Attachment”: “Novels like (this), despite their qualities, are getting increasingly difficult to sell.” Barbara Pym was also told of her novel “The Sweet Dove Died” that is was “Not the kind of thing to which people are turning.”

Finally, note that in part A.4 of this Supplementary Material we also compare the NVI derived from the Google Books corpus with alternative indices derived from other corpora including text taken from newspapers and find that they are positively correlated. We would argue that this is not surprising as newspaper publishers are also driven by the desire to sell newspapers and so match the mood of their readers.

Different Lags of the Regressors

From the discussion above, we can argue that events happening in one year could feasibly be featured in literature in the same year (if publishers correctly predict the evolution of public mood) or with a lag of several years if publication is time-consuming or delayed. The choice of appropriate lags for the different variables we are considering then becomes an empirical question.

In what follows, we compare different models determining the channels through which a country’s subjective wellbeing is factored into the different written languages based on a lag of $t - \tau$ years, with $\tau = 0, 1, 3, 5, 10$.

In Supplementary Tables 4-6 we present the estimation corresponding to the above models for life expectancy, GDP, and internal conflicts using lags 1, 3, 5 and 10. In GDP the maximal magnitude is at 5 years lag, in conflict the maximal magnitude is at a 1 year lag. For Life Expectancy, it is a bit more complicated, since it goes down after t-1, but then goes up in t-10. We preferred to use t-1 because in t-10 we lose several datapoints. From this specification the resulting lags that best explain changes in the NVI are a one-year lag for life expectancy, a five-year lag for GDP, and one-year lag for internal conflicts.

Word Stability

In this section we recalculate our main index using a set of words that have stable meanings over time. In order to identify the most stable words over time we use the following process. We use our list of ANEW words for all languages (US English, British English, Italian, German) and compute the positive pointwise mutual information (PMI) vectors using the method employed by Recchia and Louwerse[3] and initially introduced by Bullinaria and Levy.[4] For each ANEW word for every year from 1800, the PMI vector is computed as

$$PMI(x,y) = log_2(\frac{P(x,y)}{P(x)P(y)}) \tag{A-1}$$

If we wanted to calculate the PMI for the word “blossom”, then x would be “blossom” and y would be every other word in the ANEW list. $P(x, y)$ would the number of times "blossom" co-occurs with all the different ANEW words divided by the total number of words in the corpus. When calculating co-occurrences we check for ANEW words which co-occur in any 2 word window either before or after word x:

worda wordb blossom wordc worde

$P(x)$ and $P(y)$ is calculated as the frequency of x and y (respectively) divided by the total number of words in the corpus. We then take the log and set any elements containing negative values to zero. Negative values, i.e. when $P(x, y) < P(x)P(y)$, indicate less than the expected number of co-occurrences, which can arise for many reasons, including a poor coverage of the represented words in the corpus. A potentially useful variation, therefore, is to set all the negative components to zero, and use only the Positive PMI.

We then see how each word changes over time and calculate the decadal changes over time using the PMI vectors we have computed for each word for every year. We take the cosine distance of word x of $year_t$ and $year_{t+10}$, where t is every year from 1800 to 2009. The cosine distance between any two elements (u, v) is defined as $\frac{1-uv}{u_2 v_2}$.

For each word, we then take all the cosine distance values and calculate the maximum difference. As an extra robustness check, we also checked that our results held when computing the average difference of the cosine distance values for each word.

Finally, in order to identify the most stable words, we take three different methods. We order all our words in terms of average difference or maximum difference and take the top 25% or top 50% where the top words are the most stable. Supplementary Table 7 shows the most stable and least stable words identified for each language.

We then recompute our new valence indices by using only the stable set of words identified and the corresponding valence scores from ANEW.

Additionally, we also test our valence indices by computing a time-locked yearly valence score for each word based on which of the top words word x has co-occurred with. Therefore, for word x, we find the top 15 words that word has co-occurred with every year. We then calculate the valence of word x in $year_t$ as the average valence of its top 15 co-occurring words.

So, taking our word "blossom", the valence for "blossom" in 1800 will be calculating using the ANEW valence from the words 'freshness', 'flourish', 'firewood', 'canvas', 'foliage', 'ripe', 'blooming', 'glossy', 'bosom', 'awning', 'badger', 'girdle', 'pristine', 'mantle', 'gallop' whereas the valence for "blossom" in 2009 will be calculated using the ANEW valence of the words 'foliage', 'blooming', 'lavender', 'magnolia', 'leaf', 'vine', 'wreath', 'fade', 'lily', 'flourish', 'spring', 'tree', 'spray', 'rot' and 'lemon.'

The results of these analyses are shown in Supplementary Tables 8-11 with related plots of the NVI using only the most stable 50% or 25% of words in Supplementary Figure 8.

Alternative Corpora and Word Norms

In this section we highlight the similarity between our reported results on the NVI based on text derived from the Google Books corpora using the ANEW word norms and variations based on alternative corpora or word norms.

Firstly in Supplementary Figure 6 we recalculate the NVI using the COHA Corpora. The Corpus of Historical American English (COHA), collected independently of the Google Books corpus, represents a balanced and representative corpus of American English containing more than 400 million words of text from 1810 to 1990, by decade, and composed of newspaper and magazine articles.[5] Also plotted in the same figure is the NVI based on the Google Book corpus. The two display a positive correlation of 0.6144 (with a p-value of 0.0051).

In Supplementary Figure 7 we once again compare our own NVI based on the Google Books corpus but this time to an alternative derived from the “Find My Past" data from the British Library’s “British Newspaper Project" which covers 65 million newspaper and periodical articles from the UK across 200 periodicals from 1710-1953. There is a positive correlation between the two of 0.4554 (with a p-value of under 0.000). Supplementary Table 10 provides a direct comparison of the historical determinants of the two indices for the period 1820-1950.

Supplementary Tables 13 and 14 present a regression analysis of two alternative indices derived from SenticNet data, pleasantness and polarity. SenticNet is a well-known resource for sentiment analysis and offers the values for 30,000 concepts in either single word or multi-word expressions.[6] The regression analysis mirrors the analysis of the NVI in Supplementary Tables 2 and 3.

Finally, Supplementary Figure 9 presents a recalculation of the NVI using the alternative AFINN word norms rather than the ANEW word norms used in the main text. The comparison is made for British English and American English and display a positive correlation of 0.9040 and 0.7850 respectively (with p-values under 0.01).

Overview of the NVI over time

The NVI provides a first attempt to measure changes in national mood over the long-run. It also provides a way to assess how significant historical events affected national mood.

Looking at the UK some interesting patterns emerge. The NVI in the 19th century in the UK is high compared to the 20th century. The index falls with the two World Wars, and the stock market crash of 1929 and the subsequent Great Depression. In the post-World War II period the NVI reached a notable high point in 1957, the year of Harold Macmillan’s speech that most Briton’s had “never had it so good". After that the NVI falls through the 1960s and on into the 1978-79 "Winter of Discontent", with the trend rising back in the late 20th century.

Across all of the countries we consider we can see major historical events being picked up by changes in the NVI. To give a few examples: the Year of Revolutions (1848 for the European countries), the outbreak of World War I (1914 for Germany and the UK), the Wall Street Crash (1929 for the USA), Hitler takes power (1933 for Germany), the outbreak of World War II (1939 for Germany and the UK), the end of Korean War (1953 for the USA), the end of Vietnam War (1975 for the USA) and German reunification and the end of Cold War (in 1990 for all countries).

Stochastic Trends

In column 2 of Table 1 of the main text, we introduced a control for deterministic trends. However, stochastic trends may also bias our results. To address this issue we used the Augmented Dickey-Fuller unit-root test for stationarity of the NVI from 1970 onwards for all countries separately: the approach we use is typical and involves a null hypothesis defined as the presence of a unit root (a stochastic trend) and the alternative hypothesis of stationarity.

The test for a unit root can be rejected in all but Italy (MacKinnon approximate $p - value$ for $Z(t) = 0.6898$), which was integrated of order 1 (so is stationary in differences: see below). For the UK, the unit root can be rejected at 10% confidence levels (MacKinnon approximate $p - value$ for $Z(t) = 0.0696$). For these 3 countries we performed the same test on the life satisfaction variable. For life satisfaction in the UK, the test for a unit root can be strongly rejected (MacKinnon approximate $p - value$ for $Z(t) = 0.0000$). This implies that for the UK a stochastic trend is not a confounding variable in the relationship between the NVI and life satisfaction.

For life satisfaction in Italy the unit root test cannot be rejected (Italy: MacKinnon approximate $p - value$ for $Z(t) = 0.2743$), but can be rejected on the first differences; the two series are then integrated of order 1. Accordingly, there are stochastic trends in both life satisfaction and the NVI for Italy. We therefore tested for cointegration between the NVI and life satisfaction in Italy. The test for cointegration between valence and life satisfaction cannot be rejected: in the residuals of the regression of valence on life satisfaction in Italy the test allows us to reject the existence of a unit root (MacKinnon approximate $p - value$ for $Z(t) = 0.0011$). The existence of cointegration between two variables provides a further test of the existence of a link between these variables, establishing a correlation between long-term shocks in both variables. Hence a permanent shock in life satisfaction is featured in the valence as well.

In the analysis in Table 2 of the main text, we addressed the possibility that trends generated by languages, culture or other omitted factors might have biased our initial results. Here we explicitly address the possibility that omitted variables might have generated stochastic trends and biased the correlations presented above. If our estimated life satisfaction and the other regressors are integrated of order bigger than 0, this could potentially be a source of spurious correlation.

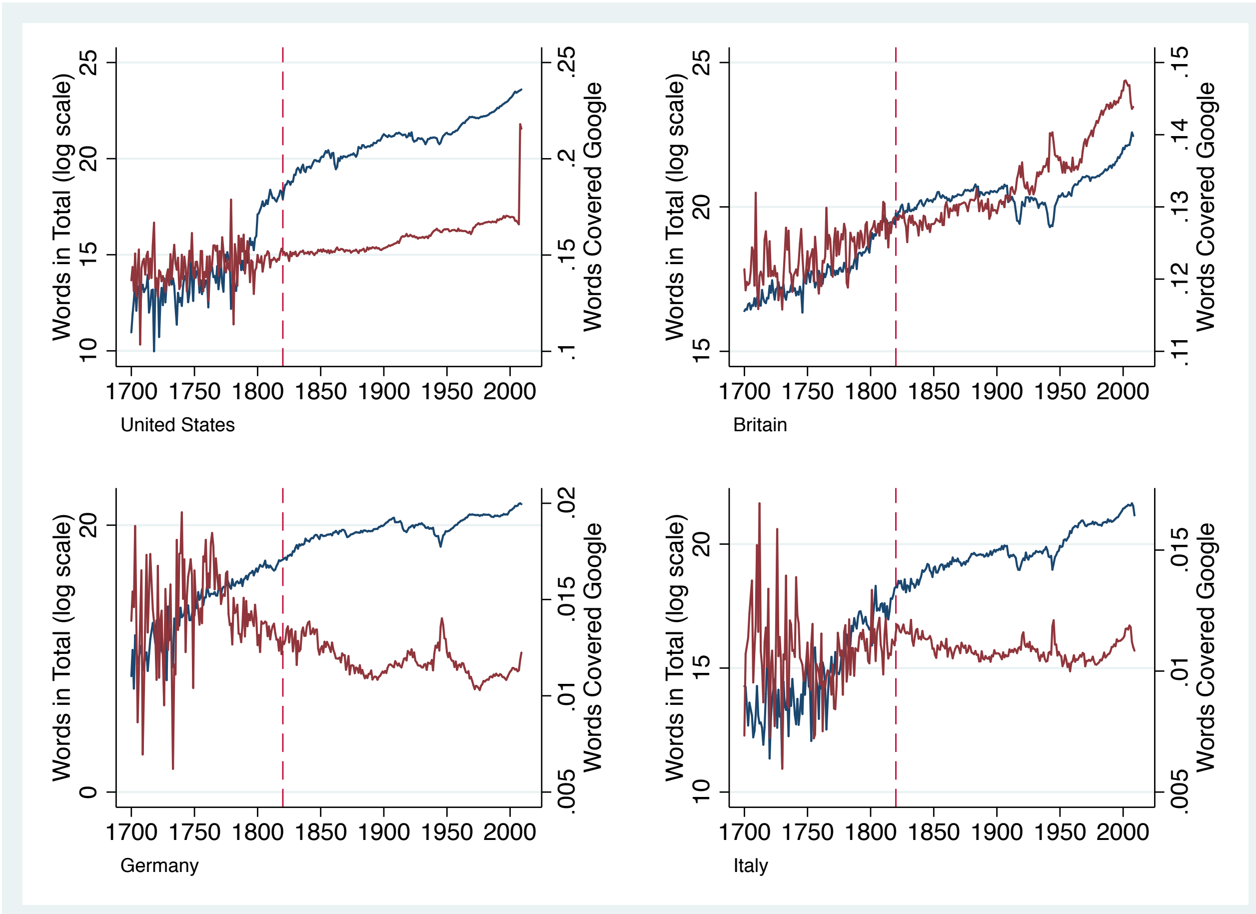
We tested the order of integration of our estimated life satisfaction for all languages and years we are considering with the Augmented Dickey-Fuller unit-root test, and we find that for all the presence of a unit root hypothesis can largely be rejected (while, as it is expected, for both GDP and life expectancy the same hypothesis cannot be rejected).

Further details of all analyses are available upon request.

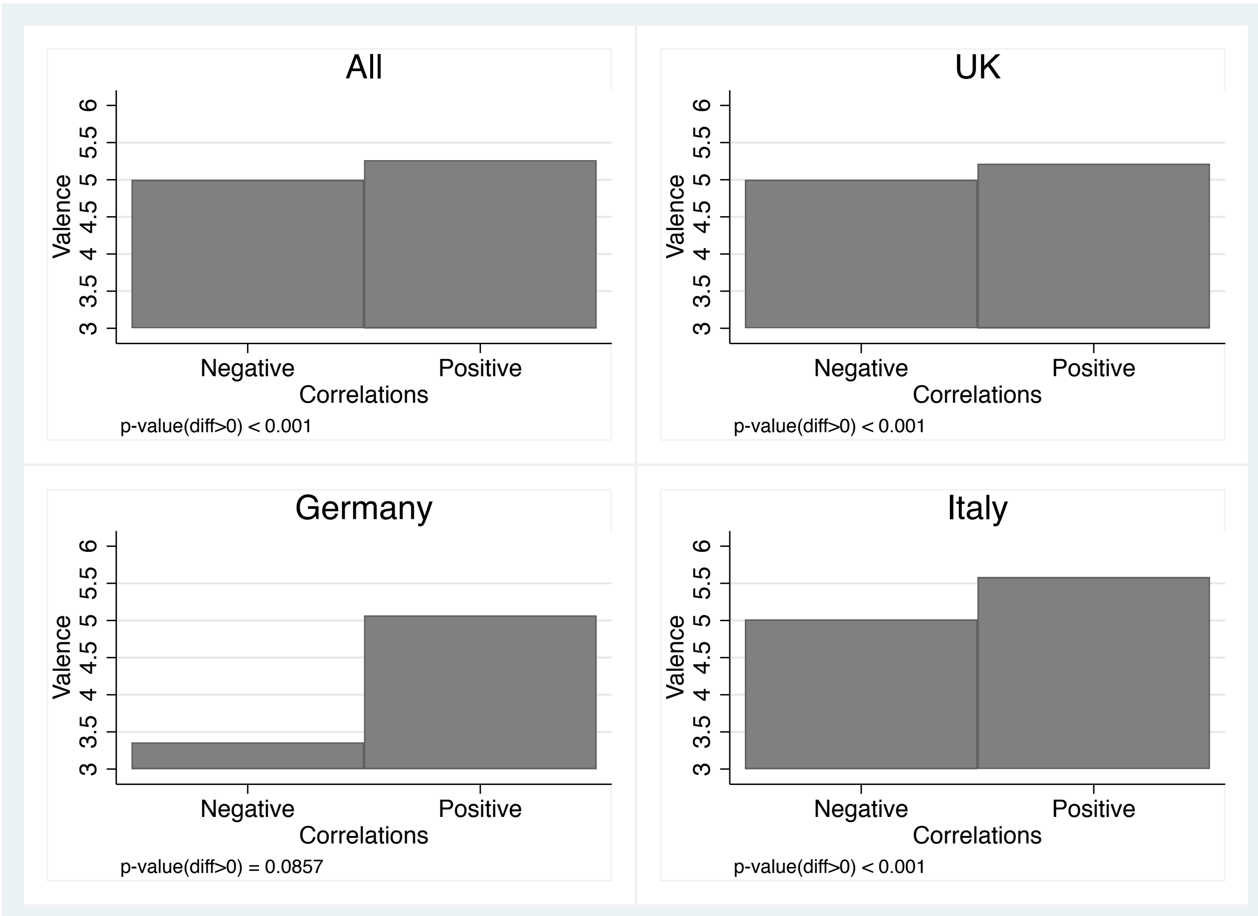
Supplementary Figures

ENGLISH	VALENCE	GERMAN	VALENCE	ITALIAN	VALENCE
aardvark	6.26	Aas	-2.6	abbaglio	3.94
abalone	5.3	Abenddämmerung	-2.35	abbandonato	2
abandon	2.84	Abendessen	2.1	abbondanza	6.82
abandonment	2.63	Abenteuer	0.81	abbraccio	7.7
abbey	5.85	Abfall	1.44	abete	6.17
abdomen	5.43	abkochen	0.4	abitante	5.67
abdominal	4.48	Abschaum	1.9	abitazione	6.46
abduct	2.42	Abscheu	-1.38	abito	7.27
abduction	2.05	Absturz	-1.6	abitudini	4.91
abide	5.52	absurd	-2.7	aborto	2.06
abiding	5.57	Abtreibung	-2.55	abuso	1.74
ability	7	aggressiv	-1.8	accettazione	5.79
abject	4	aktivieren	-0.6	accogliente	8.03
ablaze	5.15	Alarm	1.5	accomodante	6.4
able	6.64	Alimente	-0.79	accordo	6.71
abnormal	3.53	Alkoholiker	2.15	acqua	7.78
abnormality	3.05	Allee	-1.9	adorabile	7.33
abode	5.28	allein	-1.27	adulto	5.78
abolish	3.84	Allergie	-1.56	aereo	6.56
abominable	4.05	Alptraum	-1.56	affamato	4.74
abomination	2.5	anbetungswürdig	-1.22	affascinare	7.97
abort	3.1	angeekelt	0.73	affaticato	3.73
abortion	2.58	angespannt	1.53	affetto	7.48
abracadabra	5.11	Angriff	-2.1	afflizione	1.94
abrasive	4.26	ängstlich	1	affogare	1.79
abreast	4.62	Anreiz	-1.93	aggressione	2.53
abrupt	3.28	Anstellung	-2.21	aggressivo	3.48

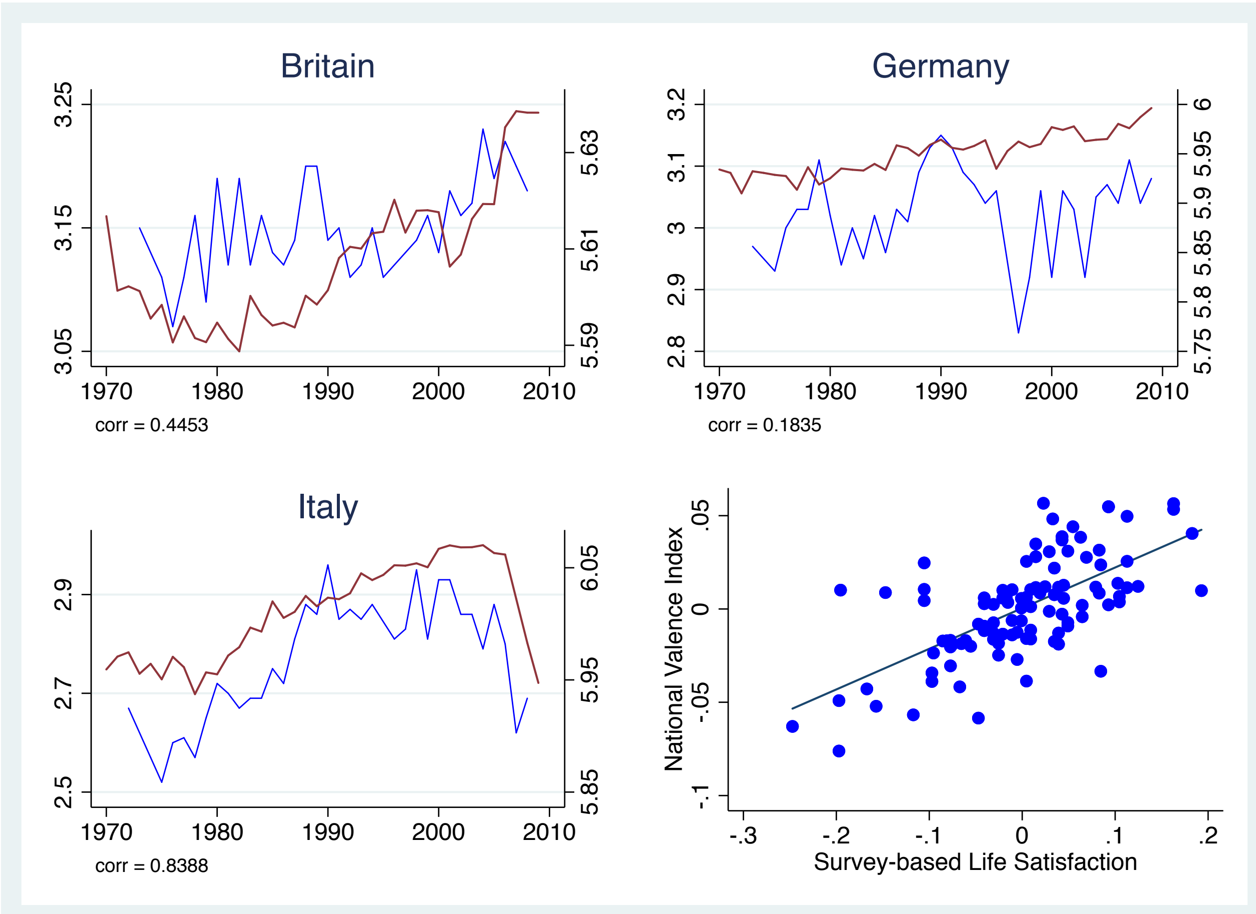
Supplementary Figure 1: **A Sample of Word Valence in Different Languages.** For English and Italian the words are scaled from 1 to 9. For Germany the valence ratings were collected on a -3 to +3 scale. The German mean values were adjusted to reflect a 1 to 9 scale in our analysis.



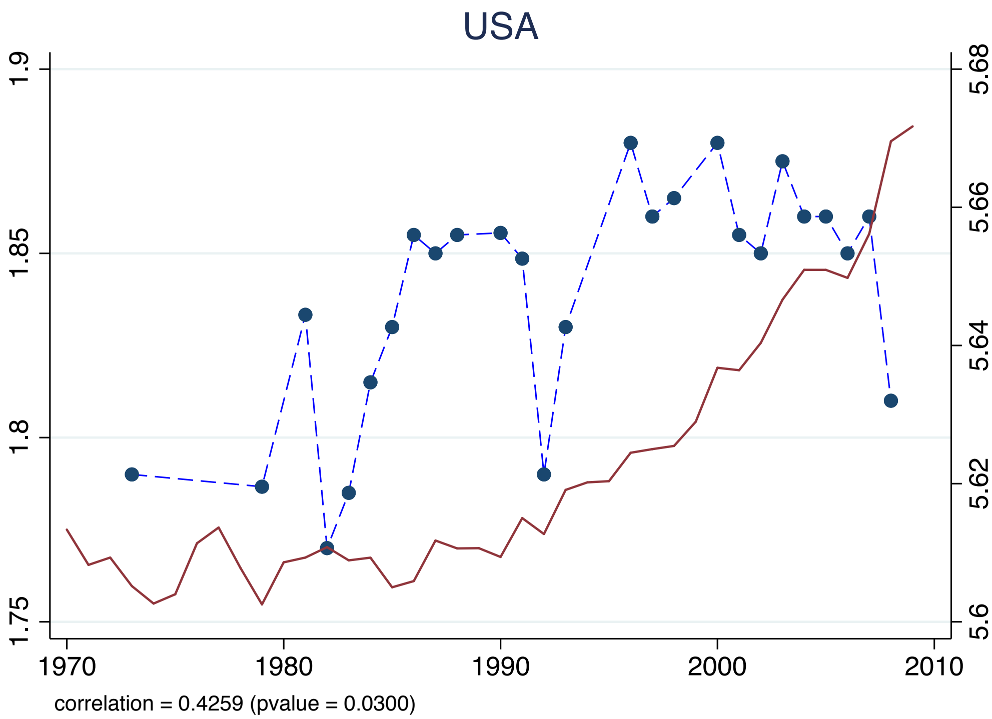
Supplementary Figure 2: **The Number of Words and Share of Words Covered.** The red line represents the proportion of words in the corpus covered in the text analysis by the valence norms and the blue line represents the total number of words—in logarithmic scale—for all countries considered in the analysis.



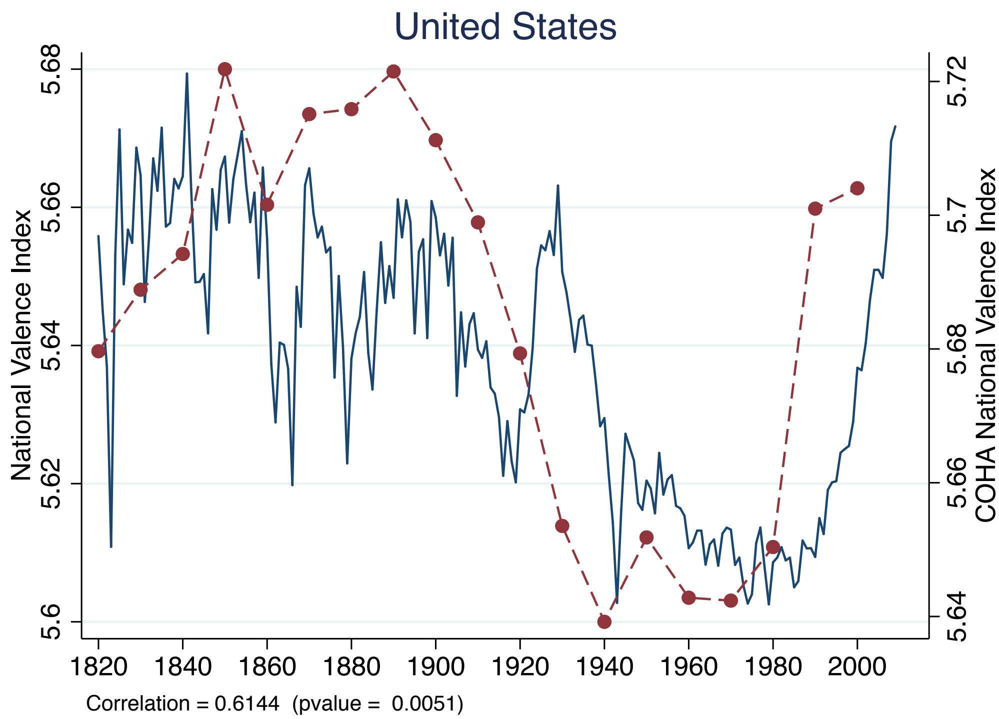
Supplementary Figure 3: **Average Valence and Correlations with Life Satisfaction: All Countries Available.** We selected the words in our dataset for which the level of correlation between valence and life satisfaction (from the Eurobaromter survey-based measure) is significant at the 0.05% level and then calculated the averages of the valence across the words correlating positively and negatively for the UK, Germany and Italy. The bars in the figure represent the average valence of words that correlate positively and negatively. By looking at the bars it is possible to see that the average valence among words that correlate positively with life satisfaction is higher than the average valence among words that correlate negatively with life satisfaction.



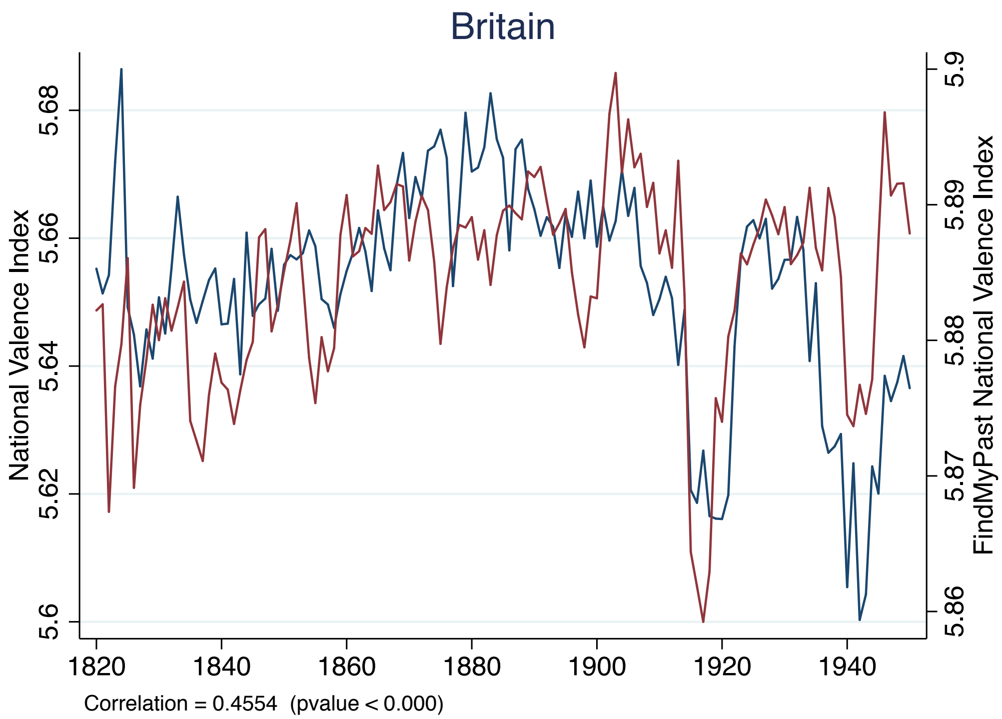
Supplementary Figure 4: **The National Valence Index and Aggregate Life Satisfaction.** In the first 3 panels which present time-series data, the National Valence Index is represented in red (values in the left axis) and life satisfaction is represented in blue (values in the right axis). In the last panel, we plotted the National Valence Index against life satisfaction for the same countries and periods; both variables are expressed in the form of residuals after controlling for country fixed-effects.



Supplementary Figure 5: **The National Valence Index and Aggregate Life Satisfaction in the US.** The National Valence Index is represented in red (values on the left axis) and life satisfaction is represented in blue (values on the right axis). Life Satisfaction data are from the World Database of Happiness[7] and are coded as 1 (= "disatisfied") and 2 (= "satisfied"). They are available only for the years 1973, 1979, 1981-1993, 1996-1998 and 2000-2008.

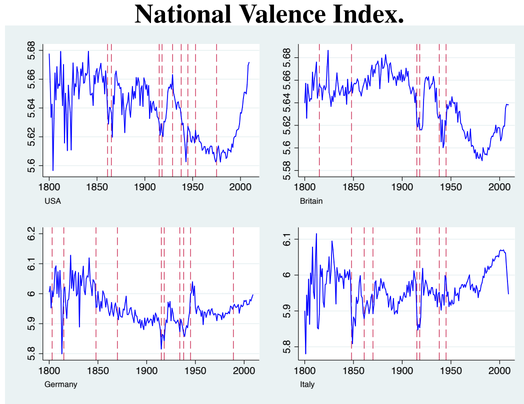


Supplementary Figure 6: **The National Valence Index Derived from Two Different Corpora of US Data.** The red line represents the National Valence Index calculated using the COHA Corpora - based on 400 million words of text from 1810 to 1990, by decade, and composed of newspaper and magazine articles. The blue line represents the US National Valence Index derived from the Google Books corpus.

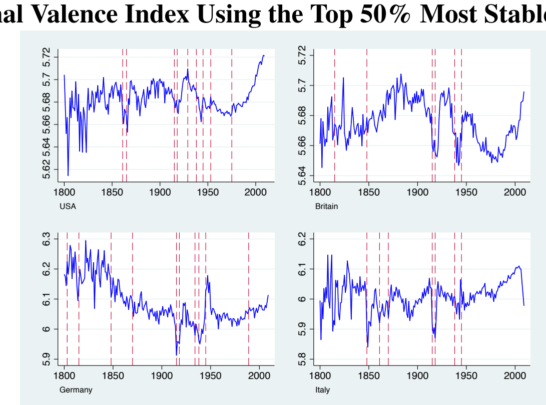


Supplementary Figure 7: **The National Valence Index Derived from Two Different Corpora of British Data.** The red line represents the National Valence Index calculated using FindMyPast data - based on 200 British periodicals from 1820-1953. The blue line represents the British National Valence Index derived from the Google Books corpus.

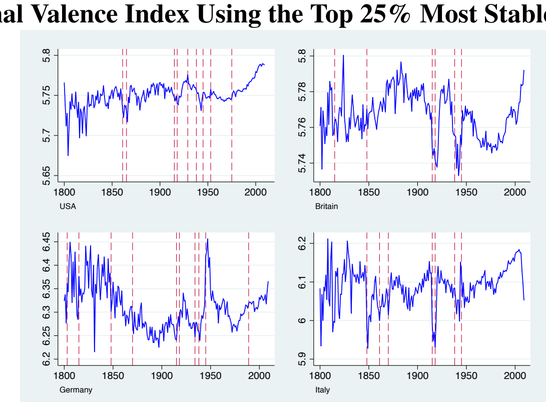
National Valence Index.



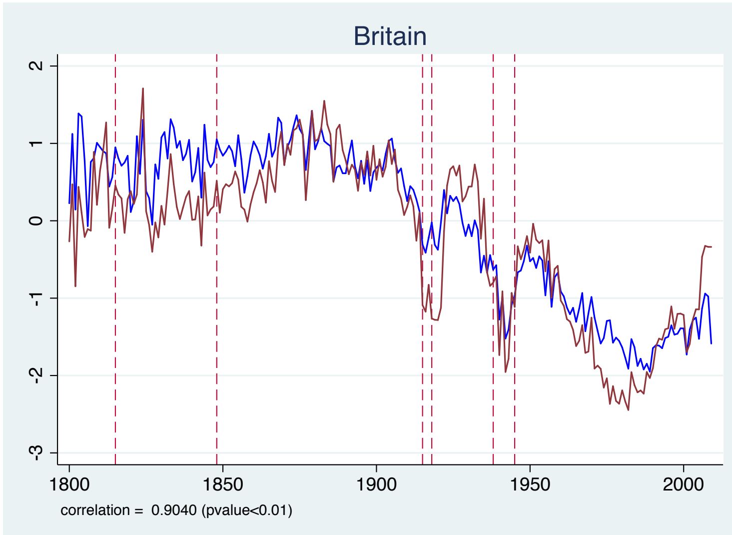
National Valence Index Using the Top 50% Most Stable Words.

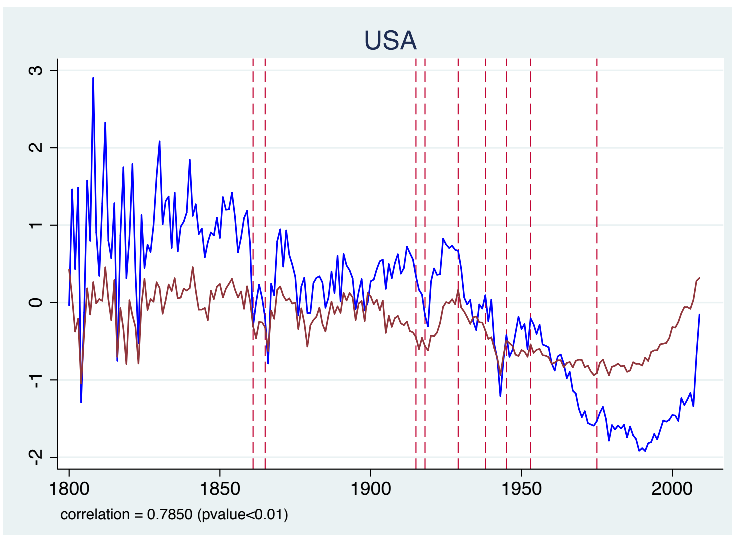


National Valence Index Using the Top 25% Most Stable Words.



Supplementary Figure 8: **A Time-Series Plot Over the Period 1800-2009.**





Supplementary Figure 9: **The NVI Derived from the AFINN Word Norm vs the ANEW Word Norm over the Period 1800-2009.** The blue line represents the National Valence Index derived from the AFINN word norm and the red line the National Valence Index derived from the ANEW word norm. The National Valence Indices are transformed in standard deviations to ease comparability.

Supplementary Tables

Supplementary Table 1: **Main Variables.** These are the mean, standard deviation, minimum value and maximum value of the key variables described in the main text.

Variable	**Mean**	**Std. Dev.**	**Min.**	**Max.**	**N**
National Valence Index	5.798	0.164	5.589	6.128	760
FindMyPast National Valence Index	5.884	0.007	5.859	5.9	131
COHA National Valence Index	5.685	0.029	5.639	5.722	19
Life Satisfaction	2.98	0.181	2.52	3.23	104
Life Satisfaction (US)	1.835	0.033	1.77	1.88	28
per capita GDP (Maddison)	11980.032	11270.36	400	50902	728
per capita GDP (Penn)	25233.999	7193.752	13069.197	43511.594	170
Life Expectancy	61.457	14.088	25.81	82.400	493
Internal Conflict	0.097	0.296	0	1	762
Democracy	5.649	5.894	-9	10	624
Education Inequality	31.526	22.722	6.111	98.935	504
Words Covered Google	0.079	0.068	0.01	0.218	759
Words Covered FindMyPast	0.016	0.001	0.015	0.018	131

Supplementary Table 2: **Differences in the National Valence Index Regressed on Differences in Aggregate Life Satisfaction.** The dependent variable is the difference between two consecutive years in the average life satisfaction per country taken from the Eurobarometer survey-based measure. Simple OLS estimator. The period covered is 1973 to 2009, the period over which both measures exist. The countries considered are Germany, Italy and the UK, the three countries for which both data exist. Per Capita GDP (expressed in terms of purchasing power parity) is from the PWT 8.0 dataset. Both regressions includes year fixed-effects (to help deal with spurious correlations over time). SE = standard error of the mean and p = p-value.

	1 Year FE b/se/p	2 Year FE+GDP b/se/p
NVI(t)-NVI(t-1)	1.2440	1.2638
	(SE= 0.7146)	(SE=0.7334)
	(p= 0.0868)	(p= 0.0901)
Log GDP(t)-Log GDP(t-1)		–0.0774
		(SE= 0.5348)
		(p = 0.8855)
Year FE	Yes	Yes
r2	0.308	0.308
N	95	95

Supplementary Table 3: **Historical Determinants of the National Valence Index – all coefficients are visible.** The countries are Germany, Italy, UK and the United States and the period considered is 1820-2009. The regressions are estimated with an OLS country fixed-effects estimator and either a year fixed-effect (to help deal with spurious correlations over time) or country fixed-effect (to help deal with spurious correlations across countries). Robust standard errors clustered at country levels are given in brackets. SE = standard error of the mean and p = p-value.

	1 Year FE b/se/p	2 Year FE b/se/p	3 Year FE b/se/p	4 CS Trends b/se/p
(log) GDP(t-5)	0.0826		0.0698	0.0550
	(SE=0.0090)		(SE=0.0106)	(SE=0.0130)
	(p=0.0027)		(p=0.0072)	(p=0.0240)
Life Expectancy(t-1)		0.0048	0.0030	0.0016
		(SE=0.0013)	(SE=0.0014)	(SE=0.0013)
		(p=0.0328)	(p=0.1187)	(p=0.2951)
Internal Conflict(t-1)				–0.0184
				(SE=0.0040)
				(p=0.0188)
Words Covered	–1.5813	–2.0859	–1.2282	0.4901
	(SE=1.3370)	(SE=2.2393)	(SE=1.3712)	(SE=0.7027)
	(p=0.3221)	(p=0.4203)	(p=0.4364)	(p=0.5357)
Democracy	0.0030	0.0024	0.0021	–0.0006
	(SE=0.0010)	(SE=0.0008)	(SE=0.0005)	(SE=0.0006)
	(p=0.0575)	(p=0.0620)	(p=0.0245)	(p=0.3339)
Education Inequality	0.0003	0.0008	0.0004	0.0001
	(SE=0.0003)	(SE=0.0002)	(SE=0.0001)	(SE=0.0002)
	(p=0.4050)	(p=0.0181)	(p=0.0341)	(p=0.6943)
Italy Trend				–0.0009
				(SE=0.0007)
				(p=0.2670)
Germany Trend				–0.0007
				(SE=0.0006)
				(p=0.3557)
UK Trend				–0.0016
				(SE=0.0005)
				(p=0.0484)
USA Trend				–0.0018
				(SE=0.0006)
				(p=0.0629)
Year FE	Yes	Yes	Yes	No
r2	0.752	0.705	0.774	0.571
N	412	412	412	412

Supplementary Table 4: **Effect of Life Expectancy on the National Valence Index, using Different Time Lags in the Regressors.** The dependent variable is the NVI at time t. OLS with country fixed-effects estimator. The countries included are Germany, Italy, UK and the United States and the period considered is 1820-2009. This table highlights the significance level of different possible lags of Life Expectancy. Robust standard errors are clustered at country levels are given in brackets. SE = standard error of the mean and p = p-value.

	1 M1 b/se/p	2 M2 b/se/p	3 M3 b/se/p	4 M4 b/se/p	5 M5 b/se/p
Life Expectancy(t)	0.0046 (SE=0.0013) (p=0.0354)				
Life Expectancy(t- 1)		0.0048 (SE=0.0013) (p=0.0328)			
Life Expectancy(t- 3)			0.0044 (SE=0.0008) (p=0.0132)		
Life Expectancy(t- 5)				0.0027 (SE=0.0010) (p=0.0717)	
Life Expectancy(t- 10)					0.0049 (SE=0.0007) (p=0.0050)
Democracy(t)	0.0026 (SE=0.0011) (p=0.0913)	0.0024 (SE=0.0008) (p=0.0620)	0.0029 (SE=0.0009) (p=0.0529)	0.0035 (SE=0.0010) (p=0.0378)	0.0026 (SE=0.0010) (p=0.0724)
Education Inequality(t)	0.0009 (SE=0.0002) (p=0.0158)	0.0008 (SE=0.0002) (p=0.0181)	0.0008 (SE=0.0002) (p=0.0195)	0.0008 (SE=0.0002) (p=0.0328)	0.0007 (SE=0.0003) (p=0.0937)
Words Covered(t)	−2.0159 (SE=2.2155) (p=0.4300)	−2.0859 (SE=2.2393) (p=0.4203)	−1.9190 (SE=2.2140) (p=0.4499)	−2.2976 (SE=2.4087) (p=0.4105)	−1.8185 (SE=2.1879) (p=0.4669)
Year FE	Yes	Yes	Yes	No	No
r2	0.696	0.705	0.699	0.672	0.698
N	412	412	408	404	394

Supplementary Table 5: **Effect of the GDP on the National Valence Index, using Different Time Lags in the Regressors.** The dependent variable is the NVI at time t. OLS with country fixed-effects estimator. The countries included are Germany, Italy, UK and the United States and the period considered is 1820-2009. This table highlights the significance level of different possible lags of GDP. Robust standard errors clustered at country levels are given in brackets. SE = standard error of the mean and p = p-value.

	1	2	3	4	5
	M1	M2	M3	M4	M5
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
(log) GDP(t)	0.0614				
	(SE=0.0072)				
	(p=0.0034)				
(log) GDP(t-1)		0.0611			
		(SE=0.0079)			
		(p=0.0046)			
(log) GDP(t-3)			0.0659		
			(SE=0.0081)		
			(p=0.0039)		
(log) GDP(t-5)				0.0735	
				(SE=0.0111)	
				(p=0.0071)	
(log) GDP(t-10)					0.0728
					(SE=0.0079)
					(p=0.0027)
Democracy(t)	0.0025	0.0026	0.0028	0.0029	0.0027
	(SE=0.0010)	(SE=0.0010)	(SE=0.0010)	(SE=0.0010)	(SE=0.0010)
	(p=0.0786)	(p=0.0763)	(p=0.0647)	(p=0.0578)	(p=0.0645)
Education Inequality(t)	0.0004	0.0004	0.0004	0.0003	0.0002
	(SE=0.0002)	(SE=0.0002)	(SE=0.0002)	(SE=0.0002)	(SE=0.0003)
	(p=0.1287)	(p=0.1366)	(p=0.1766)	(p=0.2670)	(p=0.5048)
Words Covered(t)	–2.5082	–2.4601	–2.2927	–2.1053	–2.1659
	(SE=1.4543)	(SE=1.4147)	(SE=1.2709)	(SE=1.0832)	(SE=1.0778)
	(p=0.1830)	(p=0.1804)	(p=0.1690)	(p=0.1472)	(p=0.1381)
Year FE	Yes	Yes	Yes	Yes	Yes
r2	0.707	0.707	0.718	0.735	0.728
N	459	459	459	459	459

Supplementary Table 6: **The Effect of Internal Conflicts on the National Valence Index, using Different Time Lags in the Regressors.** The dependent variable is the NVI at time t. OLS with country fixed-effects estimator. The countries are Germany, Italy, UK and the United States. This table highlights the significance level of different possible lags of Internal Conflict. Robust standard errors clustered at country levels are given in brackets. SE = standard error of the mean and p = p-value.

	1	2	3	4	5
	M1	M2	M3	M4	M5
	b/se/p	b/se/p	b/se/p	b/se/p	b/se/p
Internal Conflict(t)	–0.0372				
	(SE=0.0161)				
	(p=0.1036)				
Internal Conflict(t-1)		–0.0393			
		(SE=0.0133)			
		(p=0.0594)			
Internal Conflict(t-3)			–0.0316		
			(SE=0.0090)		
			(p=0.0392)		
Internal Conflict(t-5)				–0.0278	
				(SE=0.0064)	
				(p=0.0224)	
Internal Conflict(t-10)					–0.0224
					(SE=0.0072)
					(p=0.0523)
Words Covered(t)	0.0380	–0.0161	0.0231	–0.0876	–0.4527
	(SE=1.5854)	(SE=1.5378)	(SE=1.5244)	(SE=1.4500)	(SE=1.3528)
	(p=0.9824)	(p=0.9923)	(p=0.9889)	(p=0.9556)	(p=0.7599)
r2	0.008	0.010	0.006	0.005	0.006
N	1227	1223	1215	1207	1187

Supplementary Table 7: **The Most Stable and Least Stable Words for each Language, for Words that Existed in 1800.**

Language	Most stable words	Least stable words
UK English	hugger, can, would, will, may	daybreak, daresay, daisy, banter, irrigate
USA English	can, will, would, shall, hundred	stairs, staircase, stainless, sportsman, holly
German	frühling, räuber, liebe, gesundheit, gott	schlüssel, schnee, vogel, sauer, heu
Italian	regina, santo, colore, lago, ferro	saggio, salice, salutare, ratto, gelosia

Supplementary Table 8: **Historical Determinants of the National Valence Index (valence computed using the 50% most stable words identified using the maximum difference in cosine distances), from 1820 to 2009.** OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1 Year FE b/se/p	2 Year FE b/se/p	3 Year FE b/se/p	4 CS Trends b/se/p
(log) GDP(t-5)	0.0669		0.0507	0.0488
	(SE=0.0138)		(SE=0.0152)	(SE=0.0087)
	(p=0.0167)		(p=0.0446)	(p=0.0110)
Life Expectancy(t-1)		0.0048	0.0032	0.0024
		(SE=0.0007)	(SE=0.0010)	(SE=0.0016)
		(p=0.0066)	(p=0.0524)	(p=0.2311)
Internal Conflict(t-1)				–0.0134
				(SE=0.0011)
				(p=0.0012)
Words Covered	0.2436	0.3088	0.2814	0.9849
	(SE=0.6590)	(SE=0.6382)	(SE=0.6851)	(SE=0.4898)
	(p=0.7362)	(p=0.6616)	(p=0.7088)	(p=0.1379)
Democracy	0.0024	0.0017	0.0013	–0.0008
	(SE=0.0004)	(SE=0.0008)	(SE=0.0002)	(SE=0.0006)
	(p=0.0126)	(p=0.1086)	(p=0.0134)	(p=0.2781)
Education Inequality	0.0001	0.0005	0.0002	0.0001
	(SE=0.0002)	(SE=0.0001)	(SE=0.0001)	(SE=0.0002)
	(p=0.7621)	(p=0.0237)	(p=0.1036)	(p=0.5709)
Italy Trend				–0.0011
				(SE=0.0007)
				(p=0.1920)
Germany Trend				–0.0009
				(SE=0.0006)
				(p=0.2314)
UK Trend				–0.0015
				(SE=0.0005)
				(p=0.0716)
USA Trend				–0.0016
				(SE=0.0006)
				(p=0.0767)
Year FE	Yes	Yes	Yes	No
r2	0.691	0.673	0.725	0.464
N	412	412	412	412

Supplementary Table 9: **Historical Determinants of the National Valence Index (valence computed using the 25% most stable words identified using the maximum difference in cosine distances), from 1820 to 2009.** OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1 Year FE b/se/p	2 Year FE b/se/p	3 Year FE b/se/p	4 CS Trends b/se/p
(log) GDP(t-5)	0.0514		0.0375	0.0492
	(SE=0.0084)		(SE=0.0119)	(SE=0.0078)
	(p=0.0087)		(p=0.0507)	(p=0.0079)
Life Expectancy(t-1)		0.0041	0.0030	0.0026
		(SE=0.0010)	(SE=0.0011)	(SE=0.0019)
		(p=0.0249)	(p=0.0761)	(p=0.2605)
Internal Conflict(t-1)				–0.0102
				(SE=0.0021)
				(p=0.0175)
Words Covered	0.9801	1.0423	0.6331	1.2139
	(SE=0.7372)	(SE=0.9230)	(SE=0.5019)	(SE=0.6098)
	(p=0.2757)	(p=0.3409)	(p=0.2963)	(p=0.1406)
Democracy	0.0015	0.0008	0.0005	–0.0009
	(SE=0.0005)	(SE=0.0008)	(SE=0.0004)	(SE=0.0006)
	(p=0.0522)	(p=0.3587)	(p=0.2771)	(p=0.2665)
Education Inequality	0.0002	0.0005	0.0003	0.0003
	(SE=0.0001)	(SE=0.0001)	(SE=0.0001)	(SE=0.0002)
	(p=0.1462)	(p=0.0188)	(p=0.1029)	(p=0.2623)
Italy Trend				–0.0012
				(SE=0.0007)
				(p=0.1815)
Germany Trend				–0.0011
				(SE=0.0007)
				(p=0.2180)
UK Trend				–0.0013
				(SE=0.0005)
				(p=0.0915)
USA Trend				–0.0015
				(SE=0.0006)
				(p=0.0975)
Year FE	Yes	Yes	Yes	No
r2	0.671	0.673	0.703	0.408
N	412	412	412	412

Supplementary Table 10: **Historical Determinants of the National Valence Index (valence computed using the 25% most stable words identified using the average difference in cosine distances), from 1820 to 2009.** OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1	2	3	4
	Year FE	Year FE	Year FE	CS Trends
	b/se/p	b/se/p	b/se/p	b/se/p
(log) GDP(t-5)	0.0921		0.0708	0.0543
	(SE=0.0201)		(SE=0.0191)	(SE=0.0135)
	(p=0.0195)		(p=0.0340)	(p=0.0278)
Life Expectancy(t-1)		0.0064	0.0043	0.0023
		(SE=0.0008)	(SE=0.0009)	(SE=0.0019)
		(p=0.0034)	(p=0.0177)	(p=0.3124)
Internal Conflict(t-1)				–0.0145
				(SE=0.0037)
				(p=0.0287)
Words Covered	0.5820	0.6227	0.4637	0.8523
	(SE=0.5977)	(SE=0.6138)	(SE=0.6590)	(SE=0.6286)
	(p=0.4020)	(p=0.3850)	(p=0.5324)	(p=0.2682)
Democracy	0.0042	0.0034	0.0027	–0.0003
	(SE=0.0005)	(SE=0.0009)	(SE=0.0002)	(SE=0.0006)
	(p=0.0033)	(p=0.0290)	(p=0.0008)	(p=0.6889)
Education Inequality	0.0003	0.0009	0.0004	0.0002
	(SE=0.0004)	(SE=0.0003)	(SE=0.0002)	(SE=0.0002)
	(p=0.5949)	(p=0.0527)	(p=0.0816)	(p=0.2270)
Italy Trend				–0.0009
				(SE=0.0009)
				(p=0.4005)
Germany Trend				–0.0011
				(SE=0.0008)
				(p=0.2755)
UK Trend				–0.0016
				(SE=0.0007)
				(p=0.1083)
USA Trend				–0.0018
				(SE=0.0008)
				(p=0.1071)
Year FE	Yes	Yes	Yes	No
r2	0.739	0.711	0.780	0.605
N	412	412	412	412

Supplementary Table 11: **Historical Determinants of the National Valence Index (time-locked valences computed using the valence of the 50% most stable words, identified using the maximum difference in cosine distances, based on their co-occurence with the observed word), from 1820 to 2009.** OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1	2	3	4
	Year FE	Year FE	Year FE	CS Trends
	b/se/p	b/se/p	b/se/p	b/se/p
(log) GDP(t-5)	0.0470		0.0394	0.0362
	(SE=0.0101)		(SE=0.0101)	(SE=0.0146)
	(p=0.0187)		(p=0.0298)	(p=0.0890)
Life Expectancy(t-1)		0.0027	0.0015	0.0029
		(SE=0.0007)	(SE=0.0010)	(SE=0.0010)
		(p=0.0315)	(p=0.2282)	(p=0.0609)
Internal Conflict(t-1)				–0.0069
				(SE=0.0042)
				(p=0.1957)
Words Covered	1.1891	1.2274	1.2068	0.2085
	(SE=0.6269)	(SE=0.6296)	(SE=0.6328)	(SE=0.3967)
	(p=0.1541)	(p=0.1464)	(p=0.1526)	(p=0.6356)
Democracy	0.0018	0.0016	0.0012	0.0007
	(SE=0.0008)	(SE=0.0007)	(SE=0.0010)	(SE=0.0006)
	(p=0.1165)	(p=0.1160)	(p=0.2952)	(p=0.3597)
Education Inequality	–0.0006	–0.0002	–0.0005	–0.0004
	(SE=0.0001)	(SE=0.0001)	(SE=0.0002)	(SE=0.0003)
	(p=0.0277)	(p=0.1866)	(p=0.0464)	(p=0.2035)
Italy Trend				–0.0022
				(SE=0.0007)
				(p=0.0571)
Germany Trend				–0.0020
				(SE=0.0006)
				(p=0.0424)
UK Trend				–0.0020
				(SE=0.0006)
				(p=0.0453)
USA Trend				–0.0022
				(SE=0.0006)
				(p=0.0314)
Year FE	Yes	Yes	Yes	No
r2	0.547	0.526	0.554	0.299
N	412	412	412	412

Supplementary Table 12: **Comparing Historical Determinants of the National Valence Indices from 1820 to 2009 in Britain, using Find My Past Data and Google**. The NVI are transformed in standard deviations to ease comparability. OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1820-1950 FindMyPast b/se	1820-1950 Google b/se
GDP (log) t	0.9149***	
	(SE=0.1512)	
	(p=0.0000)	
GDP (log) t-5		0.6519**
		(SE=0.2654)
		(p=0.0154)
WW1	–2.1139***	–1.0180***
	(SE=0.2163)	(SE=0.2439)
	(p=0.0000)	(p=0.0001)
WW2	–1.4433***	–1.0039***
	(SE=0.2171)	(SE=0.2570)
	(p=0.0000)	(p=0.0002)
Words Covered(t)	146.1456	–139.5449***
	(SE= 101.3410)	(SE = 34.2593)
	(p=0.0001)	(p=0.1518)
r2	0.529	0.486
N	130	130

Supplementary Table 13: **Historical Determinants using SenticNet Pleasantness from 1820 to 2009.** OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1	2	3	4
	Year FE	Year FE	Year FE	CS Trends
	b/se/p	b/se/p	b/se/p	b/se/p
(log) GDP(t-5)	0.0224		0.0166	0.0002
	(SE=0.0062)		(SE=0.0056)	(SE=0.0020)
	(p=0.0368)		(p=0.0601)	(p=0.9117)
Life Expectancy(t-1)		0.0017	0.0012	0.0002
		(SE=0.0001)	(SE=0.0003)	(SE=0.0003)
		(p=0.0014)	(p=0.0249)	(p=0.5507)
Internal Conflict(t-1)				0.0020
				(SE=0.0036)
				(p=0.6203)
Words Covered	0.0756	0.0128	0.0574	0.1126
	(SE=0.1002)	(SE=0.1643)	(SE=0.1109)	(SE=0.0177)
	(p=0.5055)	(p=0.9428)	(p=0.6403)	(p=0.0079)
Democracy	0.0011	0.0009	0.0007	–0.0001
	(SE=0.0002)	(SE=0.0002)	(SE=0.0001)	(SE=0.0000)
	(p=0.0197)	(p=0.0184)	(p=0.0107)	(p=0.0017)
Education Inequality	0.0000	0.0002	0.0001	–0.0001
	(SE=0.0002)	(SE=0.0001)	(SE=0.0001)	(SE=0.0000)
	(p=0.8477)	(p=0.2144)	(p=0.5990)	(p=0.1370)
Italy Trend				0.0001
				(SE=0.0001)
				(p=0.3802)
Germany Trend				0.0001
				(SE=0.0001)
				(p=0.2646)
UK Trend				–0.0003
				(SE=0.0001)
				(p=0.0268)
USA Trend				–0.0003
				(SE=0.0000)
				(p=0.0041)
Year FE	Yes	Yes	Yes	No
r2	0.668	0.653	0.724	0.872
N	412	412	412	412

Supplementary Table 14: **Historical Determinants of the SenticNet Polarity from 1820 to 2009.** OLS with country fixed-effects estimator. SE = standard error of the mean and p = p-value.

	1	2	3	4
	Year FE	Year FE	Year FE	CS Trends
	b/se/p	b/se/p	b/se/p	b/se/p
(log) GDP(t-5)	0.0122		0.0089	0.0034
	(SE=0.0048)		(SE=0.0048)	(SE=0.0013)
	(p=0.0859)		(p=0.1603)	(p=0.0806)
Life Expectancy(t-1)		0.0009	0.0007	0.0002
		(SE=0.0002)	(SE=0.0003)	(SE=0.0002)
		(p=0.0144)	(p=0.0934)	(p=0.5492)
Internal Conflict(t-1)				0.0022
				(SE=0.0032)
				(p=0.5490)
Words Covered	0.1181	0.0855	0.1089	0.0927
	(SE=0.0995)	(SE=0.1090)	(SE=0.0992)	(SE=0.0105)
	(p=0.3206)	(p=0.4902)	(p=0.3526)	(p=0.0030)
Democracy	0.0007	0.0005	0.0005	–0.0000
	(SE=0.0002)	(SE=0.0001)	(SE=0.0001)	(SE=0.0000)
	(p=0.0440)	(p=0.0103)	(p=0.0198)	(p=0.4510)
Education Inequality	–0.0000	0.0000	–0.0000	–0.0000
	(SE=0.0001)	(SE=0.0001)	(SE=0.0001)	(SE=0.0000)
	(p=0.8242)	(p=0.5105)	(p=0.9332)	(p=0.1273)
Italy Trend				–0.0000
				(SE=0.0001)
				(p=0.9316)
Germany Trend				0.0001
				(SE=0.0000)
				(p=0.1943)
UK Trend				–0.0002
				(SE=0.0001)
				(p=0.0889)
USA Trend				–0.0003
				(SE=0.0000)
				(p=0.0076)
Year FE	Yes	Yes	Yes	No
r2	0.537	0.533	0.577	0.762
N	412	412	412	412

Supplementary Table 15: **Table 1 from main text with complete statistical information.** SE = standard error of the mean and p = p-value.

	1	2
	Year FE	CS trends
	b/se/p	b/se/p
National Valence Index	2.8551	1.6596
	(SE=0.2867)	(SE=0.2246)
	(p=0.0099)	(p=0.0178)
Log GDP	0.2882	0.7613
	(SE=0.0560)	(SE=0.2551)
	(p=0.0358)	(p=0.0963)
Italy Trend		–0.0125
		(SE=0.0049)
		(p=0.1236)
Germany Trend		–0.0152
		(SE=0.0045)
		(p=0.0789)
UK Trend		–0.0204
		(SE=0.0069)
		(p=0.0969)
r2	0.730	0.588
N	104	104

References

[1] Alberge, J. Publishers bypass literary agents to discover bestseller talent. Tech. Rep., *The Observer, February 22* (2015).

[2] Bernard, A. *Rotten Rejections* (Robson Books, 1990).

[3] Recchia, G. & Louwerse, M. Reproducing affective norms with lexical co-occurrence statistics: Predicting valence, arousal, and dominance. *Quarterly Journal of Experimental Psychology* **107**, 1–41 (2014).

[4] Bullinaria, J. P. & Levy, J. A. Extracting semantic representations from word co-occurrence statistics: A computational study. *Behavior Research Methods* **39**, 510–526 (2007).

[5] Davies, M. The 385+ million word corpus of contemporary American English (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* **14**, 159–190 (2009).

[6] Cambria, E., Poria, S., Bajpai, R. & Schuller, B. Senticnet 4: A semantic resource for sentiment analysis based on conceptual primitives. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, 2666–2677 (2016).

[7] Veenhoven, R. World database of happiness, collection happiness in nations, overview of happiness surveys using measure type: 121A / 2-Step verbal life satisfaction. *Viewed on 2019-01-16 at worlddatabaseofhappiness.eur.nl* (2019).