NBER WORKING PAPER SERIES

STATE OF THE ART:
ECONOMIC DEVELOPMENT THROUGH THE LENS OF PAINTINGS

Clément Gorin
Stephan Heblich
Yanos Zylberberg

Working Paper 33976
http://www.nber.org/papers/w33976

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2025, Revised June 2026

We would like to thank Dionissi Aliprantis, Elliott Ash, Laure Athias, Nicolas Baumard, Sascha Becker, Karol Borowiecki, Davide Cantoni, Don Davis, Shari Eli, Oliver Falck, Lucie Gadenne, Oded Galor, Edward Glaeser, Claudia Goldin, Ruru Hoong, Erik Hornung, Michael Hutter, Robert Johannes, Hubertus Kohle, Dominik Loibner, Bob Margo, Stelios Michalopoulos, Petra Moser, Nathan Nunn, Ömer Özak, Dominic Rohner, Stefanie Schneider, Munir Squires, Marco Tabellini, Mathias Thoenig, Daniel Trefler, Matt Turner, Hans-Joachim Voth, Fabian Waldinger; participants at Boston College, Brown University, the Cleveland Fed, Cologne, ENS-PSL, ENS Lyon, Harvard/ Boston University, Leicester, LMU, NYU, Paris Dauphine, Paris Saclay, Queen Mary University, Royal Holloway, UBC, the University of Toronto, Warwick University; and conference participants at the ACEI 2023, the 2024 CUER-UBC workshop, the 2024 German Economists Abroad conference, the 2024 IEB Workshop on Urban Economics, the Microdata in Economic History Workshop, the NBER SI, and the 2025 RUSE for helpful comments. Leyi Wang and Tony Gui provided excellent research assistance. Gabriel Benevides built the companion website. We acknowledge research support from the ORA Grant ES/V013602/1 (MAPHIS) and the Social Sciences and Humanities Research Council of Canada (SSHRC, Insight Grant 435-2025-0725). The usual disclaimer applies. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

State of the Art: Economic Development Through the Lens of Paintings
Clément Gorin, Stephan Heblich, and Yanos Zylberberg
NBER Working Paper No. 33976
June 2025, Revised June 2026
JEL No. C45, O10, R11, Z1

ABSTRACT

This paper uses 627,369 paintings since 1400 to study how societies experienced major socioeconomic transformations. We develop computer vision algorithms to extract two signals from each artwork—emotional expression and visual indicators of material living standards—and validate them against modern measures of wellbeing and economic output. Our empirical analysis documents how populations experienced socioeconomic changes by exploiting variation in emotional and material signals within artists' oeuvres and conditional on painting sub-genres. While both respond to shocks to living standards, such as climate or trade, emotional expression also varies with broader economic and institutional conditions, often without corresponding changes in material outcomes. These findings suggest that paintings provide a long-run measure of experienced welfare, complementing conventional indicators of economic performance.

Clément Gorin
University of Paris 1
Sorbonne School of Economics
Sorbonne Economics Centre (CES)
clement.gorin@univ-paris1.fr

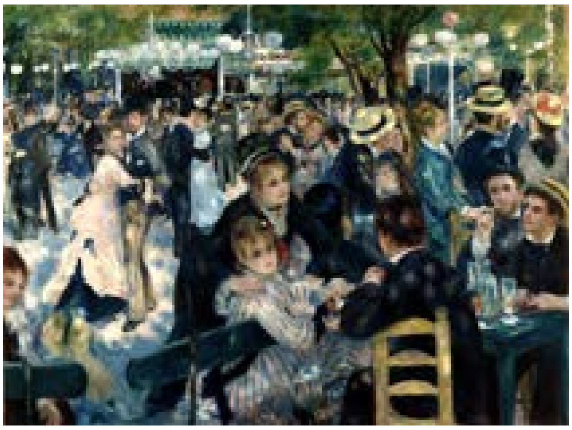
Stephan Heblich
University of Toronto
Department of Economics
and NBER
stephan.heblich@utoronto.ca

Yanos Zylberberg
University of Bristol
yanos.zylberberg@bristol.ac.uk

A data appendix is available at http://www.nber.org/data-appendix/w33976

Throughout history, painters have produced visual representations that reflect the social, political, and economic conditions of their world, offering interpretations of ritual and daily life, power and protest, and poverty and prosperity—often in contexts where other forms of data are sparse or absent. Yet despite their richness, paintings have rarely been used at scale to systematically study historical societies. This paper introduces new algorithmic methods to analyze past socioeconomic conditions through a corpus of 627,369 paintings by 32,670 artists.

Figure 1. Paintings as testaments to their times and contexts.



(a) Dance at the Moulin de la Galette

(b) The Angelus

Notes: Panel **(a)** shows *Dance at the Moulin de la Galette* (Pierre-Auguste Renoir, 1876), which depicts a lively scene at a popular Parisian dance hall. Panel **(b)** shows *The Angelus* (Millet, 1857–1859), portraying the hardship of rural life. The additional illustrations discussed in this introduction are provided in Appendix A.

Paintings may offer not only visual records of social realities but also contemporaneous interpretations of them. Pierre-Auguste Renoir's portrayals of the Parisian bourgeoisie (panel a of Figure 1) convey the excitement associated with the emerging culture of urban leisure, growing prosperity, and the rhythms of modern life during the *Belle Époque*. By contrast, Millet's nostalgic representations of peasant life (panel b) highlight the harsh conditions of rural poverty beyond urban centers. Political turbulence and protest appear in Eugène Delacroix's *Liberty Leading the People* (1830), which celebrates the fight for freedom across social classes, and contrast sharply with Pablo Picasso's *Guernica* (1937), which conveys the collective trauma and horror of the Spanish Civil War. The message may also be symbolic or indirect. For instance, Hieronymus Bosch's *Garden of Earthly Delights* renders fantastical scenes during an age of exploration and discovery, capturing both excitement and anxiety in a period marked by profound uncertainty. Finally, paintings can offer multiple, distinct interpretations of the same subject, shaped by the artist's perspective. William-Adolphe Bouguereau and Gustave Courbet both depict rural labor in 19th-century France, yet Bouguereau presents a romanticized pastoral ideal, while Courbet delivers a realist critique of social hardship. Together, these examples illustrate how paintings can capture dimensions of historical experience typically absent from standard data sources—revealing how societies *perceived* their economic, social, and political contexts.

This paper proposes to unlock new economic insights by treating cultural artifacts—specifically paintings—as structured, quantitative data. We leverage large open-access repositories of artwork and encode paintings along two dimensions: (i) a sentiment classification pipeline that identifies emotions conveyed through artwork, such as sadness, fear, contentment, or amusement, and (ii) the extraction of visual indicators of material conditions. We validate these sentiment and material signals against modern indicators of well-being (Helliwell et al., 2025) and output per capita where available (Bolt and Van Zanden, 2025). The fine-grained nature of our data allows us to document how societies experienced socioeconomic transitions through expressive changes *within* individual artists' oeuvres over time, while separately capturing variation in emotional expression across contemporaneous artworks—a measure of disagreement. We show that emotions, disagreement, and material wealth depicted in paintings provide a long-run quantitative proxy for experienced welfare, capturing how societies perceived and lived through socioeconomic transformations since 1400.

To extract contextual information embedded in historical paintings, we exploit open-access repositories from Google Arts and Culture (138,396 paintings), WikiArt (263,125 paintings), and Wikidata (225,848 paintings). Our emotion classification relies on a state-of-the-art Convolutional Neural Network (CNN) architecture designed to identify the emotions conveyed by a painting. We construct the training set using large-scale annotations comprising 79,860 paintings and over 1.5 million emotion labels, drawn from prior studies (Achlioptas et al., 2021; Mohamed et al., 2022). Annotators assign a dominant emotion to each painting from a set of predefined categories, with 15–25 distinct annotators randomly assigned per artwork. Our target probabilistic emotion scores correspond to the empirical distribution of these labels across annotators. The neural network architecture combines a visual encoder, originally trained on photographs, with a regression head that maps extracted features into emotion probabilities. Jointly estimating the regression head with the encoder allows the model to adapt to the distinctive characteristics of paintings and capture emotional expression through visual elements such as composition, movement, contrast, and texture.

The validity of our sentiment classifier rests on two hypotheses. **(H1)** Internal validity: the model accurately predicts the emotions assigned by human annotators. We present a series of exercises showing that predicted emotion probabilities closely track annotated distributions across periods, genres, and emotion categories. **(H2)** External validity: these predictions capture the emotions intended by the artist. To assess this, we benchmark annotators' free-text justifications for assigned labels against artists' own descriptions of their work.[1]

The extraction of visual features related to material living standards follows a two-stage procedure. First, a large language model generates (a) a wealth index, (b) a relevance or informativeness score, and (c) structured descriptions of visual elements conveying information about living standards. Second, we use (a) and (b) as targets to train an in-house algorithm that jointly predicts wealth and relevance directly from images. As in the sentiment pipeline—where annotators' justifications inform our understanding of their choices—the structured artwork descriptions (c) provide insight into the mechanisms underlying the inference of living standards. We map these structured descriptions into a taxonomy that aggregates object-, character-, and context-level attributes into interpretable measures of poverty and wealth. The inference of living standards by the large language model—and, by extension, by our calibrated in-house algorithm—primarily relies on physical appearance, posture, and markers of social status (people), the quality of (often domestic) objects, and scene-level interpretations. These procedures yield two complementary signals for each painting: a probabilistic vector of emotions capturing experienced conditions, and an index of depicted living standards capturing material conditions.

Using paintings as representations of their context raises two critical concerns. First, selection into digitized artworks may reflect transient or persistent tastes, commercialization, patronage, museum acquisition practices, or historical contingencies, potentially biasing inferences about the socioeconomic environment (see Cohen, 1999; Moretti, 2000, for a discussion of the "great unread" in literature). Second, the production of paintings is shaped by multiple forces—such as demand conditions—that may complicate the interpretation of their visual content. We address these concerns through a series of validation exercises on both inputs and outputs. To assess selection bias and the role of commercialization in shaping emotional expression, we compile (i) 710,121 modern paintings from *Artsy* with associated auction prices and (ii) more than one million art transactions between 1650 and 1950. We find that the signals extracted from paintings are uncorrelated with contemporaneous prices and other measures of commercialization.

Our strongest validation strategy is to document (i) the mapping between extracted signals of sentiment and the well-being of populations observed in the same country

[1]We further rely on these justifications to assess annotation quality by analyzing the semantic alignment between the selected emotion and its textual explanation. Annotators may differ in their ability to interpret artwork depending on cultural background and familiarity with art and its historical context. Despite disagreement among the 15–25 annotators per artwork, a strong wisdom-of-the-crowd effect emerges: the *distribution* of annotated emotions closely aligns with artists' stated intentions.

and year, and (ii) the *separate* mapping between representations of living standards and economic output. First, we leverage recent paintings from *Artsy* to relate the emotions conveyed by artists in a given country and year to the reported feelings of individuals from the same gender and cohort in Gallup (2,753,803 interviews; see Helliwell et al., 2025). Conditioning on year, country, and artist fixed effects, as well as output per capita and artwork attributes (including price), we find that a one-standard-deviation increase in reported life satisfaction is associated with a 0.22–0.28 standard deviation increase in average positivity conveyed by artwork. This correspondence varies by emotion type (e.g., depicted fear is most strongly correlated with reported worry), indicating that the emotional signals extracted from paintings track meaningful variation in experienced well-being. The relationship also holds within contexts: countries in which women report lower well-being exhibit more negative emotional expression among female artists. Second, we relate predicted material living standards to historical estimates of output per capita available for all countries between 1900 and 2000 (Bolt and Van Zanden, 2025). Conditional on conveyed sentiment, year and country fixed effects, and weighting artworks by their "informativeness," a one-standard-deviation increase in output per capita raises the wealth index by approximately 0.25 standard deviations. We further validate this relationship across multiple painting genres (religious, mythological, portraits, and genre scenes), with the notable exception of landscapes and abstract art, and identify the descriptive features most predictive of material living standards. Third, both mappings are primarily contemporaneous, in the sense that they hold most strongly when the timing of painting production aligns closely with the measurement of well-being or living standards.

A key challenge in using signals extracted from paintings to study how societies perceived major socioeconomic transformations is that both an artwork's subject and its perceived tone may be shaped by artist-specific characteristics (such as mindset, patrons, or style) and by the nature of the exercise (e.g., a portrait). To address these concerns, each painting is linked to rich metadata, including artist biography, artistic movements, sub-genre classifications, and physical attributes such as material and dimensions. We then estimate specifications equivalent to standard two-way fixed-effects models, with extracted signals as outcome variables and measures of socioeconomic transformation as explanatory variables, conditioning on artist fixed effects, artist-age dummies, 43 exercise dummies (e.g., battle painting, marina), and location and year fixed effects. Identification therefore relies on within-artist variation across contexts—defined at the country-year level—after adjusting for systematic differences arising from the application of our approach to particular exercises.

As a first step, we examine how conveyed sentiment responds to three macroeconomic indicators: (i) economic development (Bolt and Van Zanden, 2025), (ii) uncertainty, and (iii) inequality (Alvaredo et al., 2020).[2] Sustained growth of 4% in output per capita over 25 years reduces the probability of conveying sadness by 0.50 standard deviations, with increased depictions of awe, amusement, or excitement. Uncertainty also matters: it lowers contentment and shifts expression toward more unstable emotions (e.g., fear, disgust, or amusement). Inequality, in turn, increases disagreement in emotional signals within a given context. For example, a 10-point increase in top-1% wealth concentration—of the magnitude recently observed in advanced economies—increases disagreement by 0.21 standard deviations. These magnitudes indicate that emotional expression responds strongly to economically meaningful features of the environment.

As a second step, we examine how societies responded to a set of well-defined, distinct socioeconomic changes. We find that emotional expression tracks material representations for shocks that directly affect living standards, including climatic variation in Europe between 1500 and 2000 (highlighting the fragility of pre-industrial societies otherwise illustrated in Nunn and Qian, 2011; Waldinger, 2022), trade integration during the first wave of globalization, and technological change associated with the diffusion of the steamship (Pascali, 2017). By contrast, other forces affect emotional expression without corresponding changes in material representations. These include shifts in the social order, such as the Enlightenment—with its emphasis on freedom, rationality, experimentation, and collective progress (Mokyr, 2005)—as well as features of the geopolitical environment, including conflict, autocracy, political consolidation, and power transitions, which we measure using micro-data on academic production and political conditions. Together, these results provide new evidence on how societies experienced these transformations, showing that emotional expression captures dimensions of well-being not fully reflected in material conditions alone.

The main contributions of this paper are twofold. First, we develop a new approach to measuring material and experienced welfare from historical paintings. Second, we use these measures to provide new insights into the dynamics of welfare over time and across space (see our companion website).

By extracting emotional signals from paintings, we construct a new long-run measure of experienced welfare—capturing how societies perceived and lived through eco-

[2]Visual inspection of time series for twelve countries shows that the emotional content of paintings fluctuates in ways consistent with well-known socioeconomic transformations. Fear rises during turbulent periods, while contentment increases in stable times; sadness intensifies during hardship, often at the expense of amusement or excitement. Dispersion in visual signals across artworks produced in the same time and place carries additional explanatory power, particularly for economic inequality. Appendix A illustrates these patterns for France; Appendix D.1 covers all twelve countries; and a companion website allows for a comprehensive exploration of these time series, alongside a description of representation dynamics within individual artists' oeuvres.

nomic, institutional, and political change. These indicators, constructed from a novel data source—historical paintings—complement a large body of work relying on more direct and objective measures of past economic conditions (e.g., Bairoch, 1988; Clark, 2005, 2007; Schularick and Taylor, 2012; Bolt and Van Zanden, 2025).[3]

The most closely related studies are Safra et al. (2020), who predict perceived trustworthiness in 2,000 English portraits from 1505 to 2016; Boerner et al. (2025), who derive time series from colors in paintings; and Hills et al. (2019), who infer welfare in four countries using textual analysis of books. Like these contributions, our approach provides a proof of concept that indirect depictions of society, preserved in cultural artifacts, can serve as windows into their times. However, we differ along three key dimensions: (i) the scale of our exercise, covering 627,369 paintings by 32,670 artists across 34 present-day countries with at least 1,000 artworks each; (ii) the richness of our measurement, which leverages the full composition of a painting to infer complex emotional messages; and (iii) the scope of our analysis, which traces perceived responses to a wide range of socioeconomic conditions.

Our study contributes to a growing body of research that infers economic conditions from indirect data sources, imagery in particular. The novel use of images as data is best illustrated by recent work studying societal evolution through photographs (Voth and Yanagizawa-Drott, 2024). This emerging literature also includes studies based on satellite imagery (Overman et al., 2006; Chen and Nordhaus, 2011; Henderson et al., 2012; Jean et al., 2016; Donaldson and Storeygard, 2016) and street-level images (Gebru et al., 2017; Naik et al., 2017). Our approach relies on deep convolutional neural networks (see Dell, 2025, for a review), and our contribution is to extend these techniques to the measurement of welfare in the distant past.

Our approach rests on the premise that conveyed emotions can offer novel insights into how societies experienced socioeconomic change.[4] For instance, expressions of fear capture periods of uncertainty in ways that parallel more conventional economic indicators (Bloom, 2014; Jurado et al., 2015). Despite this potential, emotions have remained unconventional metrics in economics. While early thinkers such as Bentham recognized

[3]Prominent contributions to our understanding of historical economic performance across regions come from Angus Maddison (Maddison, 2007; Bolt and Van Zanden, 2025), or from Bairoch's work on city populations (Bairoch, 1988). These indicators have been instrumental in documenting long-run patterns of growth, structural change, and urbanization (see, e.g., Acemoglu et al., 2005a; North, 2010; Fujita et al., 2001; Galor, 2005, 2011). Some contributions construct more detailed economic indicators for smaller regions over the long run (e.g., Clark, 2005, 2007; Bouscasse et al., 2024, for England) or for broader sets of countries over more recent periods (e.g., Schularick and Taylor, 2012; Jacks et al., 2011; Jordà et al., 2017).

[4]Our emotion categories align with the concept of *basic emotions* (Ekman, 1999), which are thought to arise from evolutionarily adaptive tasks and social learning. Because these emotions are universal across cultures and relatively stable across time and space (Tooby and Cosmides, 1990), they serve as reliable indicators of how past societies responded to economic, social, and political transformations.

that preferences—and thus choices—are rooted in emotions, these constructs have historically been more difficult to quantify than observable variables such as prices, production costs, or output. Recent work, however, has reasserted the importance of emotions in shaping decision-making (Elster, 1998; Damasio, 2006), influencing risk preferences (Meier, 2022), and informing policy attitudes (Algan et al., 2025). See DellaVigna (2009) and Wälde and Moors (2017) for comprehensive reviews of the role of emotions in economics, and Hulten and Nakamura (2022) and Benjamin et al. (2024) for broader discussions of how economic activity and welfare can be measured.

Finally, our research relates to the literature on art and creativity in economics. One strand of this literature studies the lives of artists to understand the determinants of productivity and creativity (Galenson and Weinberg, 2000, 2001; Hellmanzik, 2010; Borowiecki, 2017, 2022). A second strand uses artistic production to study the economic importance of protecting novel ideas (Giorcelli and Moser, 2020; Whitaker and Kräussl, 2020). A third strand examines art markets (Ginsburgh and Jeanfils, 1995; Ashenfelter and Graddy, 2006; Ursprung and Wiermann, 2011; Graddy, 2013; Renneboog and Spaenjers, 2013; Spaenjers et al., 2015; Etro, 2018; Etro and Stepanova, 2021) and art consumption (Scitovsky, 1972; Throsby, 1994). We contribute to this literature by highlighting an additional channel through which art informs our understanding of the economy: the preserved messages conveyed by artists themselves.[5]

Our analysis proceeds as follows. Section 1 draws on insights from art history and psychology to motivate our working hypotheses. Sections 2 and 3 introduce the data and develop our measurement framework, including validation of the extracted emotional and material signals. Section 4 outlines our empirical strategy and examines responses to socioeconomic transformations. Section 5 concludes.

1 Artwork as a window into socio-economic contexts

Our formal approach uses learning algorithms to explore historical contexts through the lens of painters. This section briefly summarizes insights from art history, anthropology, evolutionary biology, and psychology that inform our working hypotheses.

The joint evolution of art and societies Research into the evolutionary roots of artistic behavior suggests that aesthetic practices were essential components of human cultural and cognitive evolution (Dissanayake, 1992; Dutton, 2009; Brown, 2021). Art has long been acknowledged as a critical factor in the evolution of societies and their

[5] This complements the work of Logothetis et al. (2025), which shows how historical critical junctures can influence artistic expression, shaping how music reflects and transmits cultural identity.

cultures, acting both as a reflection of and a catalyst for social change. Recent contributions from economics echo this perspective, e.g., Michalopoulos and Xue (2021), or Michalopoulos (2025) for a more comprehensive discussion.

An important insight from the literature is that art serves not only as an aesthetic experience but also as a record of the emotional and social life of individuals and communities (a reflection of society; see, e.g., Albrecht, 1954; Alexander, 2020). Artworks aim to convey emotions, ideas, and knowledge. This reflective function underlies our own approach, which rests on the premise that the evolution of artwork can reveal the evolution of societies. However, artistic expression also shapes social interactions and helps societies articulate cultural patterns (a catalyst channel; see Dissanayake, 1992; Gell, 1998). Indeed, interdisciplinary research at the intersection of anthropology, evolutionary biology, neuroscience, and art history paints a complex picture in which art is not merely a byproduct of human civilization but an *active* participant in the cognitive and social development of humans (Henrich et al., 2016).

Quantifying the meaning of artwork Art historians seek to recover the meaning of artwork, employing methods that have influenced our work at different stages. We broadly categorize these methodologies into three groups: (i) those focusing on the visual and symbolic aspects of the artwork, such as formalism and iconography; (ii) those focusing on the artist, such as psychoanalytic art history, which explores artists' subconscious influences on art, and biographical approaches that consider the artist's personal life and intentions; and (iii) those focusing on the broader social context, which studies the social, political, and economic conditions surrounding the creation of artworks. We draw inspiration from all these areas in our attempt to decode the information that artists embedded in their work and to relate it to societal developments.

Our algorithmic approach emulates formalism and iconography to predict emotions expressed in the artwork. The algorithm flexibly accounts for the *Elements of Art*, e.g., color, form, line, shape, space, and texture, and for the *Principles of Art*, e.g., scale, proportion, unity, variety, rhythm, mass, shape, space, balance, volume, perspective, and depth. The attempt to link art elements and composition principles to emotions is perhaps best described by Leo Tolstoy's (1898) description of the artist's creative process: "To evoke in oneself a feeling one has once experienced, and having evoked it in oneself, then by means of movements, lines, colors, sounds, or forms expressed in words, so to transmit that feeling that others may experience the same feeling." Our approach essentially inverts this process: we observe the emotions associated with artworks by a large set of individuals and train an algorithm to learn how visual elements and principles of art translate into the feelings perceived by viewers.

Our procedure to quantify the meaning of art is closely related to the emerging field of digital art history, which explores how quantitative and computational techniques can enrich traditional art-historical inquiry. A 2013 special issue of the journal *Visual Resources* provided scholars with a forum to explore future directions for the field. Among the areas of interest identified were algorithmic computational methods, including machine learning and natural language processing. As noted by Jaskot (2019), "digital art history lets us address the tradition of the social history of art in new ways," because digital methods are particularly well suited to engaging with broader art-historical debates that examine artworks in their social context and require large bodies of evidence. This objective aligns with the core of the social history of art, which "is not satisfied with a social context for art, but rather reverses this equation by arguing that an analysis of art, artist, and audience must tell us something structurally about society." Building on this intuition, we propose a method to uncover what paintings can reveal about society and its historical circumstances, as we discuss next.

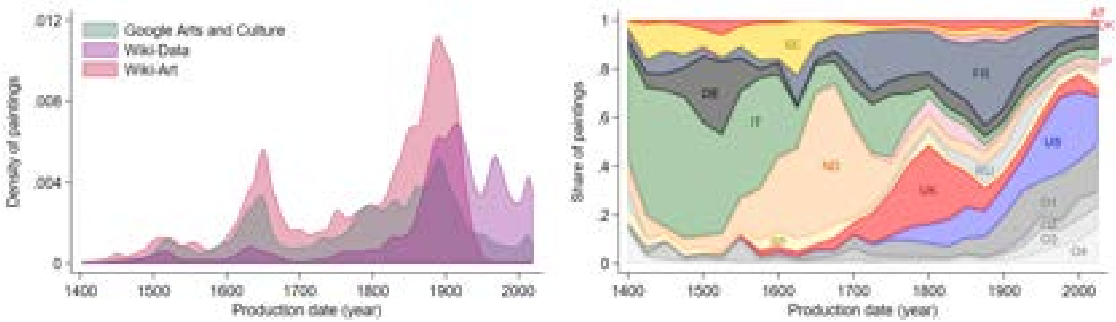
Societal influence, creativity, and personal circumstances Retrieving societal influence from the painter's creative process implies two difficulties. First, predicted emotions combine individual, artist-specific effects with broader societal influences. For instance, there was a shift in Goya's way of painting after 1793 that was likely driven by personal circumstances (a near-fatal illness) as well as societal developments (political turmoil in Spain and the effects of the Napoleonic wars). This example illustrates how both personal and social factors find reflection in artwork. In this paper, our focus lies on the latter—the social component. To disentangle social effects from personal idiosyncrasies, our empirical design controls for artist-specific circumstances and identifies group effects, i.e., the commonality behind artists' joint reflections on historical conditions and social changes. Second, a key working hypothesis is that the message and emotions perceived by a contemporaneous audience are a proper reflection of the message and emotions intended by artists of very different backgrounds.[6] The following sections provide empirical support for these working hypotheses.

[6]A related question concerns the stability of emotional expression across time and cultures. This question connects our work to (i) psychological research on the evolutionary basis of emotions, to (ii) work on the neuroscience of emotions (Damasio, 2006), and to (iii) psychoanalytic art history. We refer the interested reader to work by Ekman on basic emotions, Damasio on emotions as support to the reasoning process, and Robinson on emotions in art (Robinson, 2005). Our analysis focuses on the detection of basic emotions in historical paintings. Ekman (1999) argues that basic emotions are evolutionary adaptations to recurring life challenges. We adopt a commonly used framework that mostly relates to six such emotions: happiness; surprise; fear; disgust; anger; and sadness. According to Frijda (2007), each basic emotion is triggered by specific environmental stimuli and is associated with distinct behavioral responses. For example, fear is evoked by perceived threats, leading to defensive responses like freezing or fleeing. These patterns reflect species-wide adaptive mechanisms that evolved to solve recurrent survival problems, and such emotions are universal and largely independent of cultural variation (Tooby and Cosmides, 1990).

2 Data sources

This section briefly describes our main data sources, with further details provided in Appendix B. Our empirical analysis draws on (i) large, open-access repositories of art collections and auction data covering historical and modern artworks (Appendix B.1); (ii) a sub-sample of paintings with manually annotated emotions (Appendix B.2); and (iii) data on historical economic development and socio-economic change (Appendix B.3).

Figure 2. Temporal and spatial coverage of our image collection(s).



(a) Time of production (b) Selected countries

Notes: Panel **(a)** shows the distribution of production years in *Google Arts and Culture*, *Wiki-Art*, and *Wiki-Data*. *Google Arts and Culture* and *Wiki-Art* provide more consistent coverage of the Renaissance, the later Reformation, and the Enlightenment. The two main modes correspond to the Dutch Golden Age and the French Belle Époque. *Wiki-Data* primarily covers the more recent period from 1800 to 2000. Panel **(b)** shows the relative coverage over time for the following "harmonized" countries and country groups: Austria (AT, red), Belgium (BE, gold), Denmark (DK, light red), France (FR, navy blue), Germany (DE, black), Italy (IT, green), Japan (JP, pink), the Netherlands (ND, orange), Russia (RU, light blue), Spain (SP, yellow), the United Kingdom (UK, red), the United States (US, blue), other major European countries with more than 1,000 artworks (O1, e.g., Portugal, Sweden), other major Asian countries (O2, e.g., China, India), other major producers from other continents (O3, e.g., Argentina, Australia, Brazil, Mexico, South Africa), and smaller producers (O4, all countries with fewer than 1,000 artworks). We provide additional descriptive statistics on the coverage of these data sources in Appendix B.1, including their representation of different movements and types of artwork over time.

A harmonized collection of paintings Our primary dataset consists of approximately 630,000 high-resolution images by over 32,000 artists, retrieved from three large repositories of art collections: Google Arts and Culture; Wiki-Art; and Wiki-Data.[7] We collect information from *Google Arts and Culture* and *Wiki-Art* using a headless browser, which downloads images at the highest available resolution, along with attributes (e.g.,

[7]There is some overlap between these collections, e.g., for the most notable artwork. To identify duplicates, we rely on a fuzzy matching procedure and apply the algorithm described in Section 3.1 which produces—in an intermediary step—a vector of 1,260 embeddings for each painting. More specifically, we focus on paintings from the same artist within a given period, construct the squared difference between embeddings of separate pieces across different collections, and predict a probability to be "duplicates". The last step is informed by the constitution of a small training sample, in which we flag duplicates and non-duplicates for about 1,000 pairs of paintings from the same artists. The convenient representation of images into embeddings—allowing to create multi-dimensional indicators of distance—is used by a few recent contributions attempting to estimate demand for differentiated products (see, e.g., Han and Lee, 2025; Compiani et al., 2025).

production year, style, or movement), and retrieves additional details from the artists' Wikipedia pages. Access to *Wiki-Data* is facilitated through an API. The resulting dataset contain high-quality information on each painting's production date and physical characteristics (e.g., material and size).

Figure 2 presents the temporal coverage of these distinct data sources. *Google Arts and Culture* and *Wiki-Art* offer consistent coverage of the Renaissance, the later Reformation, and the Enlightenment, as well as the golden ages of painting (such as the Dutch Golden Age and the French Belle Époque). *Wiki-Data*, by contrast, is more focused on the period from 1800 to 2000 (panel a). Combined, these sources provide consistent coverage of the major European economies between 1500 and 2000 (panel b). Panel (b) of Figure 2 also illustrates how the geography of cultural production shifts over time in our curated sample: Italy is the dominant producer before 1600 (with the Holy Roman Empire also prominent around 1500); the Netherlands and the Low Countries lead from the late fifteenth to early seventeenth century; France emerges as a major center from 1700 to 1950; Great Britain peaks around 1800; and the United States, along with other world regions, gain prominence from 1900 onward.

Our data sources—*Google Arts and Culture*, *Wiki-Art*, and *Wiki-Data*—provide partial information about the artist or the artwork. To enrich and harmonize this information, we employ several strategies to associate each painting with approximately 100 harmonized artistic movements (e.g., post-impressionism, mannerism, fauvism), 40 harmonized exercises (e.g., allegorical painting, landscape, still life), and standardized artist identifiers, along with data on the artist's nationality, birth and death details, and possible main residences over the course of their career.[8]

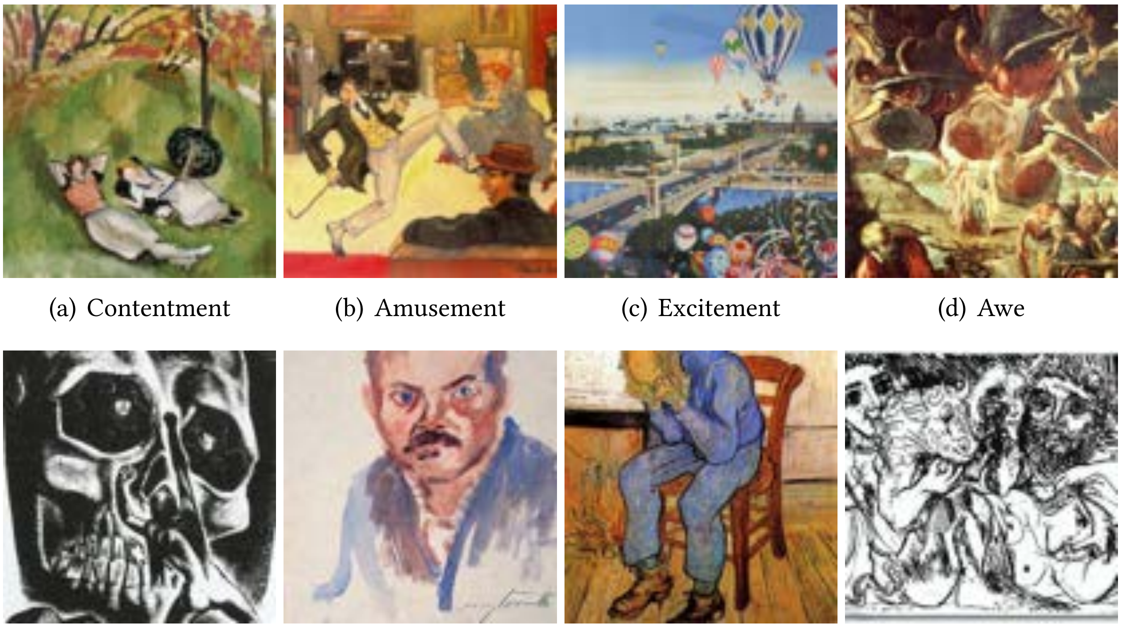
This procedure yields a dataset of 627,369 paintings by more than 32,000 artists, for which we have reliable, comprehensive information on artwork characteristics (title, movement, genre, physical features), context of production (year, location), and creator identity. These geolocated and dated artworks span 110 present-day countries, with 12 (and 34, respectively) accounting for more than 10,000 (resp. 1,000) distinct pieces. The leading contributors include Austria, Belgium, Denmark, France (115,286), Germany (40,792), Italy (60,375), Japan, the Netherlands (69,723), Russia, Spain, the United Kingdom (56,470), and the United States (68,454).

Finally, we complement this collection with 710,121 modern artworks from *Artsy*, a leading online art marketplace for which we observe information on artworks, artists

[8]We rely on three main strategies. First, many artworks in *Google Arts and Culture*, *Wiki-Art*, and *Wiki-Data* include sufficient information to retrieve the artist's *Wikipedia* page, from which we request structured data via the Wikipedia API. Second, we use pre-trained Named Entity Recognition models to systematically clean and harmonize tags across sources (e.g., artist name, genre, art movement), which we complement with tagging information from over 500,000 paintings in the ART500K dataset (Mao et al., 2017). Third, we use a large language model (GPT-4o) to classify the remaining artists and paintings.

(biographies, prizes and awards, auction histories), materials, dimensions, and prices; and with roughly 1,200,000 recorded transactions for a subset of artworks dated between 1650 and 1950 from *Getty*—including auction ID, date, object description, dimensions, artist, buyer/seller, and price—which we use to examine the role of art commercialization and the selection of preserved artworks.

Figure 3. Representative artworks by dominant *annotated* emotion.



(a) Contentment (b) Amusement (c) Excitement (d) Awe

(e) Fear (f) Anger (g) Sadness (h) Disgust

Notes: This figure shows representative artworks with high annotator agreement on a dominant emotion (amusement, anger, awe, contentment, disgust, excitement, fear, sadness, and other—not shown). Sources: ArtEmis (Achlioptas et al., 2021, approximately 450,000 annotations), ArtELingo (Mohamed et al., 2022, approximately 1,000,000 annotations).

A collection of *annotated* images Our sentiment pipeline requires labeled data—a subsample of paintings with associated annotations—in order to map each painting j onto a vector of emotion $\mathbf{p}_j$. We rely on ArtEmis (Achlioptas et al., 2021, which provides approximately 450,000 annotations for a subset of 79,860 Wiki-Art paintings discussed above) and ArtELingo (Mohamed et al., 2022, which provides over 1,000,000 annotations for the same 79,860 paintings). Both sources collect emotion scores along nine dimensions: amusement, anger, awe, contentment, disgust, excitement, fear, sadness, and "other." In practice, individuals were shown paintings and asked to select one dominant emotion from these nine categories. ArtEmis includes an average of 5.6 annotations per painting, while ArtELingo includes an average of 15.3; each annotation corresponds to a different evaluator. We treat these multiple annotations as probabilistic evaluations of one latent observer: each painting j is assigned a probabilistic emotion vector $\mathbf{q}_j$, based on the 15–25 labels provided by randomly drawn annotators. This vector records the share of evaluators who selected each of the nine emotions, and its entries sum to one.

The emotions most frequently selected are contentment (often associated with landscapes), awe (typically linked to beauty or displays of power and opulence), sadness (associated with melancholic portraits, depictions of poverty, or religious imagery), fear (often linked to chaotic scenes, whether through subject matter or visual texture—such as a ship engulfed by waves), excitement, amusement, and disgust.

Figure 4. Understanding reported emotions from annotators' free-text justifications.

(a) Contentment (0.302) (b) Amusement (0.101) (c) Excitement (0.071) (d) Awe (0.168)

(e) Fear (0.095) (f) Anger (0.018) (g) Sadness (0.101) (h) Disgust (0.048)

Notes: This figure shows the words most frequently evoked in annotators' textual justifications for selecting a given dominant emotion. We compute *Term Frequency-Inverse Document Frequency* (TF-IDF) scores to identify words that are distinctive within the corpus of justifications for each emotion label. Intuitively, TF-IDF down-weights common terms that appear across many justifications (e.g., "the") while up-weighting words that are distinctive to specific justifications. For each emotion e, we also report its average probabilistic score p^e in parentheses.

How do annotators interpret the artist's intentions and categorize emotions in paintings? Different choices may reflect variation in personality, training, cultural background, or familiarity with art and historical context. Conveniently, annotators select a dominant emotion but also provide a short description to justify their choice. We illustrate the mapping between paintings, annotators' perceptions, and their chosen emotions in Figures 3 and 4. Figure 3 shows paintings that received the highest scores for each emotion. Annotators respond both to subject matter (e.g., a clown for amusement) and to visual cues such as color, texture, and composition. Figure 4 uses annotators' textual explanations to reveal how emotion labels are interpreted: contentment is linked to calm and nature; sadness to darkness and death; fear to facial expressions and shadows; and disgust to bodily representations.[9]

Integration with additional data The main empirical novelty of our research is to exploit paintings as data points. However, our analysis also requires access to more conventional data sources, most notably: indicators of economic development and inequality, such as output per capita (Bolt and Van Zanden, 2025), life expectancy (Zijdeman and Ribeira da Silva, 2015), and wealth concentration (Alvaredo et al., 2020); measures of sentiment, e.g., Gallup (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025), World Value Surveys (270,760 interviews), and a valence index based on sentiment analysis of published books (Hills et al., 2019); and major observable socio-economic changes. These shocks include the adoption of new technologies (Cross-country Historical Adoption of Technology, see Comin and Hobijn, 2004), the first wave of globalization (Jordà-Schularick-Taylor Macrohistory Database, see Jordà et al., 2017), the intensification of knowledge production (The Rise of Universities in Medieval and Early Modern Europe, see De La Croix et al., 2024), shocks to livelihoods over the period of interest, such as political turbulence (Marshall et al., 2014), wars, climatic variability (Luterbacher et al., 2004), and short-term changes in the prices of basic commodities (Allen-Unger Global Commodity Prices Database, see Allen and Unger, 2019). We describe all these data sources in Appendix B.3.

3 Measurement and validation of emotional and material signals

This section develops our approach to measuring emotional and material signals from paintings and provides empirical validation of these measures. We first describe the convolutional neural networks used to extract emotional and material features, and then assess their internal and external validity using annotated data and modern socioeconomic indicators.

3.1 A sentiment classification pipeline

Our objective is to map each painting onto a probabilistic vector of emotions, where the target reflects the distribution of emotion labels assigned by annotators.

A convolutional network We adopt a transfer learning approach to predict the emotions conveyed by paintings (Bengio et al., 2013). Our model operates in two stages. First,

[9] We provide supplementary material in Appendix B.2, where we discuss differences across movements and types of exercise (e.g., portraits, landscapes, religious or allegorical works), as well as measurement issues related to the number of ratings per image.

a large, pre-trained image model—the encoder—extracts a hierarchy of generic features, such as edges, textures, motifs, shapes, and object parts that capture the essential visual characteristics of the artwork.[10] This encoder consists of the convolutional layers of an EfficientNetV2-S model (Tan and Le, 2021), originally trained on ImageNet (Deng et al., 2009, a multinomial classification task involving 1.4 million photographs of objects from 1,000 different classes). The model produces an array of size $12 \times 12 \times 1280$, where the 12×12 grid partitions the input image into coarse spatial regions, and each of the 1,280 variables encodes the presence of a specific feature within those regions. We then apply global average pooling by computing the mean of each feature across all spatial locations, thereby treating all regions equally, regardless of whether they correspond to central figures or peripheral details.

Second, a regression head is added on top of the encoder, mapping the 1,280 pooled feature values into nine probabilities, one for each emotion. This regression head—preceded by a dropout layer to mitigate overfitting (Tompson et al., 2014)—is a multinomial logit model, yielding a distribution over emotions with scores that sum to one and can be interpreted as probabilities. Letting $j \in \mathcal{J}$ denote a painting, $\mathbf{x}_j$ the vector of extracted features from the encoder model, and y_j the latent emotion of painting j, the regression head assumes,

$$p_j^e = P\left(y_j = e|\mathbf{x}_j\right) = \frac{e^{z_{e,j}}}{\sum_{e \in \mathcal{E}} e^{z_{e,j}}},$$

where the latent scores $z_{e,j} = f_e(\mathbf{x}_j, \mathbf{b})$ are flexible functions of the extracted features, allowing the model to capture non-linearities and interactions between input variables.

Optimization The model minimizes the Kullback-Leibler divergence metric between the target probability distribution, $\{q^e\}_{e \in \mathcal{E}}$, and the predicted probability distribution, $\{p^e\}_{e \in \mathcal{E}}$,

$$\sum_{e \in \mathcal{E}} \left(q^e \ln(q^e) - q^e \ln(p^e)\right).$$

Model parameters are estimated by maximum likelihood, using the AdamW optimizer with dynamically adjusted learning rates (Loshchilov and Hutter, 2017) and a class-balanced loss that up-weights observations with uncommon emotion distributions (Cui

[10] While convolutional networks typically require standardized input, for instance, square images of a uniform size, the raw images vary in both resolution and aspect ratio. The canvas dimensions would, however, mostly fall within a longer-side-to-shorter-side ratio between 1:1 and 1:1.5 (see Figure C1 in the Appendix), and the nature of artwork usually implies that the center of the piece is its most important part. We therefore center-crop each image while preserving its aspect ratio, and then resize it to a $384 \times 384 \times 3$ array, where the last dimension represents the three color channels: red, green, and blue. For example, Figure 3 shows pre-processed paintings, demonstrating that cropping and resizing do not significantly affect visual perception.

et al., 2019).[11]

The model optimization proceeds in two stages. First, we optimize the model while keeping the encoder parameters fixed—effectively treating the encoder's output as given and updating only the regression head that maps these features onto emotion probabilities. Directly optimizing the entire model from the start could result in large gradient updates from the randomly initialized head, potentially disrupting the encoder's ability to extract meaningful features. Second, we fine-tune the full model by (i) unfreezing the encoder parameters—except for the batch normalization layers, which remain frozen to preserve the stable feature distributions acquired during pre-training—and (ii) training the model with a low learning rate to mitigate over-fitting. This step effectively adapts the model from the "real-world photograph" domain to the painting domain, encouraging it to extract features that are predictive of emotions in artwork.[12]

We estimate and validate the model by splitting the annotated sample into three subsets: a training set (70%), a validation set (15%), and a test set (15%). The training set is used to estimate the parameters of the encoder and regression head; the validation set is used to implement early stopping and select the parameters with the best generalization performance;[13] and the test set is used to evaluate accuracy, as discussed next.

Model validation The validity of our overall procedure relies on two key hypotheses. **(H1)** Internal validity: the model predicts the emotions sensed by annotators; **(H2)** external validity: the annotator labels—and the resulting predictions—capture the artists' intentions. This subsection provides empirical evidence on model fit and potential over-fitting, both overall and across genres and periods (H1). We defer to the next subsection the discussion of the less testable assumption that the predictions meaningfully reflect the emotions that the artist(s) intended to convey (H2).

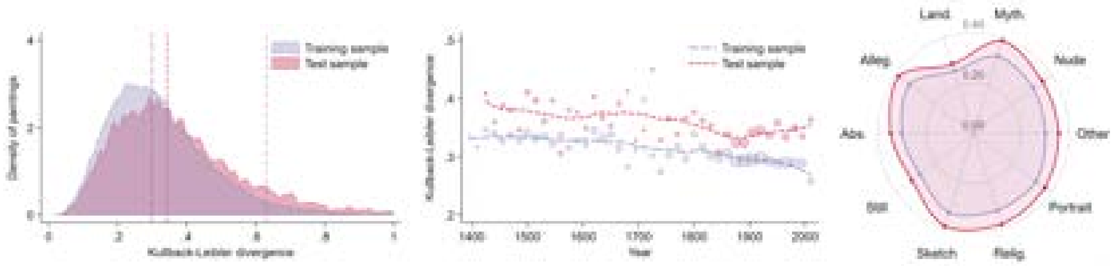
Appendix C.1 provides a detailed account of model performance, and the main takeaways appear in Figure 5: the Kullback-Leibler divergence is similar across the training and test sets, indicating that the model is not subject to major over-fitting (panel a); the divergence increases gradually with the time since production in both sets (panel b); and

[11] In practice, the model is implemented using the GPU-enabled version of the PyTorch library and can be run on a standard high-specification laptop with a properly configured Python environment; access to high-performance computing is not required.

[12] The hierarchical structure of convolutional neural networks allows them to build increasingly complex representations by combining simpler features learned in earlier layers. As illustrated in Appendix C, early layers capture low-level characteristics such as brush strokes, color gradients, and textures; mid-level layers extract more abstract elements such as shapes and lighting; and later layers focus on high-level aspects including composition, subject matter, and expressive details such as facial expressions and posture.

[13] To further limit over-fitting, we apply random transformations—such as flipping, zooming, and rotation—to each training batch. This increases the size of the training sample and provides invariance properties to these transformations. In addition, because we use dropout layers, we can leverage Monte Carlo Dropout (Gal and Ghahramani, 2016) to compute standard errors for our predictions.

Figure 5. Model validation—Kullback-Leibler divergence.



(a) Distribution (b) Period (c) Genre

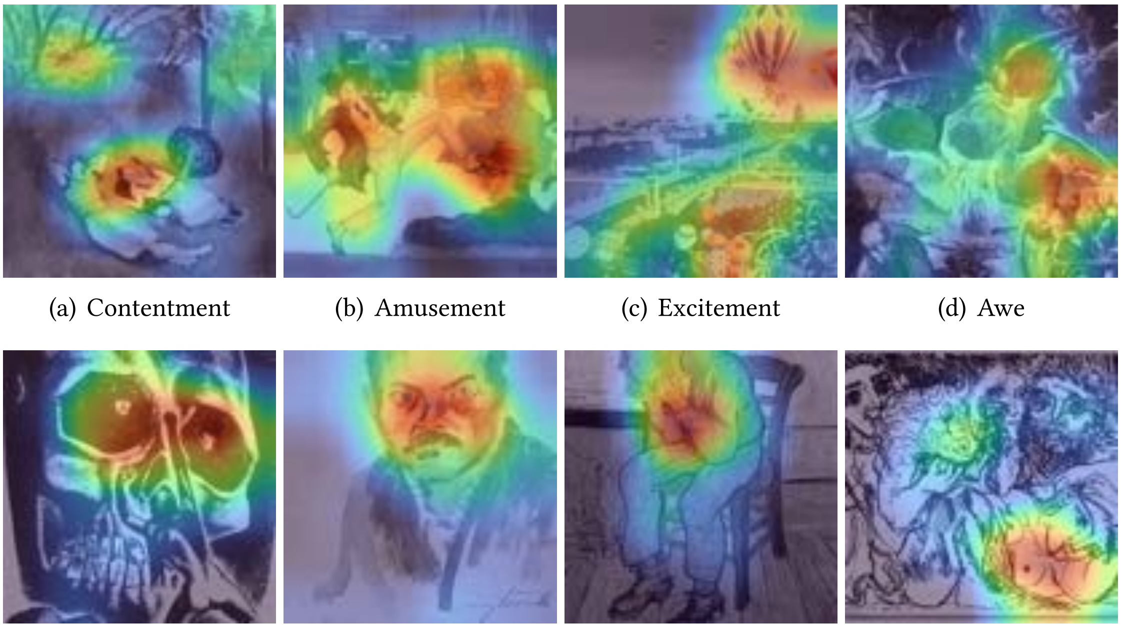
Notes: These figures present statistics on the Kullback-Leibler divergence between the predicted distribution $\{p^e\}_{e\in\mathcal{E}}$ and the target probability distribution $\{q^e\}_{e\in\mathcal{E}}$, within the training sample and the test sample. Panel **(a)** shows the distribution of this divergence metric within the training sample (light blue) and the test sample (red); the median values for each are displayed as dashed lines (long dash for training, shorter dash for test). For interpretive purposes, we also report the median divergence between predictions and randomly reshuffled labels as a fully uninformed benchmark (dashed gray line). Panels **(b)** and **(c)** show median divergence values across sub-samples defined by: (i) production period, and (ii) genre or type of exercise. The following categories are used: *Abstr.* (abstract, genre painting); *Land.* (cityscape, landscape, marina); *Still* (animal painting, flower painting, still life); *Sketch* (sketch and study); *Nude* (nude painting); *Portrait* (portrait, self-portrait); *Relig.* (religious painting); *Myth.* (mythological painting); and *Alleg.* (allegorical painting).

the divergence is slightly higher for sketches, nude, religious, and mythological paintings, reflecting their lower incidence in the annotated sample (panel c). Model performance closely follows sample size, as evidenced by the better fit for works produced between 1850 and 1950—the modal period in our corpus (see Figure 2). Overall, however, the model successfully limits over-fitting across historical periods and artistic genres.

Because the model predicts a probability distribution, $\left\{\mathbf{p}_j, \forall j \in \mathcal{J}\right\}$, and is evaluated using the Kullback-Leibler divergence (relative entropy), it is not straightforward to extract intuitive indicators of performance. To provide a more interpretable measure, we show that the actual prediction is four to five times more likely to outperform a reshuffled random prediction (panel a of Figure 5) and we compute mean squared errors in Appendix C. Within the training subsample, the mean squared error (0.056) corresponds to allocating a probability of 0.830 to the dominant emotion and 0.170 to a secondary emotion, instead of assigning probability 1 to the dominant emotion and 0 to all others. Given the limited over-fitting, the median mean squared error in the test set (0.067) is comparable—equivalent to assigning probability 0.815 to the dominant emotion and 0.185 to a secondary one.

Illustrations To illustrate the model's strengths and limitations, we contrast predictions and labels for eight images from the test sample, each selected to convey a specific emotion (see Appendix Figure C5). The model predicts well the most common emotions. For example, it identifies awe in depictions of opulent buildings or dramatic landscapes; contentment in peaceful natural scenes; excitement through warm colors, lighting, and human activity; fear through cold tones and indistinct textures that evoke uncertainty; and sadness via facial expressions or muted colors. The prediction is slightly weaker for less frequent emotions: predicting a truly multidimensional emotion vector is difficult, especially given the skewed distribution of emotions in both the annotated data and the predictions. Subtle, secondary emotions are often captured less precisely than the primary, more salient emotional tone of an image. The prediction also inherits the bias of evaluators. For instance, many images evoke contentment to evaluators, and the model tends to slightly over-predict this outcome at the expense of more distinctive emotions; and nudity is often associated with disgust, possibly reflecting moral codes and beauty standards of our times—maybe more than the artist's original intention.

Figure 6. Visualizing model attention: heatmaps of emotion-relevant image regions.



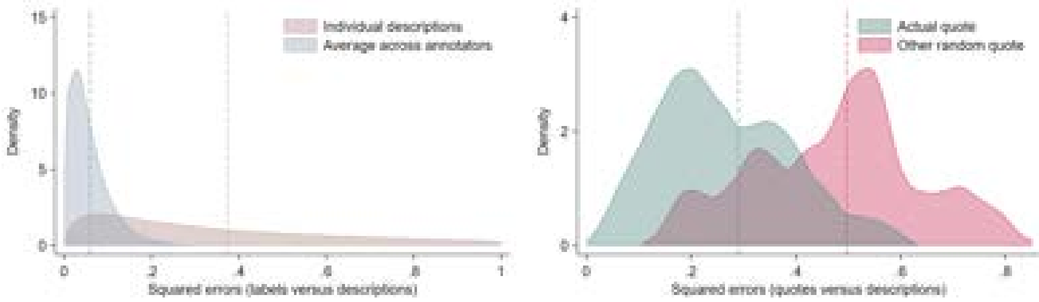
(a) Contentment (b) Amusement (c) Excitement (d) Awe

(e) Fear (f) Anger (g) Sadness (h) Disgust

Notes: This figure presents heatmaps generated using Class Activation Mapping (Zhou et al., 2016), which identifies the regions of each painting most influential in the model's prediction of latent emotional probabilities. We apply this method to the paintings with the highest emotion scores—one per emotion—among 79,860 annotated artworks; these paintings were previously shown in Figure 3. The color scale, ranging from blue to green, yellow, and red, indicates increasing activation for the corresponding emotion. For instance, amusement (panel **b**) is driven by the faces of the audience and the central dancing figure; the balloons in panel **(c)** trigger excitement; and depictions of naked bodies elicit strong signals associated with disgust (panel **h**). Additional examples are provided in Appendix C.1.

One drawback of neural networks is that they are highly complex predictors, combining numerous parameters across connected layers. This complexity often makes it difficult to identify which sources of variation are most important for prediction. While we cannot measure the contribution of individual variables to explaining the overall variance in the outcome, we illustrate the quality of the identification of probabilistic emotion scores through a set of examples. In Figure 6, we generate heatmaps using Class Activation Mapping (Zhou et al., 2016), which highlights the parts of an image most used by the network to predict a latent emotional probability. To this end, we use the eight paintings previously selected as most evocative of a dominant emotion (see Figure 3), and we report the heatmap for this dominant emotion. The results show that the algorithm often focuses on central, conspicuous elements, e.g., the angry face in panel (f), but can also infer emotion from background objects (balloons in panel c) or secondary characters (panel b). In general, bright colors and motion evoke excitement; sharp angles are associated with fear; and exposed body parts tend to trigger disgust. In Appendix C.1, we provide additional evidence, including examples of Class Activation Mapping for different emotions within the same image.

Figure 7. The internal and external validity of annotated emotions.



(a) Descriptions against labels

(b) Descriptions against quotes

Notes: Panel **(a)** displays the distribution of squared errors, $\sum_e (\vartheta^e - q^e)^2$, between the probabilistic vectors of emotions averaged across the dominant emotions chosen by annotators, $\{q^e\}_e$, and the probabilistic vectors of emotions associated with their textual justifications, $\{\vartheta^e\}_e$. The green line shows averages across textual justifications, while the beige line reports distances between the probabilistic vectors of emotions and each *individual* textual justification from each of the 15–25 annotators. Panel **(b)** displays the distribution of squared errors, $\sum_e (q^e - \rho^e)^2$, between the probabilistic vectors of emotions averaged across textual justifications provided by annotators, $\{q^e\}_e$, and the probabilistic vectors of emotions derived from quotes by the artists, $\{\rho^e\}_i$, across 34 distinct paintings (including: *The Potato Eaters* by Vincent van Gogh; *A Sunday Afternoon on the Island of La Grande Jatte* by Georges Seurat; *The Massacre of the Innocents* by Peter Paul Rubens; and *Moscow* (1916) by Wassily Kandinsky). The red distribution provides a benchmark in which quotes are randomly allocated across paintings. The textual classification of quotes into a probabilistic emotion score, $\{\rho^e\}_i$, relies on a transformer model (DistilBERT; see Sanh et al., 2019).

Predictions and the intentions of artists We exploit the free-text justifications to assess the suitability of these labels for our purposes and ask whether these annotator labels and the resulting predictions accurately capture the artist's intentions (H2).

In a first step, we test the "internal validity" of these annotations. In Section 2, we documented non-negligible heterogeneity across annotators, but also showed that the "average annotator" appears to understand the emotion categories (see Figure 3) and to be internally consistent in their classifications (see Figure 4). Panel (a) of Figure 7 analyzes the proximity between the average chosen emotions and the textual justifications:

the distribution of squared errors, $\sum_e(\vartheta^e - q^e)^2$, between the probabilistic emotion vectors averaged across annotators' chosen dominant emotions, $\{q^e\}_e$, and the probabilistic vectors of emotions derived from their textual justifications (Sanh et al., 2019), $\{\vartheta^e\}_e$, features low values. A key advantage of using a probabilistic target is that it enables a wisdom-of-the-crowd effect, attenuating the influence of annotators' idiosyncrasies (e.g., mood, cultural background, or knowledge of art and historical context).

In a second step, recognizing that our predictions essentially target the annotators' labeling process—so that any desirable or undesirable features of that process are inherited by the model—we benchmark the annotators' textual justifications against a set of artists' own descriptions of their work, covering 34 distinct paintings (panel b of Figure 7). The average emotion vector, $\mathbf{q}$, closely matches the artists' stated intentions, at least within the subset of cases for which such information is available. We further quantify the proximity between the predicted probabilistic emotion vectors and those derived from artist statements in Appendix C.1.

3.2 Identification of material representation(s)

The identification of material representations of living standards combines an external calibration with an in-house algorithm. In a first stage, we rely on a state-of-the-art large language model (GPT-5-mini) to generate: (a) an index of poverty or wealth associated with each artwork (an integer w_j between 0 and 10); (b) a relevance score measuring the informativeness of this assessment (also between 0 and 10—abstract art, for example, typically receives a score of 0); and (c) a list of depicted elements plausibly conveying information on living standards. The exact prompt used in this procedure is reported in Appendix C.2.

In a second stage, we use this calibration to train an in-house model that mirrors the structure of our sentiment prediction architecture (Section 3.1), jointly targeting the wealth index (a) and the relevance score (b) on the WikiArt sample of artworks. Given the close similarity between this procedure and our sentiment pipeline (and the lower dimensionality of the target vector), we refer the interested reader to Appendix C.2 for details on model estimation and fit.

Acknowledging that the inference of living standards via the large language model (and, by extension, our calibrated in-house algorithm) constitutes a black box, we translate the structured artwork descriptions (c) into measures of living standards using a more transparent taxonomy.[14] Specifically, we construct a wealth-leaning corpus and a

[14] As detailed in Appendix C.2, the large language model associates poverty with features such as sparse or rudimentary settings and interiors, worn objects and clothing, depictions of physical weakness, and forms of manual or rural labor. Conversely, wealth is associated with refined or ornate environments, personal luxury items (e.g., precious metals, fine textiles, jewelry), bodily presentation, contextual markers

poverty-leaning corpus across eight domains, organized into four groupings. The first grouping (0) contains vocabulary that is not informative about wealth or poverty, either because terms are intrinsically neutral (0.a) or because they require additional context to be interpreted (0.b; e.g., *textile*). The second grouping—1. Objects—comprises vocabulary related to decorative intensity, material quality, and abundance (1.a; e.g., *delicate* versus *rough*, *intact* versus *patched*, *detailed* versus *sparse*), as well as luxury goods, portable wealth, furniture, and interiors (1.b; e.g., *necklace*). The third grouping—2. People—includes terms describing presentation, expression, and posture (2.a), as well as physical markers of social status (2.b; e.g., *wig*). The final grouping captures *contextual* elements extracted from scene-level interpretations (3.a; e.g., *pleading*) and background elements establishing the general setting (3.b).

This taxonomy yields a structured set of features and descriptive attributes, each assigned a weight, which disciplines and aggregates the raw descriptions into quantitative and interpretable indicators of material living standards. For each painting j, we construct a sub-index ι_j^d for each domain $d \in \mathcal{D}$ by summing positive connotations—the number of terms in the wealth-leaning corpus—and subtracting negative connotations—the number of terms in the poverty-leaning corpus. Our main description-based indices aggregate these sub-indices: without any differential weights, $w_j^u = \sum_{d \in \mathcal{D}} \iota_j^d$; and with weights α_d derived from regressing the large-language-model wealth index (w_j) on the full set of sub-indices, $w_j^w = \sum_{d \in \mathcal{D}} \alpha_d \iota_j^d$.

3.3 Validation of extracted signals

Using emotional and material signals from artworks as proxies for their context raises two main challenges. First, digitized artworks constitute a non-random sample shaped by tastes, commercialization, patronage, museum acquisition practices, and historical contingencies. Second, the production of paintings is influenced by demand-side forces that complicate the interpretation of their visual content: art may reflect the preferences of those able to afford it, as a luxury good; artworks are assets, subject to taxation, inheritance laws, and the relative returns to alternative investments (e.g., the price of art doubled in Britain after the Slavery Abolition Act of 1833); and demand may originate from both non-market actors (e.g., the Court or the Church) and market participants, reflecting broader patronage structures. We address both concerns by validating our data inputs *and* our data outputs.

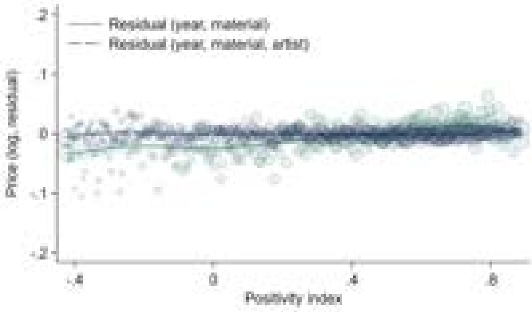
Attrition, commercialization, and selection into conservation Our collection of paintings reflects a historical selection process shaped by the tastes of audiences at the

of social status, and elaborate architecture.

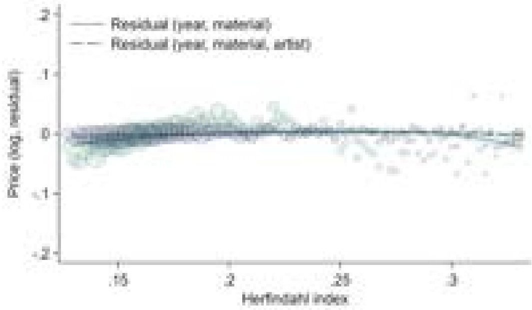
Figure 8. Attrition, commercialization, and selection into conservation.

(a) Commercialization by genre

(b) Commercialization and emotion



(c) Positivity (sentiment)



(d) Herfindahl (sentiment)

Notes: These figures summarize (i) the selection of historical artwork into commercialization—the statistics are then generated for artwork produced in a given country and year for which at least 10% of it can be associated with an artist present in the sales catalogs (*Getty Provenance Index*)—and (ii) the valuation of modern artwork as a function of its emotional content. Panel **(a)** displays the attrition rate for our artwork (i.e., the share that can be attributed to an artist whose work features in the sales catalog), as a function of its genre or type of exercise. Panel **(b)** displays the average emotion probabilities for artwork that can be attributed to an artist whose work features in the sales catalog (in darker blue) versus artwork that is attributed to "non-selling" artists (in green). Panel **(c)** displays the correlation between the (log, residualized) price and a sentiment positivity index derived from predicted emotional scores on modern artwork. Panel **(d)** displays the correlation between the (log, residualized) price and a Herfindahl index derived from predicted emotional scores. In both panels **(c)** and **(d)**, the green plain curve residualizes the (log) price by yearly dummies and object characteristics (height, width, material); the blue dashed curve residualizes the (log) price by yearly dummies, object characteristics, and artist fixed effects.

time of production and of later audiences, the idiosyncrasies of wealth accumulation, inheritance dynamics, museum donations, and the destruction of artworks through wars and revolutions. This selection process has been studied in the less complex context of non-rival goods (see Cohen, 1999; Moretti, 2000, for a discussion of the "great unread" in literature), raising the question of which works become part of the literary canon in a given context (Barré et al., 2023). By contrast, our focus on rival goods implies that the "great unseen" is, by construction, unobserved. We therefore assess the importance of these forces in shaping variation in visual content over time by leveraging (i) digitized sales catalogues covering approximately 1.2 million transactions between 1650 and 1950 in the Low Countries, France, Great Britain, and Germany, and (ii) observed recent artworks listed on the largest auction platform.

How do messages conveyed by artworks correlate with their commercialization? To evaluate the potential bias introduced by commissioned or commercialized art, Figure 8 examines whether artists represented in the auction sample—those whose works were commercialized or later appeared in digitized auctions—produced paintings that differ in technique or emotional expression from those of their peers. Panel (a) shows that between 40% and 60% of preserved paintings can be attributed to artists whose work appears in the sales catalogues, with broadly similar coverage across genres and artistic exercises. Panel (b) examines whether commercialized artworks differ in their conveyed emotions: the average emotional profile of artists included in the sales catalogues (in dark blue) closely matches that of non-selling artists (in green).

Does the emotional signal conveyed by an artwork systematically correlate with its commercial value? To assess the potential biasing role of demand, we focus on artworks advertised and sold through *Artsy* and exploit the joint information on modern artwork valuations and each painting's predicted emotional profile. Panels (c) and (d) of Figure 8 show that commercial value does not systematically correlate with overall sentiment or with the expression of extreme emotions (e.g., paintings strongly associated with a single emotion). Appendix C.3 complements this analysis by comparing artwork prices to each separate emotion and to physical characteristics.

Which artworks are selected for preservation and digitization? This question is more challenging, as non-preserved artworks are unobserved. However, we can compare auction prices between 1650–1950 for artists whose works appear in our corpus and those of "omitted" artists. We find a modest price premium of roughly 10% for artists represented in our main collection—identified within the same auction. Thus, while there is selection into preservation and commercialization, these forces may not be the *primary* driver of variation in visual content over time. The strongest argument in support of our approach and data, however, lies in quantifying the mapping between well-being and living standards, and their representation in art; we turn to this next.

The mapping between well-being and conveyed emotions Our approach is motivated by the conjecture that painters are influenced by their social context, and that their work reflects elements of these experiences. Through this channel, societal sentiment should map into the emotions conveyed by contemporaneously produced paintings. Because societal sentiment is largely unobserved for much of our period of interest, we establish the existence of this mapping using modern artwork drawn from *Artsy*, the largest art market platform. An extensive analysis is provided in Appendix C.4.

Table 1. The mapping between reported well-being and conveyed emotions.

Positivity index	(1)	(2)	(3)
Level of satisfaction (0-1, Gallup)	0.3572	0.3492	0.2817
	(0.0831)	(0.0821)	(0.0603)
	[0.278]	[0.271]	[0.219]
Observations	90,451	90,451	89,474
Fixed effects (country, year)	Yes	Yes	Yes
Fixed effects (artist)	No	No	Yes
Controls (GDP, price)	Yes	Yes	Yes
Controls (Medium, support)	No	Yes	Yes

Notes: Standard errors are reported between parentheses and are clustered at the level of a country/year. Standardized effects, obtained by normalizing the dependent and explaining variables with their standard deviations across country/year observations, are reported in square brackets. Each artwork is weighted to account for the probability that the artist is from this country—this probability is high with an average of 93.5%—and to ensure that each country/year contributes the same—thus down-weighting countries with many producers. The dependent variable is the sentiment positivity index derived from predicted emotional scores, and the explaining variable is the average satisfaction of the artist cohort—gender and age "decade"—within the country of the artist in the year of production (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025). The satisfaction variable in Gallup, originally between 0 and 10, is divided by 10 to be on a 0-1 scale. The set of controls include (log) GDP per capita across all specifications, the (log) price of the artwork across all specifications, and dummies for the material (e.g., oil gouache, ink) and support (e.g., canvas, paper, silk) in column **(2)** and column **(3)**. The set of fixed effects include: year and country fixed effects in column **(1)** and column **(2)**; and adding artist fixed effects in column **(3)**.

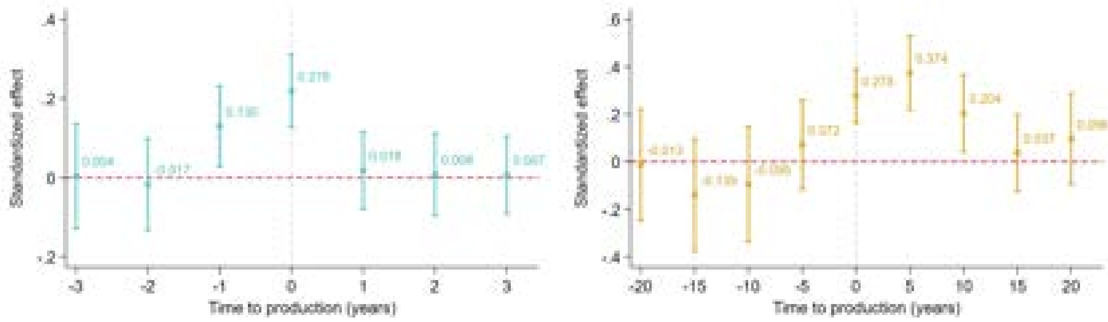
Table 1 illustrates the *general* relationship between societal sentiment and emotions expressed in artwork—conditional on material living standards. Letting j denote a given artwork, produced by an artist belonging to cohort c (defined by age and gender) in location ℓ and year t, we estimate the following fixed-effects specification:

$$l_{j,c,\ell,t} = \alpha r_{c,\ell,t} + \gamma Y_{\ell,t} + \beta \mathbf{X}_{j,\ell,t} + \nu_\ell + \eta_t + \varepsilon_{j,\ell,t}, \tag{1}$$

where $l_{j,c,\ell,t}$ is a positivity index constructed by summing the positive emotion scores associated with artwork j (contentment, excitement, amusement, awe) and subtracting the negative ones (fear, anger, sadness, disgust); $r_{c,\ell,t}$ is reported life satisfaction for the corresponding cohort within country ℓ and year t (Helliwell et al., 2025); $Y_{\ell,t}$ is (log) output per capita, capturing material living standards; $\mathbf{X}_{j,\ell,t}$ includes covariates such as the (log) price of the artwork and dummies for material (e.g., oil, gouache, ink) and support (e.g., canvas, paper, silk); and ν_ℓ and η_t denote location and year fixed effects.

Comparing average life satisfaction reported by Gallup respondents within a given country, year, gender, and age cohort to the emotions conveyed by artists with the same characteristics yields four main findings. First, an increase in reported satisfaction equiv-

Figure 9. Validation of extracted signals—timing of production.



(a) Well-being and conveyed emotions

(b) Actual and conveyed living standards

Notes: These figures provide complementary evidence to Table 1 and Table 2. Panel **(a)** reports the standardized estimate corresponding to column (3) of Table 1, but with lagged and lead values of reported life satisfaction as right-hand side variables. Negative (positive) values on the x-axis correspond to correlations between lagged (lead) satisfaction and contemporaneous positivity conveyed in paintings, l_j. Panel **(b)** reports the standardized estimate corresponding to column (3) of Table 2 (Panel A). Negative (positive) values on the x-axis correspond to correlations between lagged (lead) log output per capita and contemporaneous material living standards conveyed in paintings, w_j.

alent to the average gap between Lebanon and Denmark (about 0.45 on the normalized 0-1 scale) is associated with a 0.16 increase in positivity conveyed in modern artwork (columns 1 and 2 of Table 1), corresponding to a sizable standardized effect of approximately 0.28. Second, this relationship is robust to the inclusion of artist fixed effects, identifying changes in emotional expression within an artist's own production over time (column 3 of Table 1). Third, the mapping is primarily contemporaneous: it is observed between current measures of life satisfaction and artistic expression, with limited anticipatory or lagged effects (Figure 9, panel a). Fourth, the estimates also hold *within* contexts defined by location ℓ and year t: for example, countries in which women report lower well-being also tend to produce female artists whose work conveys more negative emotions (Appendix C.4).

Appendix C.4 provides complementary analyses: (a) decomposing the mapping between life satisfaction and conveyed emotions across the nine emotion categories; (b) using alternative welfare measures; and (c) showing that the relationship between reported emotions (e.g., worry, joy) and conveyed emotions varies by type—for instance, depicted fear is most strongly correlated with worry, while depicted contentment is most strongly correlated with joy.

Material representations as a measure of living standards To shed light on the mapping between representations of material living standards and actual living standards, we focus on artwork produced since 1900—an era for which historical estimates of output per capita are available for a large number of countries (Bolt and Van Zanden,

Table 2. Material representations as a measure of living standards.

Panel A: Wealth index	(1)	(2)	(3)
Output per capita	0.614	0.545	0.369
	(0.126)	(0.118)	(0.076)
	[0.438]	[0.388]	[0.263]
Observations	81,396	81,393	81,393
Panel B: Descriptions (unweighted)	(1)	(2)	(3)
Output per capita	1.657	1.491	1.081
	(0.442)	(0.336)	(0.224)
	[0.369]	[0.332]	[0.241]
Observations	81,396	81,393	81,393
Panel C: Descriptions (weighted)	(1)	(2)	(3)
Output per capita	0.485	0.427	0.294
	(0.119)	(0.099)	(0.067)
	[0.383]	[0.338]	[0.232]
Observations	81,396	81,393	81,393
Panel D: Predictions	(1)	(2)	(3)
Output per capita	0.560	0.502	0.305
	(0.143)	(0.142)	(0.088)
	[0.449]	[0.402]	[0.244]
Observations	81,391	81,388	81,388

Notes: Standard errors, reported between parentheses, are clustered at the level of a half-century/country. Standardized effects, obtained by normalizing the dependent and explaining variables with their standard deviations across country/year observations, are reported in square brackets. Each specification is weighted by a score from 0 to 1 capturing (i) how informative a painting is of living standards and (ii) possible duplicates across data sources. The dependent variable is (log) GDP per capita (Bolt and Van Zanden, 2025), and the explaining variables are: a wealth index w_j from 0 to 10 in panel **A**; an unweighted description-based index w_j^u adding all categorical feature sub-indices (1.a Objects: ornamentation, material quality, abundance; 1.b Objects: luxury goods, portable wealth, furniture, and interior; 2.a People: presentation, grooming, expression, and posture; 2.b People: physical markers of social status; 3.a Context: scene-level interpretation, social positioning; 3.b Context: background environment) in panel **B**; a weighted description-based index w_j^w in panel **C**; and our in-house prediction $\widehat{w}_j$ in panel **D**. See Appendix C.2 for details. The set of baseline controls include: year and country fixed effects in column **(1)**; adding artist age and exercise fixed effects in column **(2)**; and adding all emotion probabilities (p_j^e) in column **(3)**.

2025). We estimate a specification analogous in spirit to Equation (1), relating artwork-specific wealth indices extracted in Section 3.2 to (log) GDP per capita, while weighting for the relevance score and controlling for year, country, artist age, and exercise fixed effects, as well as predicted probabilities for all conveyed emotions.

Table 2 documents a strong and robust relation between representations of wealth or poverty in artwork and actual living standards. This mapping holds both in a baseline two-way fixed-effects specification, controlling for year and country effects, and in richer specifications that additionally include artist-age fixed effects, dummies for each artistic exercise, and emotion probabilities. Quantitatively, a one-log-point increase in output per capita—equivalent to sustained annual growth of 2% over 50 years—raises the wealth index derived from the large language model, w_j (ranging from 0 to 10; see Panel A), by 0.37−0.61, corresponding to a standardized effect of 0.26−0.44.

This mapping holds primarily for measures of actual and depicted living standards that are close in time: the correlation is not statistically different from zero when the production date is separated by more than 15 years from the measurement of living standards (panel b of Figure 9). Panels B to D report similar relationships for indices constructed from artwork descriptions, w_j^u and w_j^w, as well as for our in-house wealth prediction, $\widehat{w}_j$.

Appendix C.5 complements these results by (i) identifying the most predictive descriptive features—distinguishing between objects, people, and contextual clues; (ii) estimating the relationship by painting genre; and (iii) providing visual evidence illustrating both these correlations and the dynamics of material representations over time. We find that the inference of living standards primarily relies on depictions of people (including physical appearance, posture, and markers of social status), contextual cues aiding social positioning (e.g., a character begging or a figure surrounded by servants), and the quality of (often domestic) objects. For example, a 10% increase in output per capita is associated with 0.058 additional descriptive words such as "refined", "delicate" or "patterned" in artwork descriptions (c) by the large language model. Finally, this inference is shown to be generally feasible across many painting genres (religious, mythological, portraits, and genre scenes), with the notable exception of landscapes and abstract art.

The previous mapping underscores that the common variation underlying our measure ($\widehat{w}_j$), weighted by the informativeness of each artwork, closely tracks variation in underlying economic conditions.

4 Emotional and material welfare through socioeconomic transformations

The previous sections describe how we extract two types of signals from paintings: emotional indicators, captured by the probabilistic vector of emotions $\mathbf{p}_j$; and material indicators, summarized by predicted wealth $\widehat{w}_j$ (and the informativeness measure $\widehat{z}_j$). This section leverages these painting-specific characteristics to study emotional and material welfare during periods of socioeconomic transformation.

Our focus is not on causal identification, but on the co-movement between painting-based indicators and contemporaneous economic and social conditions. We proceed in three steps. First, we document the temporal dynamics of these signals across countries. Second, we relate them to summary proxies for economic conditions—economic development, uncertainty, and inequality. Third, we examine how societies respond to a set of well-defined socioeconomic changes: climate, trade, technology, social order (as captured by academic production), and the (geo)political environment.

4.1 The dynamics of extracted signals across countries

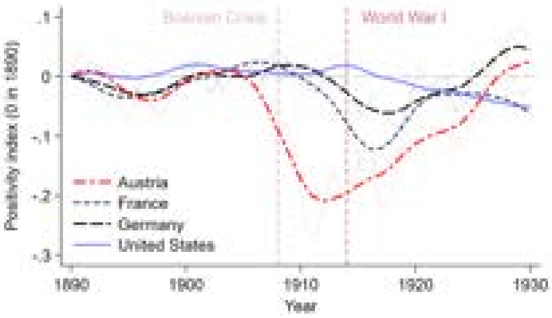
We begin by assessing whether aggregated signals extracted from paintings in a given historical context (ℓ, t) exhibit systematic and meaningful variation across countries and over time. Interested readers can refer to our companion website, which provides a comprehensive exploration of these time series, along with a description of representation dynamics within individual artists' oeuvres.

In Appendix C.5 and Appendix D.1, we document the long-run dynamics of wealth indices and the evolution of emotional expression within the present-day borders of the twelve countries with the largest number of paintings in our sample: France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. We distinguish between negative emotions (fear, sadness, and anger; Appendix Figure D1), positive emotions (excitement, amusement, and contentment; Appendix Figure D2), and remaining emotions (awe, disgust, and other; Appendix Figure D3). We also aggregate all emotion probabilities into a positivity index, $l_{\ell,t}$, and analyze its cross-country behavior. Finally, we examine within-context dispersion in expressed sentiment as a proxy for inequality.[15] The evidence points to cross-country differences in emotional expression. For example, Great Britain and the United States are persistently more positive before 1900, whereas Italy, Spain, and Germany exhibit more negative expressions; France, Spain, Belgium, and Austria display greater volatility in conveyed emotions over time. There are also common long-run trends: the pronounced negativity observed during the Little Ice Age and the twentieth century contrasts with a more positive period spanning the eighteenth and nineteenth centuries. However, a large share of medium-run fluctuations is country-specific, suggesting that the extracted signals respond to local historical conditions rather than merely reflecting broad shifts

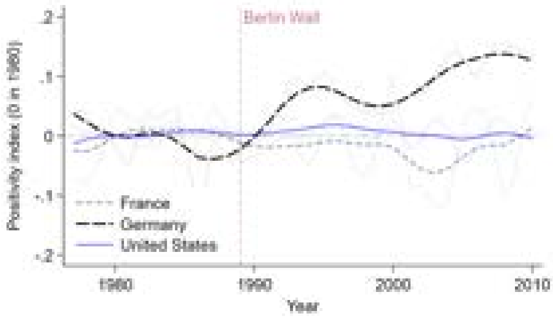
[15] To assess divergent views within a given context, one could alternatively compute average emotions for different groups of artists, for example by socio-economic status. While artists reflect the spirit of their time, their experiences may differ systematically with their social position—as illustrated, in a modern setting, by our within-context, across-demographic-group mapping between reported and expressed well-being (Appendix C.5).

in artistic style.[16]

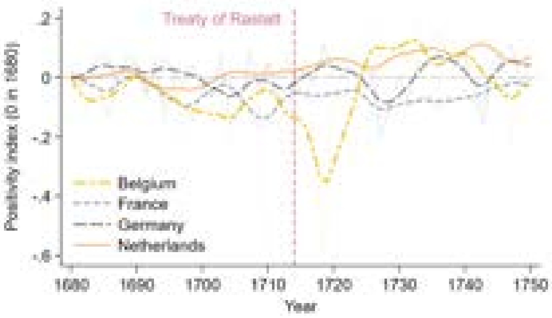
Figure 10. The dynamics of emotions around selected events.



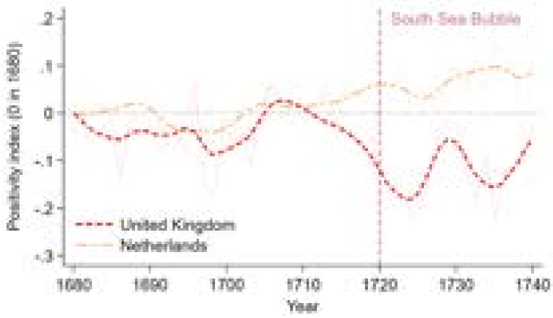
(a) World War I



(b) The fall of the Berlin Wall



(c) Treaty of Rastatt (Spanish Netherlands)



(d) South Sea Bubble

Notes: These figures illustrate the dynamics of the positivity index, $l_{\ell,t}$, constructed by summing positive emotion scores (contentment, excitement, amusement, awe) and subtracting negative ones (fear, anger, sadness, disgust). Panel **(a)** shows positivity in Austria, France, Germany, and the United States around World War I; dashed lines mark the Bosnian Crisis (1908) and the assassination of Franz Ferdinand (1914). Panel **(b)** shows positivity in France, Germany, and the United States around the fall of the Berlin Wall (1989). Panel **(c)** shows positivity in Belgium, France, Germany, and the Netherlands around the Treaty of Rastatt, whereby Austria obtained the Spanish Netherlands—including most of present-day Belgium but excluding the Netherlands. Panel **(d)** shows positivity in the Netherlands and the United Kingdom around the burst of the South Sea Bubble. For illustration, raw time series are shown as thin lines, while smoothed series—obtained using a Hodrick-Prescott filter with smoothing parameter 6.25—are shown as thicker lines. All time series are normalized to 0 in 1890 (**a**), 1980 (**b**), and 1680 (**c**, **d**).

To further illustrate these country-specific variations and the informational content of the extracted emotional signals, we focus on four well-identified historical events: (i) World War I, a large-scale external conflict; (ii) the fall of the Berlin Wall, a major political transition; (iii) the Treaty of Rastatt, which transferred control of the Spanish Netherlands (most of modern-day Belgium) to Austria; and (iv) the burst of the South Sea Bubble, a financial crisis with uneven international exposure. These events provide clear and interpretable variation in exposure across countries, allowing us to compare how emotional responses differ between directly affected and less affected regions. They differ along three dimensions: the type of shock (external conflict, political transition,

[16]Overall, idiosyncratic variation dominates these dynamics: less than 2% of the variance is explained by common yearly fluctuations across countries or by fixed cross-country differences.

or financial crisis), the degree of geographic concentration, and the expected impact on local populations. Panel (a) of Figure 10 shows a pronounced decline in positivity around World War I in Austria, France, and Germany, with drops of up to 0.20–0.25 in the years surrounding the conflict, while the United States exhibits a more muted response. Notably, the decline begins as early as 1908 in Austria, following the Bosnian Crisis and the resulting tensions between Austria-Hungary, Serbia, Italy, and Russia. Panels (b), (c), and (d) illustrate more localized events. Panel (b) highlights the fall of the Berlin Wall, where positivity increases markedly in Germany relative to France and the United States, consistent with the political and economic significance of reunification. The Treaty of Rastatt is associated with a sharp but temporary decline in positivity in Belgium, the region most directly affected by the territorial transfer, while neighboring countries display more limited responses (panel c). Finally, the burst of the South Sea Bubble is associated with a pronounced and persistent decline in positivity in Great Britain relative to another major trading power, the Netherlands (panel d).

Across all panels, two patterns emerge. First, major events are associated with sizable shifts in emotional expression, often on the order of 0.10 to 0.30 in the positivity index—comparable to or larger than a standard deviation in context-level positivity $l_{\ell,t}$ (approximately 0.14). Second, these responses are highly heterogeneous across countries, reflecting differences in exposure to and interpretation of the same historical shocks. Together, these patterns suggest that the extracted emotional signals capture meaningful, context-specific variation in experienced well-being.

4.2 Empirical strategy

To assess how the signals extracted from paintings reflect socioeconomic conditions, we relate painting-level indicators to the economic environment prevailing at their time and place of production. Each painting thus constitutes an observation linking artistic expression to the broader historical context in which it was created.

The data construction described in Section 3 yields a dataset of paintings with the following attributes: extracted emotional and material welfare, artistic sub-genres, title, author, as well as year and location of production. While our objective is not to establish causal effects in a strict sense, we implement a demanding empirical design to identify credible co-movements between the signals conveyed in paintings and contemporaneous socioeconomic conditions.

Let j denote a painting and y_j an associated welfare indicator—for instance, the predicted probability of evoking emotion e. The painting is produced by artist a in location ℓ and year t, characterized by socioeconomic conditions $\mathbf{s}_{\ell,t}$. We estimate the equivalent to a standard two-way fixed effects specification,

$$y_{j,\ell,t} = \alpha \mathbf{s}_{\ell,t} + \beta \mathbf{X}_{j,\ell,t} + \nu_\ell + \eta_t + \varepsilon_{j,\ell,t}, \tag{2}$$

where $\mathbf{X}_{j,\ell,t}$ includes covariates such as artist fixed effects, $\alpha_{a(j)}$, as well as indicators for the artist's age and the type of exercise; ν_ℓ and η_t denote location and year fixed effects.

This specification is demanding, as it exploits within-artist variation in signals over the production cycle, comparing how an artist's work evolves across different socioeconomic conditions. By controlling for persistent differences across artists, locations, and time, the approach limits the scope for spurious correlations driven by patronage, stylistic differences, long-run trends, and country-specific idiosyncrasies. In Appendix D.2, we quantify the relative importance of these components: country differences account for less than 2–3% of the overall variance in depicted emotions; common yearly dynamics across countries explain less than 8% for most emotions; and unobserved artist-specific factors account for 40–50% of the variance. Augmenting artist fixed effects with context-specific effects, (ℓ, t), increases explanatory power by 3–4 percentage points. The subsequent analysis shows that the remaining *within-artist variation*, while limited, is nonetheless sufficient to capture systematic fluctuations in emotional and material signals over time.[17]

One variant of the previous model replaces observable measures of socioeconomic conditions with a full set of context-specific effects for each location and year. This approach, closely related to the AKM decomposition in the labor literature (Abowd et al., 1999), recovers a latent measure of welfare at the context level while controlling for artist-specific factors. In our setting, it amounts to comparing how artists' emotional and material expressions vary across contexts, net of persistent differences across artists. In practice, this approach relies on strong assumptions, including the separability of artist and context effects and sufficient mobility of artists across contexts (see, e.g., Bonhomme et al., 2020). Given potential concerns about non-random mobility, we view this specification as a useful conceptual benchmark, but focus on specifications that relate painting-level signals to observable socioeconomic variables while controlling for artist, location, and time fixed effects.

Derivatives of the emotion scores Our empirical evidence focused on emotional welfare primarily relies on the probabilistic vector of emotions, $\mathbf{p}_j$, associated with each painting j. In addition, we associate indices to this probabilistic vector of emotions. The

[17] In Appendix D.2, we use a sample of 30 prominent artists and highlight variation in the content of their work and how their emotional expression may reflect broader societal changes. Our companion website provides a more comprehensive exploration of variation across artists and within their oeuvre.

first index combines the nine emotion scores into a single value,

$$l_j = \sum_{e \in \mathcal{E}^+} p_j^e - \sum_{e \in \mathcal{E}^-} p_j^e,$$

where l_j represents a positivity index ranging from -1 to 1, calculated by summing the positive emotion scores (contentment, excitement, amusement, awe) and subtracting the negative ones (fear, anger, sadness, disgust). This dimensionality reduction might help better visualize the general dynamics of emotions conveyed by different paintings, but such a simplistic index may overlook important nuances: both fear and sadness are negative emotions, but fear may be forward-looking and tied to uncertainty, while sadness is more often backward-looking; similarly, contentment and excitement are both positive, though contentment tends to reflect calmness. Emotions like awe and disgust are also more difficult to interpret or categorize.[18] We finally construct an artwork-specific measure of disagreement, d_j, defined as the distance to the average context vector, $\bar{\mathbf{p}}_{\ell,t}$,

$$d_j = \frac{1}{|\mathcal{E}|} \sum_{e \in \mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|.$$

4.3 The perception of changes in living standards

We examine how emotional signals extracted from paintings relate to key indicators of the macroeconomic environment: economic development, uncertainty, and inequality.

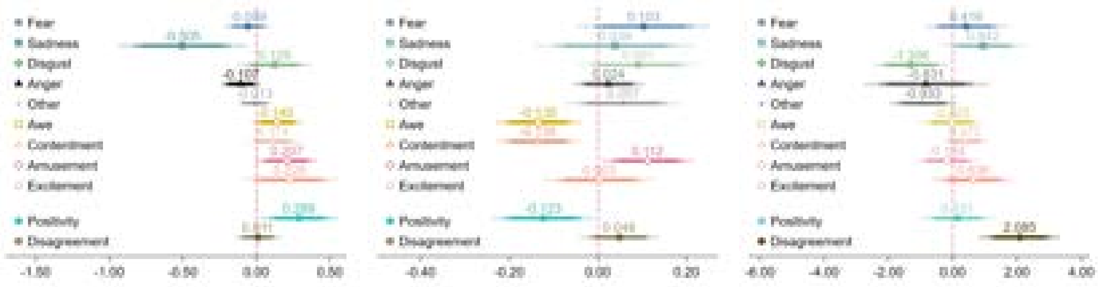
The perception of changes in living standards A visual inspection of time variation in depicted emotions across major economies offers intuitive insights into their economic and social significance. Sadness peaks during periods of hardship for vulnerable populations—such as the Industrial Revolution or the Little Ice Age—whereas emotions such as excitement or amusement are more prevalent during periods of rising living standards (Appendix Figures D1, D2, and D3). Depictions of fear coincide with periods of heightened uncertainty, including civil conflicts (e.g., the Huguenot diaspora in France), regime transitions (e.g., the fall of communism), and major political shifts (e.g., the Meiji Restoration in Japan). Conversely, contentment moves inversely with fear and becomes more prominent during stable periods. Finally, episodes of high economic in-

[18] To better understand the latent variation underlying our nine correlated variables, we perform a Principal Component Analysis (PCA) on the probabilistic vector of emotions, $\mathbf{p}_j$, and select the eigenvectors of the covariance matrix with eigenvalues above or close to 1. The loadings for each emotion are reported in Appendix D.2 and Appendix Table D1. The first principal component ($p_j^{c_1}$) loads positively on negative emotions (fear, anger, sadness) and negatively on contentment. The second principal component ($p_j^{c_2}$) loads positively on positive emotions (amusement) and disgust. The third component ($p_j^{c_3}$) captures awe, while the fourth component ($p_j^{c_4}$) predominantly reflects the residual category "other".

equality are associated with elevated emotional disagreement—for instance, in France and Great Britain at the turn of the twentieth century (Appendix Figure D5).

In Figure 11, we formalize these patterns by examining whether the predicted probability of evoking each emotion $e \in \mathcal{E}$, denoted p_j^e, as well as the positivity index l_j and the disagreement index d_j, correlate with three key macroeconomic indicators: (a) economic development, measured as the logarithm of GDP per capita (Bolt and Van Zanden, 2025); (b) economic uncertainty, proxied by six-year volatility in GDP per capita (Bloom, 2014); and (c) wealth inequality, measured by the share of total wealth held by the top 1% (Alvaredo et al., 2020). These relationships are estimated using specification (2), which leverages within-artist variation while controlling for year, country, harmonized exercise, and artist age fixed effects.

Figure 11. The perception of changes in living standards.



(a) GDP per capita (b) Uncertainty (c) Wealth inequality

Notes: These figures display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to indicators of economic development and inequality. We consider Equation (2), with location, year, artist, age, and genre fixed effects, and replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with: **(a)** recently revised estimates of (log) GDP per capita (Maddison Project Database, see Bolt and Van Zanden, 2025); **(b)** the volatility of GDP per capita calculated over a window of [-2,+4] years ($\sum_{\tau=-2}^{4} |\widehat{y}_{t+\tau}|$, controlling for $\sum_{\tau=-2}^{4} \widehat{y}_{t+\tau}$); and **(c)** the concentration of wealth among the top-1% wealthiest individuals, mostly available from the mid-nineteenth century onward (Alvaredo et al., 2020). For **(b)**, the measures $\widehat{y}_t$ are residualized using a Hodrick-Prescott filter with a smoothing parameter of 100. The coefficients represent separate regressions with the following left-hand side variables: the predicted probability of expressing emotion e, p_j^e for painting j, where e is *fear* (dark-blue, plain circle), *sadness* (light-blue, plain square), *disgust* (teal, plain diamond), *anger* (black, plain triangle), *other* (gray, cross), *awe* (gold, hollow square), *contentment* (salmon, hollow triangle), *amusement* (pink, hollow diamond), and *excitement* (orange, hollow circle); the *positivity* index, $l_j = \sum_{e\in\mathcal{E}^+} p_j^e - \sum_{e\in\mathcal{E}^-} p_j^e$, in light blue; and a measure of disagreement, defined as the sum of deviations from the mean emotion in a given context, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|$. All left-hand side variables are standardized, so the reported estimates can be interpreted in terms of standard deviations across country/year observations. The bands represent 10%, 5%, and 1% confidence intervals, and standard errors are clustered at the level of a country/half-century.

The effect of economic development Figure 11 (a) shows that a 169% increase in GDP per capita—equivalent to sustained annual growth of 2% over 50 years—is associated with a 0.50 standard deviation decline in sadness, alongside moderate increases in amusement, excitement, and awe. Overall, the positivity index rises by 0.29 standard deviations, driven primarily by the reduction in sadness and the presence of more positive emotional tones.[19]

To gauge the magnitude of these effects, we leverage the empirical mapping between conveyed emotions and reported life satisfaction documented in Section 3.3: the increase in positivity associated with sustained annual GDP per capita growth of 2% over 50 years corresponds to a very significant one-standard-deviation improvement in life satisfaction. These sizable effects do not, however, capture the full emotional response to changes in living standards: conveyed emotions also react strongly to other dimensions of the economic environment, notably uncertainty and inequality.

The perception of uncertain times Figure 11 (b) shows that economic uncertainty, defined as the sum of absolute short-run output deviations ($\sum_{\tau=-2}^{4} |\widehat{y}_{t+\tau}|$), triggers a very distinct emotional pattern. In uncertain times, artists depict higher levels of *unstable* emotions (fear and amusement) and lower levels of stable emotions such as contentment. Quantitatively, moving from the 10th to the 90th percentile of economic uncertainty (approximately 0.42) is associated with a 0.058 standard deviation decline in both awe and contentment, alongside a mirrored 0.043 standard deviation increase in fear.

The perception of wealth inequality Finally, wealth inequality primarily influences the *dispersion* of emotions. For instance, the share of total wealth held by the top 1% in Europe declined by roughly 40 percentage points between the late nineteenth century and 1980, before rising again by about 10 points in recent decades. Such a 10-percentage-point increase in wealth concentration would be associated with a very substantial 0.21 standard deviation increase in the disagreement index (Figure 11, panel c).

Taken together, these findings indicate that the extracted emotional signals respond systematically to economically meaningful features of the environment. While rising living standards constitute the dominant long-run socioeconomic transition over this period, their emotional imprint is complemented by forces related to economic uncertainty and inequality. More broadly, the results provide a framework for interpreting emotional fluctuations: changes in poverty map into changes in conveyed sadness; shifts in contentment and fear reflect transitions between stable and uncertain environments; and variation in emotional disagreement signals changes in inequality.

4.4 The effect of large socio-economic transformations

Having established that indicators extracted from paintings respond to core macroeconomic trends, we next examine how material and emotional signals behave across

[19]In Appendix D.1, we contrast our main emotion-based indicators—p_j^e, l_j, and d_j—against additional measures of living standards, including sentiment trends in books from select countries (Hills et al., 2019) and historical life expectancy at birth (Zijdeman and Ribeira da Silva, 2015).

different types of historical transformations. The key question is whether these two dimensions of welfare move together, or whether emotional expression also captures aspects of societal experience that are not reflected in material living standards.

We organize the analysis around three broad types of changes: (i) shocks affecting subsistence conditions, where emotional and material welfare are expected to move closely together; (ii) forces driving medium-run economic development, where the relationship may be weaker or heterogeneous; and (iii) institutional and political environments, where changes in welfare may arise even in the absence of contemporaneous changes in material conditions. Across all exercises, we estimate specification (2) using the predicted emotion probabilities $\{p_j^e\}$, the positivity index l_j, the disagreement index d_j, and the wealth prediction $\widehat{w}_j$ as dependent variables.

Subsistence shocks We begin with shocks that directly affect contemporaneous living standards. Prior to industrialization, much of the population lived close to subsistence levels, making livelihoods highly sensitive to climate.

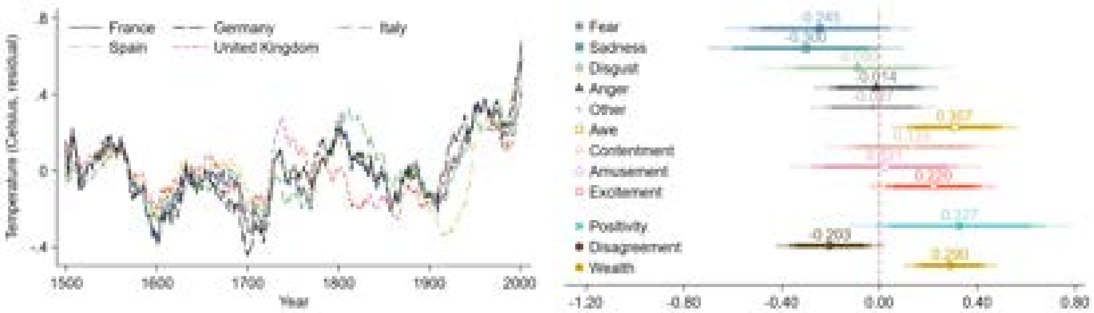
Many studies have examined the long-run impact of climatic shocks on European economies prior to the Industrial Revolution. For example, Waldinger (2022) investigates the effects of climatic variability during the Little Ice Age on agricultural productivity and mortality, while Oster (2004) explores scapegoating and witch trials in response to the same climatic shock. Some paintings directly reflect these climatic fluctuations—most notably the increased prevalence of winter landscapes, such as *The Hunters in the Snow* by Pieter Bruegel the Elder. However, the broader societal effects of these fluctuations (see panel a of Figure 12 for variation across a subset of countries) may have been more subtly and systematically embedded in artistic expressions.

To investigate this, Figure 12 (b) presents regression estimates relating the predicted emotions conveyed in a painting, $\{p^e\}_{e \in \mathcal{E}}$, to the annual temperature (in degrees Celsius) during its year of production, available from 1500 onward for all European economies. A one-degree Celsius increase is associated with a 0.25 standard deviation decrease in the predicted probability of evoking fear and a 0.30 standard deviation decrease in sadness, alongside increases in positive emotions such as awe (0.31) and excitement (0.22). These magnitudes are economically meaningful: relatively modest climatic fluctuations are associated with sizable shifts in the emotional tone of artistic production.

Climatic variation may have had such strong well-being effects because much of the population at the time lived just above subsistence level. A one-degree Celsius increase is indeed associated with a 0.29 standard deviation increase in the wealth prediction, indicating significant improvements in depicted material living standards.[20]

[20]We also incorporate data on commodity prices across markets and over time (Appendix Figure D11)

Figure 12. Climatic variability and the emotions conveyed through paintings.



(a) Temperature (evol.) (b) Temperature (effect)

Notes: Panel **(a)** illustrates variation in climate, defined as the (residualized) 30-year moving average of predicted annual temperature in Celsius degrees in France, Germany, Italy, Spain, and the United Kingdom between 1500 and 2000 (Luterbacher et al., 2004). Panel **(b)** displays the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to climatic variability across all countries and periods. We consider Equation (2), with location, year, artist, and genre fixed effects, and replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with the 30-year moving average of predicted annual temperature in Celsius degrees (Luterbacher et al., 2004). The coefficients represent separate regressions with the following left-hand side variables: the predicted probability of expressing emotion e, p^e_j for painting j, where e is *fear* (dark-blue, plain circle), *sadness* (light-blue, plain square), *disgust* (teal, plain diamond), *anger* (black, plain triangle), *other* (gray, cross), *awe* (gold, hollow square), *contentment* (salmon, hollow triangle), *amusement* (pink, hollow diamond), and *excitement* (orange, hollow circle); the *positivity* index, $l_j = \sum_{e\in\mathcal{E}^+} p^e_j - \sum_{e\in\mathcal{E}^-} p^e_j$, in light blue; a measure of disagreement, defined as the sum of deviations from the mean emotion in a given context, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |p^e_j - \bar{p}^e_{\ell,t}|$; and the wealth prediction, $\widehat{w}_j$. All left-hand side variables are standardized, so the reported estimates can be interpreted in terms of standard deviations across country/year observations. The bands represent 10%, 5%, and 1% confidence intervals, and standard errors are clustered at the level of a country/half-century.

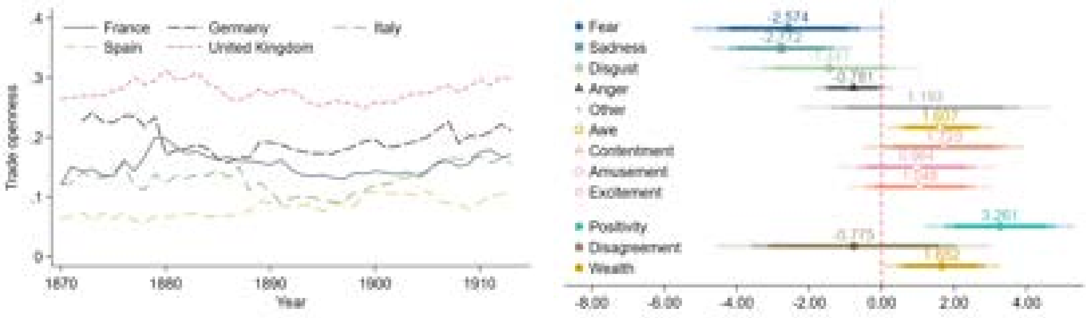
Taken together, these results point to a setting in which material and emotional welfare move closely together, reflecting the central role of subsistence constraints in shaping both economic outcomes and lived experience.

Economic integration and technological change We next turn to structural forces driving economic development. One key driver during our period of interest was the expansion of commerce and the opening of trade routes (Acemoglu et al., 2005b; Pascali, 2017). This shift spurred the rise of a merchant class, increased demand for luxury goods such as silk and spices, and encouraged the adoption of new crops (Nunn and Qian, 2011). However, reliable measures of trade prior to the nineteenth century are limited. To address this, we draw on two sources: (i) a measure of trade openness from the mid-nineteenth century onward (imports as a fraction of output, from Jordà et al., 2017), and (ii) the prices of luxury goods in commodity markets before 1850, shown in Appendix Figure D11 (Allen and Unger, 2019).

Figure 13 (b) shows that the first wave of globalization—along with access to cheaper

and examine the relationship between bread prices and conveyed emotions. A 50% increase in the price of bread raises the probability of evoking sadness (often linked to depictions of poverty) by 0.075 standard deviations, at the expense of awe and contentment. Both climatic variability and basic subsistence costs have a pronounced effect on our measure of emotional disagreement.

Figure 13. Trade and the first wave of globalization.



(a) Trade openness (evol.) (b) Trade openness (effect)

Notes: Panel **(a)** illustrates the variation in trade openness—defined as the ratio of imports to output—in France, Germany, Italy, Spain, and the United Kingdom (Jordà-Schularick-Taylor Macro-history Database, see Jordà et al., 2017). Panel **(b)** displays the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to this indicator of trade across all countries and periods. We consider Equation (2), with location, year, artist, and genre fixed effects, and replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with trade openness. All left-hand side variables are standardized.

food and raw materials—coincided with a markedly more positive emotional outlook in artwork across countries and time periods. A 20 percentage point increase in trade openness (as observed in England during the mid-nineteenth century) is associated with a 0.50–0.55 standard deviation decline in the probability of fear or sadness, alongside a substantial increase in the positivity index of 0.65 standard deviations. These magnitudes are economically meaningful, indicating that trade integration is associated with significant improvements in the emotional tone of artistic production.

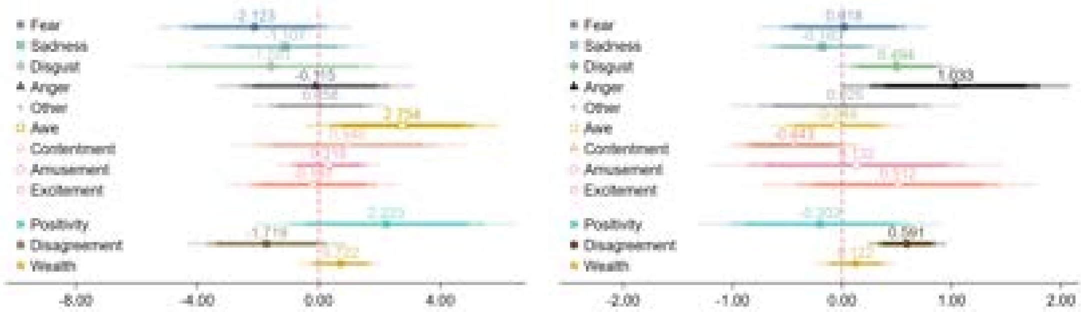
This positive emotional outlook might reflect a corresponding rise in living standards: a 20 percentage point increase in trade openness raises the predicted wealth index by 0.33 standard deviations, suggesting that improvements in material living standards are an important channel underlying these emotional changes.[21] Taken together, these results highlight a setting in which economic development improves both material and emotional welfare, consistent with the view that trade integration raised both living standards and the broader outlook of societies.

The first wave of globalization was partly driven by technological advances in maritime transportation (Pascali, 2017). How societies perceive technological progress is, in general, a fascinating question (see, e.g., Mokyr et al., 2015, on the "cultural anxiety" often associated with major innovations). Studying this question is challenging, not only due to limited data availability before 1800, but also because major technological shifts typically unfold gradually over time, whereas our empirical strategy relies on "shorter-

[21] Appendix Figure D11 (b) shows that higher prices for luxury goods—potentially reflecting increased demand from elites *before* the first wave of globalization—are associated with a less positive outlook: more frequent depictions of excitement, but also of anger.

horizon" variation within an artist's production cycle.

Figure 14. Technology adoption.



(a) Steamship (b) Radio

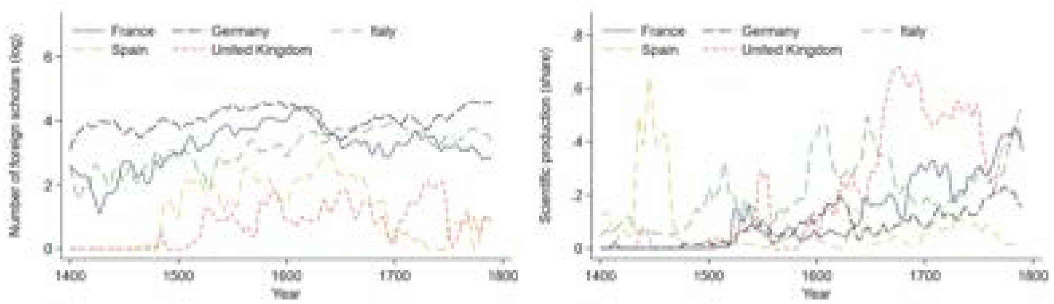
Notes: These figures display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to indicators of technology adoption. We consider Equation (2), with location, year, artist, and genre fixed effects, and replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with: **(a)** the annual ratio of tons of cargo transported using steamships over the population between 1788 and 1880; and **(b)** the ratio of radios per inhabitant between 1923 and 1950. Both variables are extracted from the Cross-country Historical Adoption of Technology (Comin and Hobijn, 2004). All left-hand side variables are standardized.

Figure 14 examines the rapid and uneven diffusion of two innovations from the nineteenth and early twentieth centuries: the steamship, which transformed trade and mobility, and the radio, which influenced entertainment, information, and, in some cases, propaganda (Adena et al., 2015; Castiello, 2025).

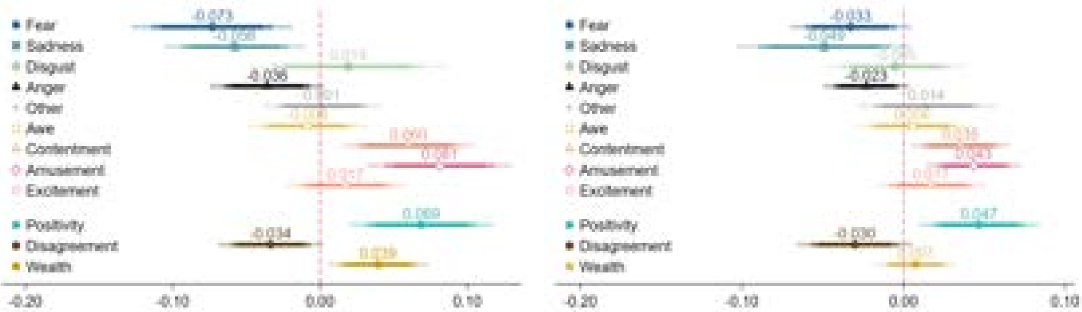
The emotional response to these technologies markedly differs. Steamship adoption is associated with more positive emotional expression: awe increases by 0.083 standard deviations following an increase in steamship cargo equivalent to the rise observed in the United States during the nineteenth century (approximately 0.03 tons per capita per year), and disagreement declines. These patterns are consistent with the expansion of economic opportunities and improved integration into global markets. By contrast, the diffusion of radio corresponds to more polarized emotional responses. The 15 percentage point increase in radio ownership observed within the French or German populations between 1900 and 1950 would be associated with a 0.15 standard deviation increase in depicted anger, alongside increases in other intense emotions, and higher emotional disagreement. Notably, these changes occur without measurable effects on material representations. This contrast suggests that technological change can affect how societies experience their environment even when it does not substantially alter material living standards, highlighting a divergence between material and emotional welfare.

Institutional and political environments The previous cases largely concern transformations that affect economic conditions directly. We now turn to changes in the social and political environment, where the main effects may operate through experienced welfare rather than material living standards. Recent efforts have helped to document the evolution of academic production from 1200 onward (De La Croix et al., 2024, see Appendix B.3)—a broad expansion with important implications for *longer-run* economic development (Cantoni and Yuchtman, 2014; Squicciarini and Voigtländer, 2015). We draw on this dataset of scholars and their host institutions to characterize knowledge production and the rigidity of the social order within a given context, focusing on a country's global academic influence, its openness to foreign scholars, the role of religiosity in its academic output (Cantoni et al., 2018; Blasutto and De La Croix, 2023; Cummins and Noble, 2025), and the emergence of scientific research.

Figure 15. Universities and the scientific revolution.



(a) Foreign scholars (evol.) (b) Sciences (evol.)



(c) Foreign scholars (effect) (d) Sciences (effect)

Notes: The top panels illustrate the variation in the intensity and nature of knowledge production in France, Germany, Italy, Spain, and the United Kingdom between 1400 and 1800 (De La Croix et al., 2024). Panel **(a)** reports a measure of academic openness, defined by the (log) number of foreign scholars; and panel **(b)** reports the weighted share of knowledge production in sciences. The bottom panels display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e \in \mathcal{E}}$, to these indicators of knowledge production. We consider Equation (2), with location, year, artist, and genre fixed effects. All left-hand side variables are standardized.

Figure 15 (panels a and b)—openness to foreign scholars and scientific research—and Appendix Figure D12 (panels a and b) illustrate the substantial variation in the extent and nature of academic production within the borders of five present-day countries between 1400 and 1800. For example, Great Britain experiences a pronounced rise in scientific production during the seventeenth century, which persists through 1800 and predates the Industrial Revolution, while Spain and Italy undergo episodes of scientific expansion followed by counter-reformations.

Figure 15 (panel c) shows that the overall quality and openness of academic production is associated with sizable shifts in emotional expression. The increase in foreign scholars observed in France between 1500 and 1600 (a rise of 2 in the openness measure) is associated with declines in fear (0.15 standard deviations), sadness (0.12), and anger (0.07), alongside increases in amusement (0.16) and contentment (0.12). In parallel, contexts dominated by scientific production—typically at the expense of religiosity—exhibit less frequent depictions of sadness and more amusement (panel d), despite limited *contemporaneous* effects on material representations. These magnitudes suggest that intellectual transformations shape emotional expression in meaningful ways, even when material conditions remain broadly unchanged in the short run.

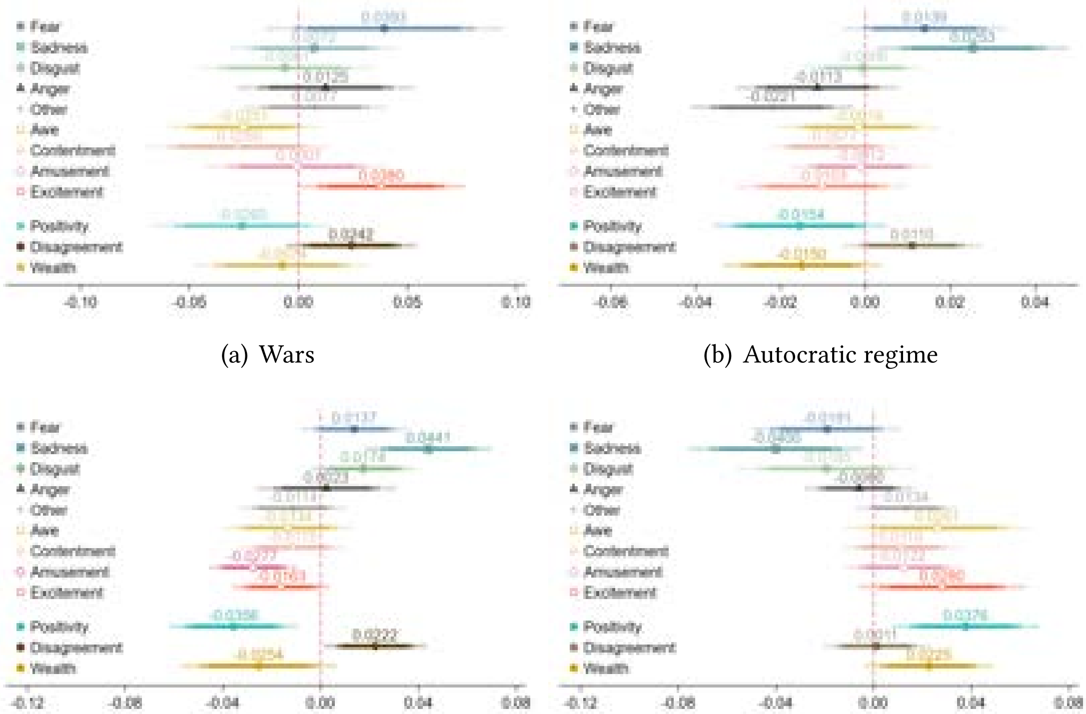
Building on our claim that emotions offer insight into how societies perceived their historical moment, these results help illuminate the broader transformation associated with the Enlightenment—typically less directly observable with conventional indicators of economic activity. Prior to this period, hierarchical structures and religious doctrine reinforced social order and constrained intellectual inquiry. The emergence of scientific production and openness to new ideas coincides with a shift toward more positive emotional expression, reflecting a climate of optimism, intellectual curiosity, and forward-looking thought. This pattern resonates with Mokyr's concept of the Industrial Enlightenment (Mokyr, 2005), emphasizing that important socioeconomic transformation may not only be material.

We next turn to the political environment. We estimate specification (2) using several short- to medium-term sources of variation in the (geo)political context, including participation in major wars, measures of institutionalized autocracy, regime durability, and the nature of political transitions.

Figure 16 shows that political instability significantly affects the emotional tone of artworks. Wars increase the depiction of fear and reduce contentment, while also raising emotional disagreement, reflecting heterogeneous responses to conflict (panel a). In parallel, a one standard deviation increase in the autocracy score is associated with modest increases in fear, sadness, and disagreement (0.01–0.025 standard deviations, panel b).

Regime durability exhibits a stronger effect (panel c): a one standard deviation increase is associated with a decline in positive emotions (e.g., amusement decreases by 0.028 standard deviations), an increase in negative emotions (e.g., sadness increases by 0.044), and higher emotional disagreement. In periods of regime transition (panel d), positive democratic shifts are associated with increases in awe and excitement, while

Figure 16. Regime characteristics, transitions, wars, and the emotions conveyed through paintings.



(a) Wars (b) Autocratic regime

(c) Regime durability (d) Regime transitions

Notes: These figures display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to indicators of political turbulence and conflicts. We consider Equation (2), with location, year, artist, and genre fixed effects. All left-hand side variables are standardized.

negative political shifts tend to elicit sadness.[22] The response of emotional expression to conflict, governance, and political transitions indicates that the political environment shapes well-being over and above material conditions.

A summary Taken together, these results reveal a consistent pattern. Changes in trade or climate—expected to affect material living standards—are associated with sizable shifts in both emotional expression and material representations. By contrast, forces such as social rigidity or political instability weakly affect material representations, even when emotional signals respond strongly. This asymmetry highlights a central distinction: material indicators capture changes in living standards, while emotional signals capture how these conditions were experienced. More importantly, the latter also respond to institutional and political environments in ways that are not reflected in material conditions alone.

[22]In Appendix D.3, we complement this evidence with (i) a specific focus on labor and emancipation reforms as well as transitions to democratic regimes, and (ii) an event study approach that further examines the dynamics of conveyed emotions.

5 Concluding remarks

This paper shows how artworks can be used to measure both material living standards and the emotional experience of societies. In particular, we use paintings to construct a measure of experienced welfare—a dimension largely absent from conventional historical data. While traditional indicators document economic performance, emotional signals extracted from artistic production reveal how societies perceived and lived through major economic, institutional, and political transformations.

Our empirical analysis shows that material and emotional signals behave differently. Material representations and extracted emotions closely track changes in living standards associated with economic development, trade integration, and subsistence shocks. However, emotional signals also respond to economic uncertainty, institutional change, and political instability—often without corresponding changes in material conditions. The magnitude of these effects indicates substantial variation in experienced well-being over time and across contexts that is not captured by conventional measures of economic performance. In this sense, emotional signals embedded in artistic production provide a systematic proxy for experienced welfare, capturing dimensions of societal change not reflected in material living standards alone.

This research will benefit from ongoing digitization efforts as museums continue to expand access to their collections. Future work could extend coverage to earlier periods, non-European regions, and less prominent artists. In particular, (i) refining the spatial attribution of paintings to precise locations would allow the identification of finer within-country patterns; and (ii) accounting for heterogeneity in artists' living standards within locations—by socioeconomic status, gender, religion, or ethnicity—would deepen our understanding of the geography and dynamics of inequality.

References

Abowd, John M, Francis Kramarz, and David N Margolis, "High wage workers and high wage firms," *Econometrica*, 1999, *67* (2), 251–333.

Acemoglu, Daron, Simon Johnson, and James A Robinson, "Institutions as a fundamental cause of long-run growth," *Handbook of Economic Growth*, 2005, *1*, 385–472.

_ , _ , and James Robinson, "The rise of Europe: Atlantic trade, institutional change, and economic growth," *American Economic Review*, 2005, *95* (3), 546–579.

Achlioptas, Panos, Maks Ovsjanikov, Kilichbek Haydarov, Mohamed Elhoseiny, and Leonidas Guibas, "ArtEmis: Affective language for visual art," *CoRR*, 2021, *abs/2101.07396.*

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya, "Radio and the Rise of the Nazis in Prewar Germany," *The Quarterly Journal of Economics*, 2015, *130* (4), 1885–1939.

Albrecht, Milton C, "The relationship of literature and society," *American Journal of Sociology*, 1954, *59* (5), 425–436.

Alexander, Victoria D, *Sociology of the arts: Exploring fine and popular forms*, John Wiley & Sons, 2020.

Algan, Yann, Eva Davoine, Thomas Renault, and Stefanie Stantcheva, "Emotions and policy views," Technical Report, Harvard University Working Paper 2025.

Allen, Robert C and Richard W Unger, "The Allen-Unger Global Commodity Prices Database: Social and Economic History," *Research Data Journal for the Humanities and Social Sciences*, 2019, *4* (1), 81–90.

Alvaredo, Facundo, Anthony B Atkinson, Thomas Blanchet, Lucas Chancel, Luis Bauluz, Matthew Fisher-Post, Ignacio Flores, Bertrand Garbinti, Jonathan Goupille-Lebret, Clara Martínez-Toledano et al., "Distributional national accounts guidelines, methods and concepts used in the world inequality database," Technical Report 2020.

Ashenfelter, Orley and Kathryn Graddy, "Art auctions," *Handbook of the Economics of Art and Culture*, 2006, *1*, 909–945.

Bairoch, Paul, *Cities and economic development: from the dawn of history to the present*, University of Chicago Press, 1988.

Barré, Jean, Jean-Baptiste Camps, and Thierry Poibeau, "Operationalizing Canonicity," *Journal of Cultural Analytics*, 2023, *8* (1), 88113.

Bengio, Yoshua, Aaron Courville, and Pascal Vincent, "Representation learning: A review and new perspectives," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, *35* (8), 1798–1828.

Benjamin, Daniel J., Kristen Cooper, Ori Heffetz, and Miles Kimball, "From Happiness Data to Economic Conclusions," *Annual Review of Economics*, August 2024, *16* (1), 359–391.

Blasutto, Fabio and David De La Croix, "Catholic censorship and the demise of knowledge production in early modern Italy," *The Economic Journal*, 2023, *133* (656), 2899–2924.

Bloom, Nicholas, "Fluctuations in uncertainty," *Journal of Economic Perspectives*, 2014, *28* (2), 153–76.

Boerner, Lars, Tim Reinicke, Samad Sarferaz, and Battista Severgnini, "Colors of Growth," *KOF Working Papers*, 2025, *527.*

Bolt, Jutta and Jan Luiten Van Zanden, "Maddison-style estimates of the evolution of the world economy: A new 2023 update," *Journal of Economic Surveys*, 2025, *39* (2), 631–671.

Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler, "How Much Should we Trust Estimates of Firm Effects and Worker Sorting?," Technical Report 2020.

Borowiecki, Karol Jan, "How Are You, My Dearest Mozart? Well-Being and Creativity of Three Famous Composers Based on Their Letters," *The Review of Economics and Statistics*, 2017, *99* (4), 591–605.

— , "Good Reverberations? Teacher Influence in Music Composition since 1450," *Journal of Political Economy*, 2022, *130* (4), 991–1090.

Bouscasse, Paul, Emi Nakamura, and Jón Steinsson, "When did growth begin? New estimates of productivity growth in England from 1250 to 1870," *The Quarterly Journal of Economics*, 2024, p. qjae046.

Brown, Steven, *The Unification of the Arts*, Oxford University Press, 11 2021.

Cantoni, Davide and Noam Yuchtman, "Medieval universities, legal institutions, and the commercial revolution," *The Quarterly Journal of Economics*, 2014, *129* (2), 823–887.

— **, Jeremiah Dittmar, and Noam Yuchtman**, "Religious competition and reallocation: The political economy of secularization in the protestant reformation," *The Quarterly Journal of Economics*, 2018, *133* (4), 2037–2096.

Castiello, Maxence, "Spread the Word: Mass Media, Language and Propaganda in Fascist Italy," 2025. Working paper.

Chen, Xi and William D Nordhaus, "Using luminosity data as a proxy for economic statistics," *Proceedings of the National Academy of Sciences*, 2011, *108* (21), 8589–8594.

Clark, Gregory, "The condition of the working class in England, 1209–2004," *Journal of Political Economy*, 2005, *113* (6), 1307–1340.

— , "The long march of history: Farm wages, population, and economic growth, England 1209–1869 1," *The Economic History Review*, 2007, *60* (1), 97–135.

Cohen, Margaret, *The Sentimental Education of the Novel*, Princeton University Press, 1999.

Comin, Diego and Bart Hobijn, "Cross-country technology adoption: making the theories face the facts," *Journal of Monetary Economics*, 2004, *51* (1), 39–83.

Compiani, Giovanni, Ilya Morozov, and Stephan Seiler, "Demand estimation with text and image data," *arXiv preprint arXiv:2503.20711*, 2025.

Croix, David De La, Frédéric Docquier, Alice Fabre, and Robert Stelter, "The academic market and the rise of universities in medieval and early modern Europe (1000–1800)," *Journal of the European Economic Association*, 2024, *22* (4), 1541–1589.

Cui, Yin, Menglin Jia, Tsung-Yi Lin, Yang Song, and Serge Belongie, "Class-balanced loss based on effective number of samples," in "Proceedings of the IEEE/CVF conference on computer vision and pattern recognition" 2019, pp. 9268–9277.

Cummins, Neil J. and Aurelius Noble, "Real and Nominal Belief in God: Evidence from Wills, England: 1300-1850," Technical Report 2025.

Damasio, Antonio R, *Descartes' error*, Random House, 2006.

Dell, Melissa, "Deep learning for economists," *Journal of Economic Literature*, 2025, *63* (1), 5–58.

DellaVigna, Stefano, "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, June 2009, *47* (2), 315–72.

Deng, Jia, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei, "Imagenet: A large-scale hierarchical image database," *IEEE conference on computer vision and pattern recognition*, 2009, pp. 248–255.

Dissanayake, Ellen, *Homo Aestheticus: Where Art Comes from and why*, Free Press, 1992.

Donaldson, Dave and Adam Storeygard, "The view from above: Applications of satellite data in economics," *Journal of Economic Perspectives*, 2016, *30* (4), 171–98.

Dutton, Denis, *The art instinct: Beauty, pleasure, and human evolution*, Bloomsbury Press, 2009.

Ekman, Paul, "Basic emotions," in T. Dalgleish and M. J. Power, eds., *Handbook of cognition and emotion*, John Wiley & Sons Ltd., 1999, pp. 45–60.

Elster, Jon, "Emotions and Economic Theory," *Journal of Economic Literature*, 1998, *36* (1), 47–74.

Etro, Federico, "The economics of Renaissance art," *The Journal of Economic History*, 2018, *78* (2), 500–538.

_ **and Elena Stepanova**, "Art return rates from old master paintings to contemporary art," *Journal of Economic Behavior & Organization*, 2021, *181*, 94–116.

Frijda, Nico H., *The laws of emotion*, Mahwah, NJ: Lawrence Erlbaum, 2007.

Fujita, Masahisa, Paul R Krugman, and Anthony Venables, *The spatial economy: Cities, regions, and international trade*, MIT Press, 2001.

Gal, Yarin and Zoubin Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in "international conference on machine learning" PMLR 2016, pp. 1050–1059.

Galenson, David W and Bruce A Weinberg, "Age and the quality of work: The case of modern American painters," *Journal of Political Economy*, 2000, *108* (4), 761–777.

__ **and** __ , "Creating modern art: The changing careers of painters in France from impressionism to cubism," *American Economic Review*, 2001, *91* (4), 1063–1071.

Galor, Oded, "From stagnation to growth: unified growth theory," *Handbook of Economic Growth*, 2005, *1*, 171–293.

__ , *Unified growth theory*, Princeton University Press, 2011.

Gebru, Timnit, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Erez Lieberman Aiden, and Li Fei-Fei, "Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States," *Proceedings of the National Academy of Sciences*, 2017, *114* (50), 13108–13113.

Gell, Alfred, *Art and Agency: An Anthropological Theory* Art and Agency: An Anthropological Theory, Clarendon Press, 1998.

Ginsburgh, Victor and Philippe Jeanfils, "Long-term comovements in international markets for paintings," *European Economic Review*, 1995, *39* (3-4), 538–548.

Giorcelli, Michela and Petra Moser, "Copyrights and creativity: Evidence from Italian opera in the napoleonic age," *Journal of Political Economy*, 2020, *128* (11), 4163–4210.

Graddy, Kathryn, "Taste endures! The rankings of Roger de Piles (1709) and three centuries of art prices," *The Journal of Economic History*, 2013, *73* (3), 766–791.

Han, Sukjin and Kyungho Lee, "Copyright and Competition: Estimating Supply and Demand with Unstructured Data," *arXiv preprint arXiv:2501.16120*, 2025.

Helliwell, John F, Richard Layard, Jeffrey D Sachs, Jan-Emmanuel De Neve, Lara B Aknin, and Shun Wang, "World happiness report 2025," 2025.

Hellmanzik, Christiane, "Location matters: Estimating cluster premiums for prominent modern artists," *European Economic Review*, 2010, *54* (2), 199–218.

Henderson, J Vernon, Adam Storeygard, and David N Weil, "Measuring economic growth from outer space," *American Economic Review*, 2012, *102* (2), 994–1028.

Henrich, Joseph, Robert Boyd, Maxime Derex, Michelle A. Kline, Alex Mesoudi, Michael Muthukrishna, Adam T. Powell, Stephen J. Shennan, and Mark G. Thomas, "Understanding cumulative cultural evolution," *Proceedings of the National Academy of Sciences*, 11 2016, *113.*

Hills, Thomas T., Eugenio Proto, Daniel Sgroi, and Chanuki Illushka Seresinhe, "Historical analysis of national subjective wellbeing using millions of digitized books," *Nature Human Behaviour*, 2019, *3* (12), 1271–1275.

Hulten, Charles R. and Leonard I. Nakamura, "Is GDP Becoming Obsolete? The "Beyond GDP" Debate," *NBER Working Paper 30196*, 2022.

Ioffe, Sergey and Christian Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *CoRR*, 2015.

Jacks, David S, Kevin H O'Rourke, and Jeffrey G Williamson, "Commodity price volatility and world market integration since 1700," *Review of Economics and Statistics*, 2011, *93* (3), 800–813.

Jaskot, Paul B., "Digital Art History as the Social History of Art: Towards the Disciplinary Relevance of Digital Methods," *Visual Resources*, 4 2019, *35*, 21–33.

Jean, Neal, Marshall Burke, Michael Xie, W Matthew Davis, David B Lobell, and Stefano Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, 2016, *353* (6301), 790–794.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor, "Macrofinancial history and the new business cycle facts," *NBER macroeconomics annual*, 2017, *31* (1), 213–263.

Jurado, Kyle, Sydney C Ludvigson, and Serena Ng, "Measuring uncertainty," *American Economic Review*, 2015, *105* (3), 1177–1216.

Kingma, Diederik P. and Jimmy Lei Ba, "Adam: A method for stochastic optimization," in "International Conference on Learning Representations" 2015.

Logothetis, Vassilis, Stelios Michalopoulos, and Elias Papaioannou, "Songs of Refugees," Technical Report 2025.

Loshchilov, Ilya and Frank Hutter, "Decoupled Weight Decay Regularization," in "International Conference on Learning Representations" 2017.

Luterbacher, Jurg, Daniel Dietrich, Elena Xoplaki, Martin Grosjean, and Heinz Wanner, "European seasonal and annual temperature variability, trends, and extremes since 1500," *Science*, 2004, *303* (5663), 1499–1503.

Maddison, Angus, *Contours of the world economy 1-2030 AD: Essays in macro-economic history*, OUP Oxford, 2007.

Mao, Hui, Ming Cheung, and James She, "Deepart: Learning joint representations of visual arts," in "Proceedings of the 25th ACM international conference on Multimedia" ACM 2017, pp. 1183–1191.

Marshall, Monty G, Ted Robert Gurr, and Keith Jaggers, "Polity IV project: Political regime characteristics and transitions, 1800–2013," *Center for Systemic Peace*, 2014, *5.*

Meier, Armando N., "Emotions and Risk Attitudes," *American Economic Journal: Applied Economics*, July 2022, *14* (3), 527–58.

Michalopoulos, Stelios, "Ethnographic Records, Folklore, and AI," *National Bureau of Economic Research*, 2025.

— **and Melanie Meng Xue**, "Folklore," *The Quarterly Journal of Economics*, 2021, *136* (4), 1993–2046.

Mohamed, Youssef, Mohamed Abdelfattah, Shyma Y. Alhuwaider, Feifan Li, Xiangliang Zhang, Kenneth Ward Church, and Mohamed Elhoseiny, "ArtELingo: A Million Emotion Annotations of WikiArt with Emphasis on Diversity over Language and Culture," Technical Report 2022.

Mokyr, Joel, "The intellectual origins of modern economic growth," *The Journal of Economic History*, 2005, *65* (2), 285–351.

_ , Chris Vickers, and Nicolas L Ziebarth, "The history of technological anxiety and the future of economic growth: Is this time different?," *Journal of Economic Perspectives*, 2015, *29* (3), 31–50.

Moretti, Franco, "The slaughterhouse of literature," *MLQ: Modern Language Quarterly*, 2000, *61* (1), 207–227.

Naik, Nikhil, Scott Duke Kominers, Ramesh Raskar, Edward L Glaeser, and César A Hidalgo, "Computer vision uncovers predictors of physical urban change," *Proceedings of the National Academy of Sciences*, 2017, *114* (29), 7571–7576.

North, Douglass C, "Understanding the process of economic change," in "Understanding the Process of Economic Change," Princeton University Press, 2010.

Nunn, Nathan and Nancy Qian, "The potato's contribution to population and urbanization: evidence from a historical experiment," *The Quarterly Journal of Economics*, 2011, *126* (2), 593–650.

Oster, Emily, "Witchcraft, weather and economic growth in Renaissance Europe," *Journal of Economic Perspectives*, 2004, *18* (1), 215–228.

Overman, Henry, Diego Puga, Matthew Turner, and Marcy Burchfield, "Sprawl: A Portrait from Space," *The Quarterly Journal of Economics*, 02 2006, *121*, 587–633.

Pascali, Luigi, "The wind of change: Maritime technology, trade, and economic development," *American Economic Review*, 2017, *107* (9), 2821–2854.

Renneboog, Luc and Christophe Spaenjers, "Buying beauty: On prices and returns in the art market," *Management Science*, 2013, *59* (1), 36–53.

Robinson, Jenefer, *Deeper than reason: Emotion and its role in literature, music, and art*, Oxford University Press, 2005.

Safra, Lou, Coralie Chevallier, Julie Grèzes, and Nicolas Baumard, "Tracking historical changes in perceived trustworthiness in Western Europe using machine learning analyses of facial cues in paintings," *Nature Communications*, 2020, *11* (1), 1–7.

Saleh, Babak and Ahmed Elgammal, "Large-scale Classification of Fine-Art Paintings: Learning The Right Metric on The Right Feature," *International Journal for Digital Art History*, 2016.

Sanh, Victor, Lysandre Debut, Julien Chaumond, and Thomas Wolf, "DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter," *arXiv preprint arXiv:1910.01108*, 2019.

Schularick, Moritz and Alan M Taylor, "Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008," *American Economic Review*, 2012, *102* (2), 1029–61.

Scitovsky, Tibor, "What's wrong with the arts is what's wrong with society," *The American Economic Review*, 1972, *62* (1/2), 62–69.

Spaenjers, Christophe, William N Goetzmann, and Elena Mamonova, "The economics of aesthetics and record prices for art since 1701," *Explorations in Economic History*, 2015, *57*, 79–94.

Squicciarini, Mara P and Nico Voigtländer, "Human capital and industrialization: Evidence from the age of enlightenment," *The Quarterly Journal of Economics*, 2015, *130* (4), 1825–1883.

Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich, "Going deeper with convolutions," in "Proceedings of the IEEE conference on computer vision and pattern recognition" 2015, pp. 1–9.

Tan, Mingxing and Quoc V. Le, "EfficientNetV2: Smaller Models and Faster Training," *Proceedings of the 38th International Conference on Machine Learning*, 2021, *139*, 10096–10106.

Throsby, David, "The production and consumption of the arts: A view of cultural economics," *Journal of Economic Literature*, 1994, *32* (1), 1–29.

Tolstoy, Leo, *What is Art?*, New York: Brotherhood Publishing, 1898.

Tompson, Jonathan, Ross Goroshin, Arjun Jain, Yann LeCun, and Christopher Bregler, "Efficient object localization using convolutional networks," *Computer Vision and Pattern Recognition*, 2014.

Tooby, John and Leda Cosmides, "The past explains the present," *Ethology and Sociobiology*, 7 1990, *11* (4-5), 375–424.

Ursprung, Heinrich W and Christian Wiermann, "Reputation, price, and death: An empirical analysis of art price formation," *Economic Inquiry*, 2011, *49* (3), 697–715.

Voth, Hans-Joachim and David Yanagizawa-Drott, "Image(s)," July 2024. Working paper.

Wälde, Klaus and Agnes Moors, "Current Emotion Research in Economics," *Emotion Review*, 2017, *9* (3), 271–278.

Waldinger, Maria, "The economic effects of long-term climate change: evidence from the Little Ice Age," *Journal of Political Economy*, 2022, *130* (9), 2275–2314.

Whitaker, Amy and Roman Kräussl, "Fractional equity, blockchain, and the future of creative work," *Management Science*, 2020, *66* (10), 4594–4611.

Zeiler, Matthew D. and Rob Fergus, "Visualizing and understanding convolutional networks," in "Computer Vision – ECCV 2014" 2014, pp. 818–833.

Zhou, Bolei, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba, "Learning deep features for discriminative localization," in "Proceedings of the IEEE conference on computer vision and pattern recognition" 2016, pp. 2921–2929.

Zijdeman, Richard and Filipa Ribeira da Silva, "Life Expectancy at Birth (Total)," 2015.

Online Appendix for "State of the Art"—not for publication

This appendix **(A)** complements introductory remarks in the main paper, **(B)** provides details about the data sources, **(C)** sheds additional light on the encoding of paintings through computer vision, and **(D)** presents additional empirical results.

A Motivation **2**

B Data description **6**

B.1 A harmonized collection of paintings 6

B.2 A collection of *annotated* images 11

B.3 Other data sources 13

C The encoding of paintings through computer vision **19**

C.1 A sentiment classification pipeline 19

C.2 The identification of material representation(s) 27

C.3 Attrition, commercialization, and selection into conservation 31

C.4 The mapping between well-being and conveyed emotions 34

C.5 Material representations as a measure of living standards 38

D Emotional and material welfare through socioeconomic transformation **44**

D.1 The dynamics of extracted signals across countries 45

D.2 Empirical strategy 50

D.3 The effect of large socio-economic transformations 57

A Motivation

This section complements our introductory remarks by: (i) displaying artworks referenced in the motivational discussion (see, e.g., p.2–3); and illustrating the dynamics of emotions in French paintings.

Figure A1. Paintings as testaments to their times and contexts—additional illustrations.

(a) Guernica

(b) Liberty Leading the People

(c) The Garden of Earthly Delights

Notes: Panel **(a)** displays *Guernica* (Pablo Picasso, 1937). Panel **(b)** shows *Liberty Leading the People* (Eugène Delacroix, 1830), which captures the short-lived July Revolution of 1830 in France. Panel **(c)** shows *The Garden of Earthly Delights* (Hieronymus Bosch, around 1500), which depicts novel fruits and imaginary creatures at a time of exploration and discovery of the New World.

Additional illustrations In the introduction, we refer to Pablo Picasso's *Guernica*, Delacroix's *Liberty Leading the People*, and Hieronymus Bosch's *The Garden of Earthly Delights*. These paintings are reproduced in Appendix Figure A1. In particular, *The Garden of Earthly Delights* is a prime example of (i) symbolic or indirect messaging in visual art and (ii) the emotional duality provoked by uncertain times, simultaneously evoking excitement and anxiety. The left panel depicts *The Garden of Eden*, the center shows *The Garden of Earthly Delights*, and the right panel presents *Hell*. While the triptych has been extensively studied, its meaning remains open to interpretation. Nonetheless, the imagery—featuring "exotic" fruits, animals, and characters—appears influenced by contemporaneous global explorations (e.g., the visit to Egypt by Cyriac of Ancona) and the discovery of the New World. These themes also resonated with contemporaries such as Albrecht Dürer, who shared a fascination with these discoveries.

Figure A2. Paintings as a testament to their times and contexts—different perspectives.

(a) The Shepherdess (b) The Stone Breakers

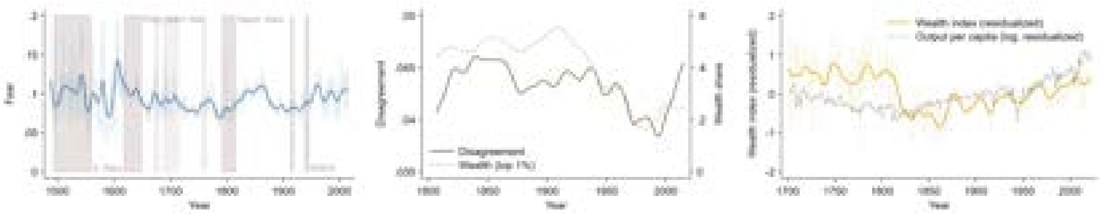
(c) Napoleon Crossing the Alps (d) Snow Storm: Hannibal and his Army Crossing the Alps

Notes: Panel **(a)** shows *The Shepherdess* (1889, William Bouguereau), an idealized vision of pastoral life. Panel **(b)** displays *The Stone Breakers* (1849, Gustave Courbet), a painting typical of the realism movement. Interestingly, Bouguereau was a very successful painter in his lifetime, whereas Gustave Courbet gained national prominence later in life and worldwide recognition after his death. Panel **(c)** shows one of the five versions of *Napoleon Crossing the Alps* (1801–1805, Jacques-Louis David), an allegorical portrait of Bonaparte leading his army (France) through the Col du Grand Saint-Bernard—an image of the early challenges of the French Consulate, seen from a French perspective. Panel **(d)** presents a response by William Turner, showing the struggle of Hannibal in his earlier crossing (as a possible metaphor for Napoleon's later difficulties), from an English perspective.

In the introduction, we also highlight variation in emotional expression across contemporaneous artworks, even when addressing similar themes. For instance, William

Bouguereau, born into a Catholic merchant family, portrays an idealized vision of pastoral life in *The Shepherdess* (panel a of Figure A2) and related works, in which peasants often resemble goddesses or nymphs. By contrast, Gustave Courbet—raised in an anti-monarchical family in the Doubs—depicts rural life with what was, at the time, considered radical realism. *The Stone Breakers* (panel b) captures what he described as "the most complete expression of poverty." A further contrast appears in depictions of *The Crossing of the Alps*—a symbol of Napoleonic ambition. Jacques-Louis David and William Turner both painted this event, but from markedly different perspectives (Figure A2, panels c and d). David's portrayal casts Napoleon as an allegorical figure, imperiously leading a nation on horseback—a distinctly French perspective from the early Consulate years. In contrast, Turner's version depicts Hannibal's crossing of the Alps, with the army engulfed by a snowstorm—an English view painted during the waning years of Napoleon's reign.

Figure A3. The dynamics of fear, disagreement, and wealth in French paintings over time.



(a) Fear (b) Disagreement (c) Wealth index

Notes: Panel **(a)** illustrates the evolution of the average emotion probability allocated to *fear*, $p^f_{\ell,t}$, for paintings produced at time t between 1491 and 2015 in France. The starting date is set to exclude the first percentile of French artworks in terms of production time, and each observation within a year is weighted to account for duplicates and the overall production of each artist. For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a coefficient of 100—as a thicker line. Shaded areas highlight the periods of major wars (e.g., 1792–1815 for the Revolutionary and Napoleonic Wars). Panel **(b)** displays the disagreement in portrayed emotions between 1807 and 2015, alongside the interpolated/smoothed share of wealth owned by the top 1%, extracted from the World Inequality Database (Alvaredo et al., 2020). The context-level *disagreement* index is computed as the average level of disagreement across all paintings produced in a given year t and location ℓ, $d_{\ell,t} = \frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |r^e_j - \bar{r}^e_{\ell,t}|$, where both r^e_j and $\bar{r}^e_{\ell,t}$ are derived from the earlier measures, p^e_j and $\bar{p}^e_{\ell,t}$, but residualized by genre, movement, age, and artist fixed effects. Panel **(c)** aggregates the wealth index at the yearly level to show the dynamics of such index over time in France.

The dynamics of fear, disagreement, and wealth in French paintings In other parts, we provide a systematic visualization of how emotions in paintings vary across countries and over time (see Section 4.1 and Appendix D.1). Here, we briefly illustrate some of these meaningful fluctuations using the case of France (Appendix Figure A3). Panel (a) shows that depictions of fear increase by approximately 3 percentage points during periods of historical turbulence, such as the Religious Wars, the end of the Napoleonic Wars, and World War II. Panel (b) indicates that within-context emotional variance tracks long-run trends in inequality since 1807, suggesting a relationship between social inequality and emotional diversity in artistic expression within a given environment. Panel (c) displays the dynamics of material living standards, as extracted from artwork—contrasting it against more traditional measurement of output per capita (Bolt and Van Zanden, 2025).

B Data description

This section provides complements to Section 2.

B.1 A harmonized collection of paintings

Google Arts and Culture *Google Arts and Culture* is a partnership between Google and numerous cultural institutions (museums, galleries, etc.) aimed at digitizing and publishing art collections in high resolution for online access. Many major art museums collaborate with Google, resulting in the availability of a large share of publicly accessible, well-known paintings as online images on the platform. For example, the Musée d'Orsay (Paris) shares hundreds of items, including *The Church in Auvers-sur-Oise* by Vincent van Gogh and *The Ballet Class* by Edgar Degas. The Museum of Modern Art also shares its collection, including *Hope II* by Gustav Klimt and *Turning Road at Montgeroult* by Paul Cézanne. Overall, more than 9,000 artists are represented on *Google Arts and Culture*, with approximately 360,000 art pieces, about two-thirds of which are paintings. Prominent artists—including Albrecht Dürer, Pierre-Auguste Renoir, Rembrandt, Vincent Van Gogh, Edgar Degas, Henri de Toulouse-Lautrec, and William Turner—typically have between 500 and 2,000 paintings featured. While the collections primarily consist of European works, they also include Chinese calligraphy, Japanese prints and water-based works, so-called "primitive" art, landscapes from the Hudson River School, and American realism.

Artists and paintings are accompanied by descriptions and a range of metadata tags. Figure B1 provides an example of this information and illustrates how it is stored across the different, interlinked pages associated with an artist or a specific artwork. Data collection relies on a headless browser, followed by post-processing and cleaning procedures. The main challenge is the correct identification of the location and date of production. First, we retain the completion year of the artwork or, when a time interval is reported (e.g., 1900–1906), we use the final year provided by *Google Arts and Culture*. Second, we use entity-recognition models and large language models to clean location information and assign a (geo)location. Location data are often noisy—partly because they are frequently unobserved by museum curators, art collectors, or auction houses—unless explicitly tied to a specific place, as is often the case for landscapes. When the exact location is unavailable, we retain the country of production. If the country is also missing, we reconstruct it using artist biographies.

Figure B2 shows the locations of a subset of precisely geolocated paintings across Northern Europe and Japan, where only about one-third of artworks can be assigned to an exact location. The figure also illustrates that (i) the geography of artwork produc-

Figure B1. Artist biographies and artworks in *Google Arts and Culture*.

(a) Artist information

(b) List of artworks

(c) Artwork image

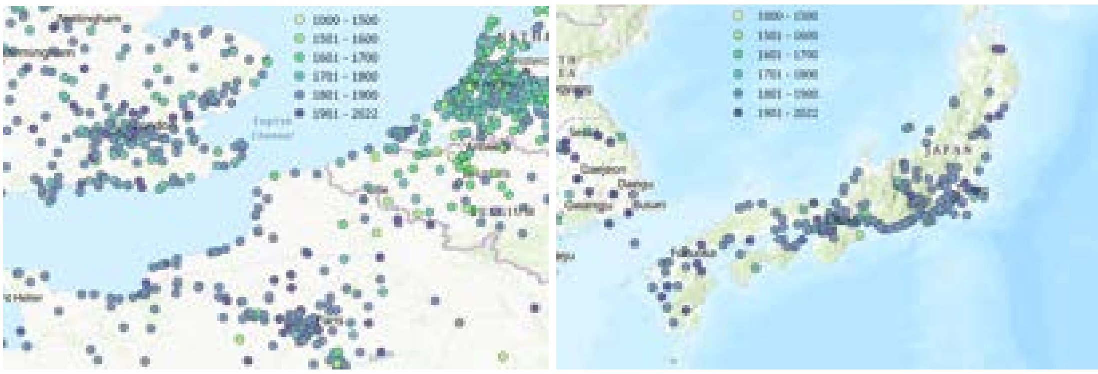
(d) Artwork attributes and tags

Notes: This figure shows an example of the information available in *Google Arts and Culture* for an artist (here, Vincent van Gogh). Panel **(a)** is the biographical page; panel **(b)** shows the list of artworks; panel **(c)** displays the image of *The Harvest*; and panel **(d)** shows the associated attributes and tags. The artist information includes: name, birth and death dates, description, list of artworks, and Wikidata ID for linking with Wikipedia information. The painting information includes: the image, title, date, location of creation, and medium.

tion closely aligns with the geography of economic activity, and (ii) the golden ages of painting (e.g., the Dutch Golden Age or the French Belle Époque) are clearly visible in the preserved artworks—with the Netherlands predominantly represented by paintings from before 1700 and France mainly covered by paintings from 1800–1900.

Wiki-Art *WikiArt* is an online repository of approximately 250,000 high-resolution images of paintings (Saleh and Elgammal, 2016). The selection of artworks and associated metadata follows principles similar to those of Wikipedia, namely a collaborative process augmented by checks and moderation to ensure the quality and reliability of the information provided. As a result, the collection primarily focuses on the golden ages of painting—such as the Northern Renaissance and the Belle Époque—and on very well-known artists, often with extensive coverage of their oeuvre (e.g., Pablo Picasso, Claude Monet, Salvador Dalí, Vincent van Gogh, Henri Matisse, Paul Cézanne, Wassily Kandinsky, Edward Hopper, Leonardo da Vinci, Rembrandt, Gustav Klimt, Egon Schiele, Frida Kahlo, Albrecht Dürer, and Caravaggio). *WikiArt* also contains detailed information on

Figure B2. The geography of (selected) paintings in Northern Europe and in Japan.



(a) Northern Europe (b) Japan

Notes: This figure shows the geolocation of paintings in a subset of *exactly* geolocated production locations in **(a)** Northern Europe (Netherlands, Belgium, North of France, South-East England) and **(b)** Japan. The colors of the dots indicate the period: 1000–1500, 1600–1700, 1700–1800, 1800–1900, 1900–2020.

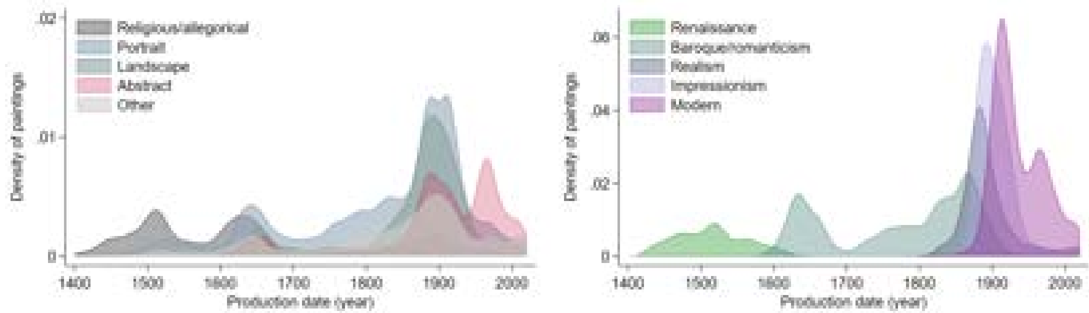
artists, most notably their networks of "influencers" and "influencees."[23]

The collection provides well-organized information on both paintings and artists. It is further enriched by two annotation projects conducted on a subsample of artworks, in which individuals were asked to associate each painting with one of nine emotions. One advantage of *WikiArt* is its highly detailed and structured metadata. For example, artworks are classified into more than 90 artistic movements and a similarly rich set of genres (e.g., abstract, genre painting, landscape, cityscape, animal painting, flower painting, marina, sketch and study, still life, nude painting, portrait, self-portrait, mythological painting, allegorical painting, religious painting).

To harmonize these classifications across sources, we combine structured queries to the Wikipedia API, tags from over 500,000 paintings in ART500K (Mao et al., 2017), and large language models (GPT-4o) to construct 100 harmonized movements (e.g., impressionism, mannerism, fauvism) and 40 harmonized genres (e.g., allegorical painting, landscape, still life). This information allows us to condition the analysis on movement and genre, which may influence how paintings are interpreted by annotators. The temporal distribution of the main movements and genres across all data sources is shown in Appendix Figure B3.

[23]For instance, Paul Cézanne was influenced by: *Gustave Courbet, El Greco, Charles-Francois Daubigny, Nicolas Poussin, Pierre-Auguste Renoir, Eugene Delacroix, Jean-Baptiste-Simeon Chardin*; he had influence on: *Pablo Picasso, Amedeo Modigliani, Jackson Pollock, Fernand Leger, Chaim Soutine, Piet Mondrian, Francis Bacon, Man Ray, Vilhelm Lundstrom, Paul Gauguin, Wassily Kandinsky, Roman Selsky, Adalbert Erdeli, Michel Kikoine, Giorgio Morandi, Jozef Pankiewicz, Robert Falk, Harry Phelan Gibb, Marjorie Acker Phillips, Thomas Hart Benton, Beauford Delaney, William Balthazar Rose*; and he was friend and co-worker with *Paul Gauguin.*

Figure B3. Movements and genre/exercise.



(a) Exercise/genre

(b) Movement

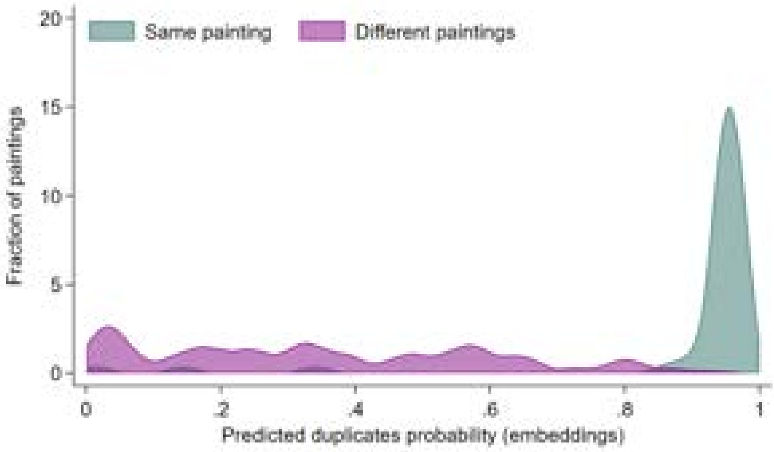
Notes: Panel **(a)** shows the distribution of production years across different categories of exercise (abstract: abstract, genre painting; landscape: landscape, cityscape, animal painting, flower painting, marina, sketch and study, still life, nude painting; portrait: portrait, self-portrait; religious/allegorical: mythological painting, allegorical painting, religious painting; other). Panel **(b)** shows the distribution of production years across different categories of movement (renaissance: early renaissance, northern renaissance, mannerism; modern: symbolism, cubism, expressionism, abstract expressionism, art nouveau, pop art; impressionism: impressionism, post-impressionism; baroque: baroque, rococo, romanticism; realism: realism). These distributions are calculated from artworks in *Google Arts and Culture*, *Wiki-Art*, and *Wiki-Data*.

Wiki-Data The WikiData collection—officially titled the *WikiProject Sum of All Paintings*—is dedicated to the systematic identification and linking of "notable paintings." More specifically "[a] painting is considered notable if one of these options is true: the painting is in a notable collection of a museum, library, or archive (e.g., a painting in the Metropolitan Museum of Art); or the painting is made by a notable artist (e.g., a Van Gogh painting in a private collection)." A direct implication of this criterion is that the collection primarily consists of artworks held in major national and regional museums.

A live summary of coverage is available here. It shows that the largest contributors include many of the world's most prestigious institutions—for example, the Louvre (10,129 pieces), the Metropolitan Museum of Art (13,007 pieces), the Museo del Prado (6,418 pieces), and the Rijksmuseum–as well as prominent national galleries such as the Finnish National Gallery, the Statens Museum for Kunst in Copenhagen, and the National Museum in Warsaw.

Duplicates The previous collections include more than 650,000 paintings, but overlaps exist due to the shared interest in notable artwork. Identifying duplicates is surprisingly challenging because titles often differ across sources—sometimes slightly, sometimes significantly—and many lesser-known paintings lack formal titles altogether. Rather than relying on fuzzy matching based on these short or missing textual descriptions, we use the algorithm described in Section 3.1, which generates a vector of 1,260 embeddings for each painting k, denoted as $z_{e,k}$ in Section 3.1. Letting $\mathbf{z}_i$ represent the vector of 1,260 embeddings associated with painting i, we filter all paintings from the same artist

Figure B4. Detecting duplicates.



Notes: This figure displays the distributions of the embedding-based probability of being duplicates, $\widetilde{\varphi}_{i,j}$, between paintings in a "training" sample. We distinguish between pairs of paintings that have been identified as duplicates (green) and those that have not (purple).

and period across different collections and calculate the distance to these paintings j as:

$$d_{i,j} = (\mathbf{z}_i - \mathbf{z}_j)'(\mathbf{z}_i - \mathbf{z}_j) = \sum_e \left(z_{e,i} - z_{e,j}\right)^2 .$$

We then identify duplicates and non-duplicates for about 1,000 pairs of paintings from the same artists, i.e., we construct an indicator $d_{i,j} \in \{0, 1\}$ within this sub-sample, and estimate the following probit specification,

$$\varphi_{i,j} = P\left(d_{i,j} = 1 \mid d_{i,j}\right) = \Phi\left(\alpha + \beta d_{i,j}\right) .$$

The estimates $(\widetilde{\alpha}, \widetilde{\beta})$ allow us to compute the predicted probability of being a duplicate for any pair of paintings (i, j) across collections, $\widetilde{\varphi}_{i,j}$, which we subsequently use to re-weight our sample. Figure B4 illustrates the predictive power of the embedding-based distance in distinguishing duplicates from non-duplicates within the "training" sample: approximately 90% of true duplicates have a predicted probability above 0.90, compared to only 2% of false duplicates.

Auction and online marketplace data To assess the potential biasing role of (i) commercialization and (ii) selective preservation and digitization, we collect:

- auction data for historical artworks, including pieces that might have been destroyed, remain in private collections, or be held in curated public collections but fall outside our main dataset;

- modern paintings from a large online marketplace, providing access to *both* images and auction outcomes during a period in which material and emotional living standards can be readily measured across contexts.

First, we retrieve and clean entries from sales catalogs digitized by the Getty Research Institute (the *Getty Provenance Index*). These catalogs contain detailed records of public auctions, primarily from Belgium, France, Germany, Great Britain, and the Netherlands, with country coverage varying over time between 1650 and 1945 (e.g., Great Britain in the eighteenth and early nineteenth centuries). After cleaning approximately 20,000 catalogs—including variables such as price, clearly identified painter, and artwork title—we retain roughly 1,200,000 transactions, sometimes involving the same artwork. The dataset enables the identification of artists (e.g., recurring figures such as Brueghel the Elder, Van Dyck, Poussin, Rembrandt, Rubens, and Teniers the Younger across different artworks and over time), as well as recurring sellers and buyers, typically painters, art collectors, aristocrats, or wealthy investors. We use the high-quality reporting of artist names to link catalog entries to artists in our primary, harmonized collection of preserved paintings. This linkage allows us to assess the extent and direction of attrition, commercialization, and selection into conservation.

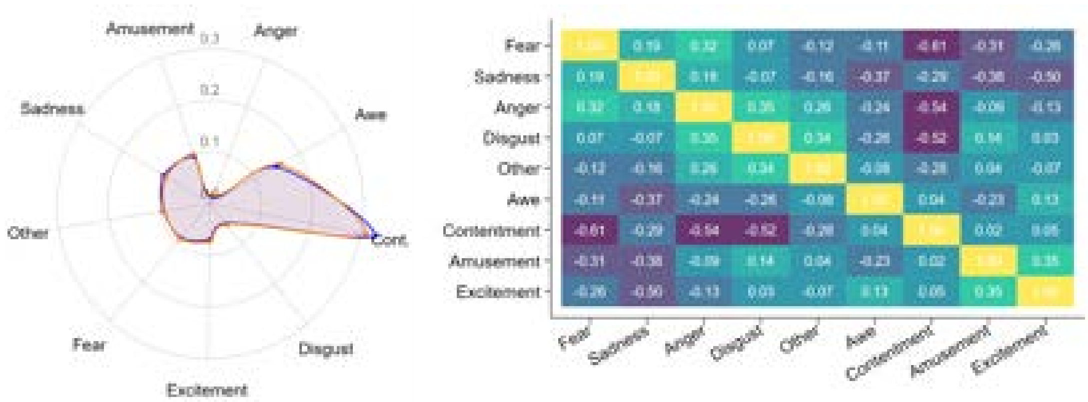
Second, we collect modern artwork from the online marketplace *Artsy*. The platform records realized auctions, provides high-quality images, and includes extensive artwork- and artist-level metadata, such as biographies, prizes and awards, auction histories, and—at the artwork level—title, material, dimensions, tags generated by the Art Genome Project, and price. We restrict the sample to 710,121 *paintings*.

B.2 A collection of *annotated* images

A collection of *annotated* images Our training sample consists of two sources of emotion annotations associated with *WikiArt* images: ArtEmis (Achlioptas et al., 2021, which provides approximately 450,000 labels for 79,860 paintings); and ArtELingo (Mohamed et al., 2022, which provides more than 1,000,000 labels for the same 79,860 paintings). Both datasets include textual explanations justifying the assigned emotion. The latter dataset further leverages variation across annotators in language and cultural background, thereby covering a wide range of countries.

Figure B5 displays the average emotion scores and the cross-correlations across the following categories: amusement, anger, awe, contentment, disgust, excitement, fear, sadness, and other. As discussed in Section 2, the most frequently assigned emotions are contentment (e.g., landscapes), awe, amusement, sadness, and fear (panel a). Panel (b) reports correlations among the resulting emotion scores: fear, anger, and disgust form a closely related cluster of negative emotions and appear to function as substitutes. At the other end of the spectrum, contentment emerges as a largely distinct emotion and is rarely mentioned alongside other emotions in a systematic way.

Figure B5. Average emotion scores and cross-correlations.



(a) Scores (b) Correlations

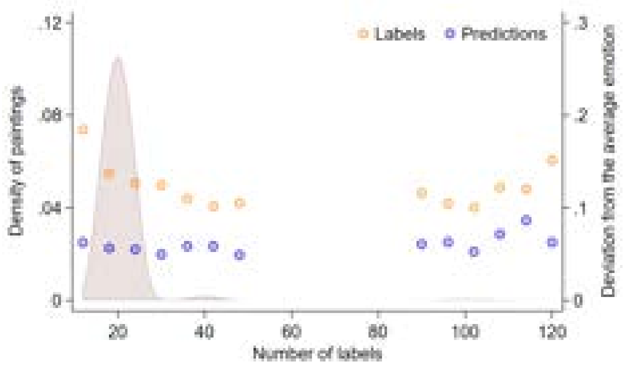
Notes: Panel **(a)** shows the distribution of emotion scores across 79,860 annotated paintings—the annotations in orange and the predictions in blue. The respective scores are (standard deviations in parentheses): *amusement*, 0.101 (0.058); *anger*, 0.018 (0.019); *awe*, 0.168 (0.074); *contentment*, 0.302 (0.123); *disgust*, 0.048 (0.048); *excitement*, 0.071 (0.039); *fear*, 0.095 (0.090); *sadness*, 0.101 (0.086); and *other*, 0.094 (0.044). Panel **(b)** shows the correlation between each dimension of the emotion score associated with the 79,860 annotated paintings.

A probabilistic vector of emotions In our baseline strategy, we construct a probabilistic vector of emotions, $\mathbf{q}_j$, using the 15–25 labels associated with each painting j. Specifically, we interpret these multiple labels as a probabilistic evaluation aggregated across individuals. Each painting is assigned a vector of scores representing the share of annotators who selected each of the nine emotions, normalized so that the scores sum to one for each painting. In the presence of measurement error and idiosyncratic annotator traits or preferences, the estimated scores could, in principle, depend on the number of labels provided. However, Appendix Figure B6 (a) shows that the average score does not vary systematically with the number of annotators.

While the number of annotations does not appear to affect the average emotion attributed to a painting, it could influence the dispersion of scores across artworks. For example, if each annotator provided a purely random, idiosyncratic assessment, the resulting scores would be expected to converge to a stable value as the number of annotations increases. We find no evidence supporting this pattern. As shown in Appendix Figure B6 (b), artworks with fewer annotations are not significantly more likely to deviate from the average rating observed in the overall population of paintings.

Figure B6. Average and dispersion in annotations, as a function of the number of labelers.

(a) Scores by labels



(b) Variance by labels

Notes: This figure explores the distribution of ratings and their number per artwork. Panel (a) replicates panel **(a)** of Figure B5 for artworks associated with fewer than 17 labels (dark blue), between 18 and 22 labels (blue), and more than 23 labels (light blue). Panel **(b)** shows a measure of deviation from the average emotion within the annotated sample, based on labels (orange) and predictions (blue), as a function of the number of provided labels. To construct these deviations from the mean, we collapse the different emotion ratings for a painting into an empirical probability distribution, i.e., the probability that one of the associated annotators reports a given emotion, and compute the squared difference from the overall sample average (represented as dots). Data sources: ArtEmis (Achlioptas et al., 2021) and ArtELingo (Mohamed et al., 2022).

B.3 Other data sources

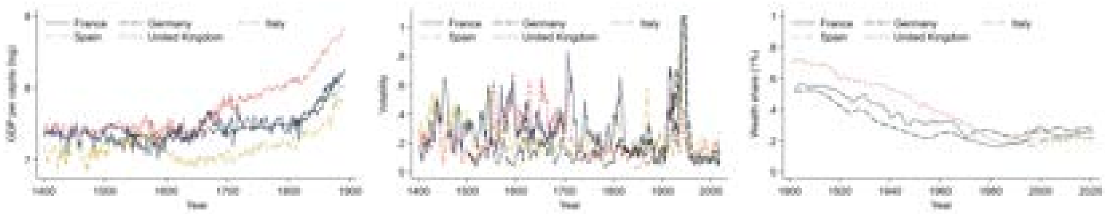
We have assembled a range of data sources capturing historical economic conditions. Below, we provide a more detailed description of these sources.

Sentiment measures In Section 3.3, we validate our sentiment-extraction procedure from artwork against measures of happiness and emotions in recent periods. We do so by collecting all available interviews from Gallup (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025) and World Value Surveys (270,760 interviews) from 2004 onward. Gallup provides near-annual coverage for each country, with at least 1,000 interviews per country and year. Our baseline analysis exploits the following questions (and responses): *How is life today?* (from *worst possible* to *best possible*, on a 0–10 scale); *feelings about household income* (from *finding it very difficult on present income* to *living comfortably on present income*, on a 1–4 scale); *not enough money for [food/shelter]* (*yes*/*no*); and *experienced [emotion] yesterday* (*yes*/*no*, where emotions include stress, anger, worry, and sadness, among others). For exposition purposes, we normalize these variables between 0 and 1. The World Values Survey covers fewer countries and is conducted at irregular intervals (three waves between 2004 and 2021); it also includes fewer questions related to life satisfaction. Our sensitivity analysis exploits the following questions (and responses): *satisfaction with your life* (from *completely dissat-*

isfied to *completely satisfied*, on a 1–10 scale); and *feeling of happiness* (from *not at all happy* to *very happy*, on a 1–4 scale).

For historical periods, available data sources are much more limited. We therefore incorporate a "valence index" derived from sentiment analysis of published books across four countries over roughly 100 years (Hills et al., 2019).

Figure B7. The variation in standard economic measures.



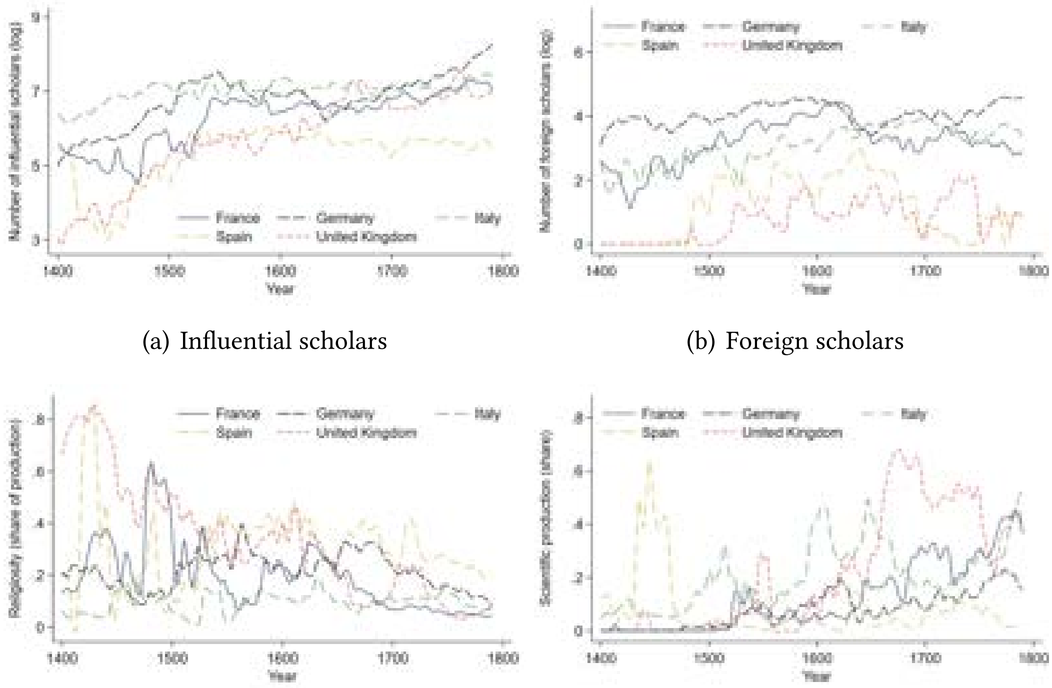
(a) GDP per capita (b) Uncertainty (c) Wealth inequality

Notes: These figures display the variation in France, Germany, Italy, Spain, and the United Kingdom for: **(a)** recently revised estimates of (log) GDP per capita (Maddison Project Database, see Bolt and Van Zanden, 2025); **(b)** the volatility of GDP per capita calculated over a window of [-2,+4] years ($\sum_{\tau=-2}^{4} |\hat{y}_{t+\tau}|$, controlling for $\sum_{\tau=-2}^{4} \hat{y}_{t+\tau}$); and **(c)** the concentration of wealth among the top-1% of individuals, mostly from the mid-nineteenth century onward (Alvaredo et al., 2020). This variation is used more systematically across countries and periods in Figure 11 of the paper.

Indicators of economic development In Section 4, we contrast fluctuations in emotions against standard indicators of economic development. First, the Maddison Project provides estimates of output and population for a selected sample of countries from 1400 onward; our analysis relies on their most recent revisions (Maddison Project Database; see Bolt and Van Zanden, 2025). We use these data to construct key macroeconomic variables, illustrated in Appendix Figure B7: (a) economic development, measured as (log) GDP per capita (Bolt and Van Zanden, 2025); and (b) economic uncertainty, proxied by six-year implied volatility in GDP per capita (Bloom, 2014). Second, the World Inequality Database allows us to capture economic inequality through the concentration of wealth held by the top 1% or 10% of individuals, with particularly strong coverage for France (Alvaredo et al., 2020, see also Appendix Figure 11 and Appendix Figure B7). Finally, we use historical estimates of life expectancy at birth (Zijdeman and Ribeira da Silva, 2015) in a robustness check.

Large socioeconomic changes Our analysis aims to capture several key transformations in Medieval Europe, the Renaissance, the later Reformation, the Enlightenment, and the Industrial Revolution. To this end, we first draw on data documenting the adoption of new technologies (Cross-country Historical Adoption of Technology; see Comin and Hobijn, 2004), focusing on two technologies highlighted in Figure 14: (i) steamships,

Figure B8. The production of knowledge over time.



(a) Influential scholars

(b) Foreign scholars

(c) Religiosity

(d) Sciences

Notes: This figure illustrates the variation in the intensity and nature of knowledge production in France, Germany, Italy, Spain, and the United Kingdom between 1400 and 1800 (De La Croix et al., 2024). Panel **(a)** reports a measure of "quality" or influence, as captured by the (log) number of scholars with more than 10 Wikipedia pages (in different languages). Panel **(b)** reports a measure of openness, defined by the (log) number of foreign scholars. Panel **(c)** reports the weighted share of knowledge production in theology. Panel **(d)** reports the weighted share of knowledge production in the sciences. This variation is used more systematically across countries and periods in Figure 15 of the paper.

which contributed to the first wave of globalization, and (ii) radio (including its potential use as a tool of propaganda; see Adena et al., 2015; Castiello, 2025).

The Macro History Database provides more conventional measurements of macroeconomic variables from the mid-nineteenth century onward for 18 economies (Jordà-Schularick-Taylor Macrohistory Database, see Jordà et al., 2017). One of these variables—trade openness (the ratio of imports and exports to GDP), which we analyze in Section 4 and Figure 13—allows us to capture the dynamics of early globalization.

This period is marked by the rise of universities, the democratization of knowledge, a decline in religiosity, and the onset of the scientific "revolution." We capture these long-run secular changes using an extensive database of scholars and their host institutions across Europe (The Rise of Universities in Medieval and Early Modern Europe; see De La Croix et al., 2024).

The scholars data are recorded at the individual level, and we aggregate them to the country-year level in three steps. First, we assign characteristics to each scholar, including measures of influence (such as the number of listed publications, Wikipedia page length, and number of language editions), whether the scholar is foreign (based on country of birth), and their main academic field (theology, law, arts, medicine, or sciences). Second, we construct a scholar-year panel to account for spells across different institutions. Third, we aggregate these data at the country-year level using present-day borders. We focus on three main outcomes: (i) the number of influential scholars (defined as those with more than ten Wikipedia language editions); (ii) the number of foreign scholars; and (iii) the weighted share of knowledge production across fields. For this third measure, we restrict attention to influential scholars, though results are robust to alternative weighting schemes.

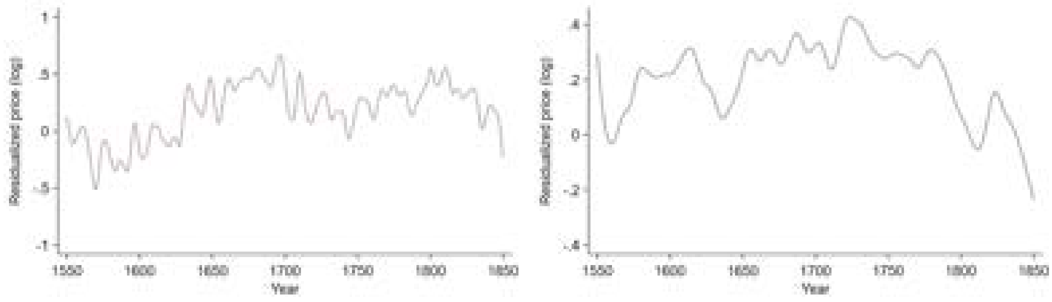
Figure B8 illustrates historical patterns in France, Germany, Italy, Spain, and the United Kingdom, based on the scholars database: (a) knowledge production rises over time in all countries; (b) France, Germany, and Italy heavily rely on foreign scholars—Great Britain less so—and this reliance reflects both secular trends and idiosyncratic movements; (c) religiosity follows markedly different trends across economies, declining sharply in Great Britain from 80% of academic output in 1400 to under 10% by 1700; (d) scientific production follows the opposite pattern, with a notable rise during the mid-seventeenth century in Great Britain (and to a more moderate extent in France)—a pattern not clearly evident in the Holy Roman Empire, which was in its Baroque period. Time variation in scientific production is particularly pronounced in Italy, with major peaks during the early Scientific Revolution (e.g., Leonardo da Vinci, Galileo Galilei), followed by a sharp decline after 1630 with the onset of the Counter-Reformation and church censorship. We use—and in some cases reproduce—this variation in Figure 15.

Shocks to livelihoods The previous section focused on longer-term, secular changes in livelihoods across the period of study. In this section, we turn to more transient fluctuations throughout the period of interest.

First, we capture political turbulence using regime characteristics and transitions documented from 1800 onward by the *Polity IV* project (Marshall et al., 2014). Second, we identify key historical events that affected political entities corresponding to 26 present-day countries, e.g., major wars.[24] Third, we capture short-lived changes in

[24] We define a sample of present-day territories for which we track socio-economic changes and shocks: Argentina, Australia, Belgium, Brazil, Canada, China, Czech Republic, Denmark, France, Germany, India, Italy, Japan, Mexico, New Zealand, Norway, Poland, Portugal, Russia, South Africa, Spain, Sweden, Switzerland, Turkey, the United Kingdom, and the United States. In each of these countries/territories, we identify country-specific shocks (e.g., colonization, international and civil wars, pandemics, revolutions, political upheaval, new trade arrangements, famines, and financial crises), as well as the timing of major socio-economic transformations (e.g., the emancipation of slaves, female emancipation, the arrival of new communication and transportation technologies, industrialization, mass urbanization, and the emergence and regulation of labor unions). For each event and country, we record a unique event identifier, the start

Figure B9. The price of commodities over time.



(a) Bread

(b) Luxury

Notes: This figure illustrates the variation in the (residualized) price of commodities in the United Kingdom (mostly England) between 1550 and 1850. The procedure involves residualizing (log) prices of standardized measures and aims to account for differences in units and currencies, the selection of markets entering or exiting the database, and the possible aggregation of similar commodities. Panel **(a)** displays the variation in a "basic" commodity, bread, and panel **(b)** focuses on a "luxury" commodity based on several different goods (e.g., including silk, spices). This variation is used more systematically across countries and periods in Appendix Figure D11 and Appendix D.3.

the price of basic commodities. To do so, we rely on a database of commodity prices at the market level (Allen-Unger Global Commodity Prices Database; see Allen and Unger, 2019). Constructing measures at the country-year level is challenging due to differences in units and currencies, variation in market coverage over time, and the possible aggregation of similar commodities. We consider the (log) price of a standardized measure of commodity c within a market m in year t, and consider,

$$p_{cmt} = \alpha_c + \beta_m + \gamma_t + \varphi_{cmt}.$$

Next, we aggregate the residuals $\{\varphi_{cmt}\}$ within commodity category d, a given country i, and year t, using the prevalence of a market/commodity over time as weights. We illustrate the variation in those quantities for two "commodities", bread (basic, used in Appendix Figure D11, panel (a) and a luxury basket (e.g., including silk, spices, used in Appendix Figure D11, panel (b), in the United Kingdom (see Figure B9).

Finally, Section 4.4 examines the emotional response conveyed through paintings to the secular but non-negligible climate variability affecting Europe from 1500 onward. Our approach relies on the harmonized temperature reconstructions from Luterbacher et al. (2004), which are based on multiple sources, including tree-ring data and ice cores. We report their temperature predictions for the main European "purveyors" of paintings in Figure B10. Two notably cold decades, around 1600 and 1700, are common across most European countries, though more pronounced in Germany, Denmark, and the Nether-

year (and, where applicable, the end year), a textual description, and whether and which other countries are involved.

Figure B10. Climatic conditions over time.

(a) France (b) United Kingdom (c) Germany (d) Italy (e) Spain

(f) Netherlands (g) Austria (h) Denmark (i) Belgium (j) Russia

Notes: This figure illustrates the average temperature between 1500 and 2000 in France, Great Britain, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, and Russia—as extracted from Luterbacher et al. (2004). One can observe secular fluctuations in temperature, sometimes common across European countries, and sometimes more country-specific. For visualization purposes, we report a 30-year moving average. This variation is used more systematically across countries and periods in Figure 12 of the paper.

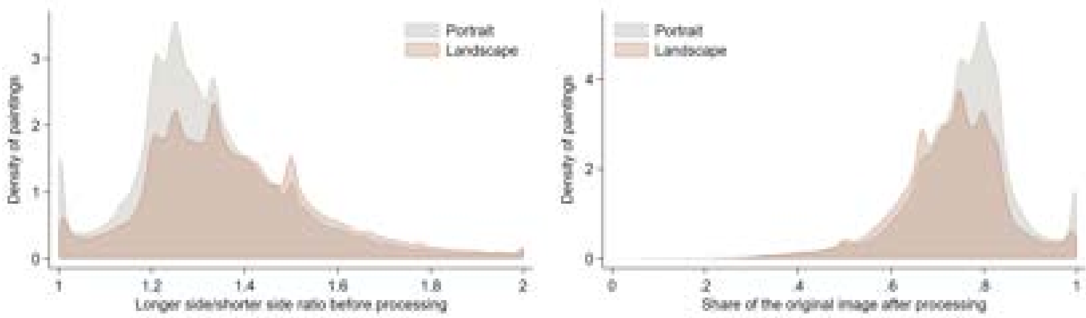
lands. Approximately 60% of the within-country temporal variation is explained by a “common” component—a share amplified by the anthropogenic nature of climate warming in the late twentieth century.

C The encoding of paintings through computer vision

C.1 A sentiment classification pipeline

Image pre-processing While convolutional neural network architectures offer flexibility in how pixel values interact and contribute to the final prediction, they are more rigid in the required input format. Therefore, all images undergo pre-processing to convert them into a standardized $384 \times 384 \times 3$ array (height, width, color). Each pixel in this array contains Red-Green-Blue (RGB) information, which is stored as an integer between 0 and 255.

Figure C1. Image pre-processing.



(a) Longer side/shorter side ratio

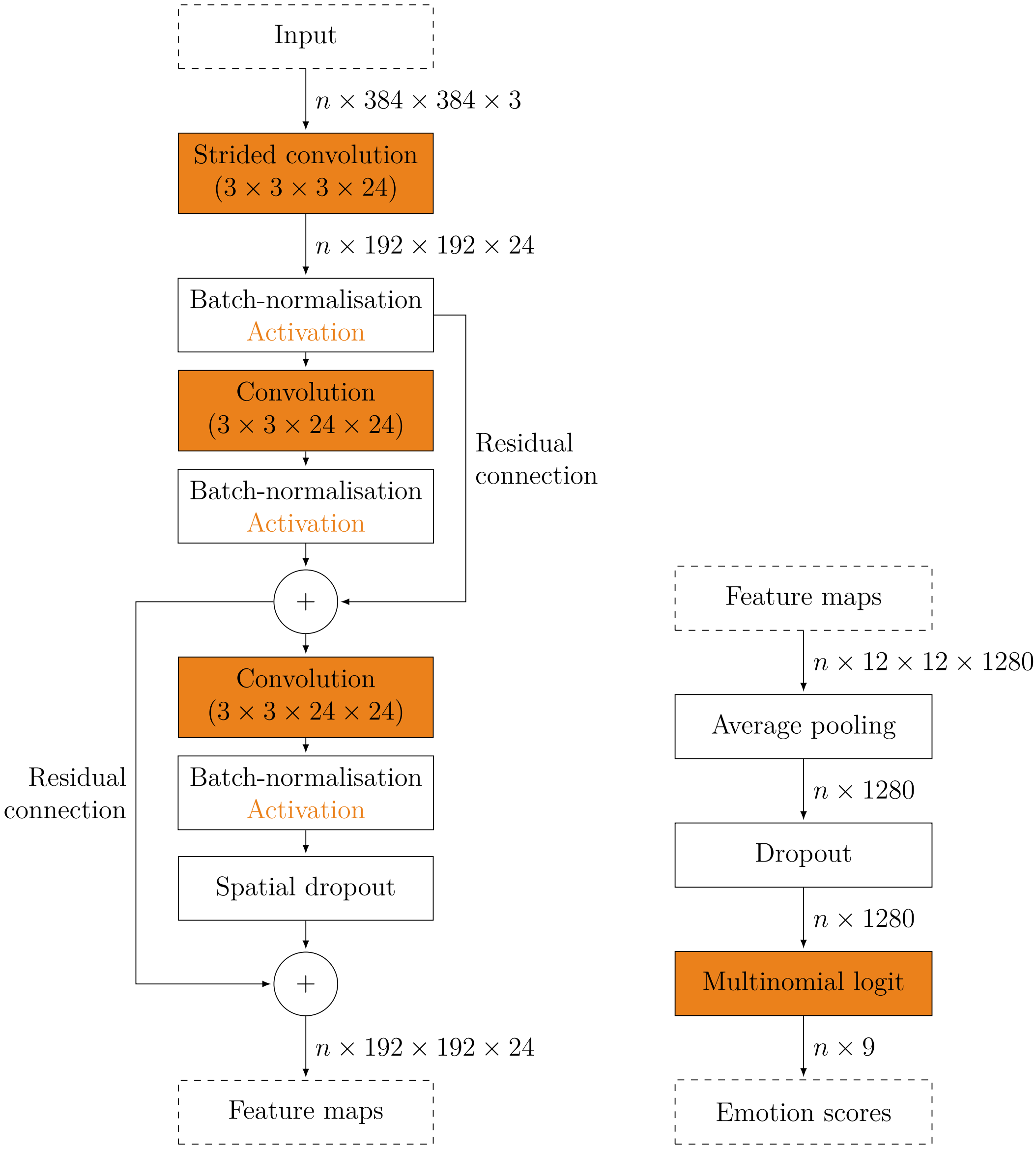
(b) Coverage

Notes: Panel **(a)** shows the distribution of the longer side/shorter side ratio across images in our harmonized collection; panel **(b)** displays the distribution of the average share of these initial images that is retained through our image pre-processing procedure. The areas of the two distributions reflect their contribution to the final sample of artwork. These distributions are calculated from artworks in *Google Arts and Culture*, *Wiki-Art*, and *Wiki-Data*.

This pre-processing step has two main implications: (i) the resulting image is slightly blurred compared to the original painting, and (ii) the final image may not retain all features of the original, as only the largest possible square with the same central point as the original image is preserved. Despite this, Figure 3 illustrates that reducing the original image to a 384×384 collection of pixels does not significantly alter our general perception of the painting. To further explore potential losses due to the exclusion of non-central pixels, we analyze the distribution of the longer side to shorter side ratio for paintings oriented as portraits or landscapes (see Figure C1). The results indicate that about 80% of centered images retain more than 70% of the original pixels. In theory, these preserved sections of the image should encompass the most important areas, typically around the focal point.

Structure, estimation, and optimization The prediction model integrates a pre-trained encoder model with a regression head. Intuitively, the encoder extracts a large

Figure C2. Structure of the model.



(a) EfficientNet block (b) Regression head

Notes: This figure illustrates the architecture used to generate emotion scores from input images. Panel **(a)** shows the first EfficientNet block, which processes input images of size $n \times 384 \times 384 \times 3$ and outputs features of dimension $n \times 192 \times 192 \times 24$. The block includes strided and residual convolutions, batch normalization, activation functions, and spatial dropout. Subsequent blocks contain attention mechanisms (i.e., "squeeze and excite" blocks) that enable the model to dynamically focus on specific features depending on global context information (i.e., global pooling and fully connected layers). Panel **(b)** displays the regression head, which takes as input feature maps of size $n \times 12 \times 12 \times 1280$ produced by the last EfficientNet block, applies global average pooling and dropout, and outputs a probabilistic vector of nine emotion scores through a multinomial logit layer.

number of features from the images, while the regression head uses these features to predict emotion scores. The encoder is based on the convolutional layers of an EfficientNet V2-S model (Tan and Le, 2021), with the original output layer removed and replaced by our regression head. One advantage of using a pre-trained model is that it leverages extensive training—here, on 1,400,000 photos of objects across 1,000 different classes. However, the model was originally designed to predict object classes, not emotions, and was trained on real-life pictures rather than paintings. To address this, we attach a regression head to the encoder, mapping the extracted features to scores associated with each emotion. Letting j denote a painting, $\mathbf{x}_j$ the vector of extracted features from the encoder model, and y_j the latent emotion of painting j, the regression head assumes,

$$p_j^e = P\left(y_j = e|\mathbf{x}_j\right) = \frac{e^{z_{e,j}}}{\sum_{e\in\mathcal{E}} e^{z_{e,j}}},$$

where latent factors, $z_{e,j} = f_e(\mathbf{x}_j, \mathbf{b})$, are estimated through a non-parametric estimation, allowing to capture non-linearities and interactions between input variables.

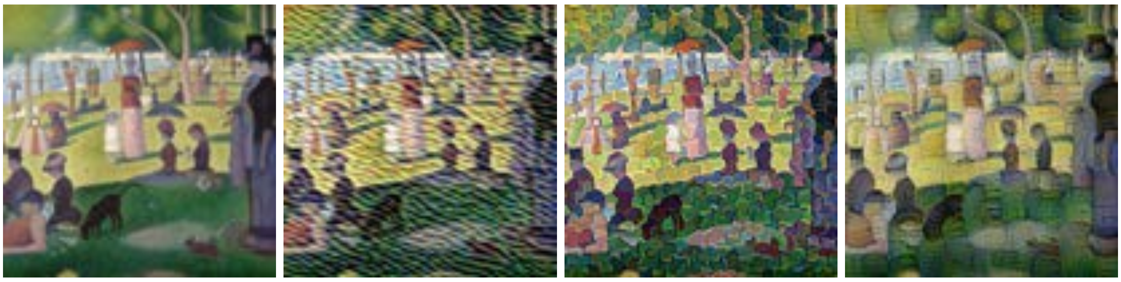
Figure C2 provides a detailed description of both the encoder and the regression head—with the important components highlighted in orange. The regression head processes the extracted features through two densely connected hidden layers, each containing 64 units with ReLU activation, followed by batch-normalization (Ioffe and Szegedy, 2015) and dropout layers (Tompson et al., 2014). The final layer applies the previous multinomial logistic transformation, ensuring the emotion scores sum to 1 and allowing them to be interpreted as probabilities.

The model estimation minimizes the Kullback-Leibler divergence between the predicted and observed score distributions, better accounting for the probabilistic nature of our target and limiting the occurrence of false negatives—instances where true high emotion scores are predicted as low. The Kullback-Leibler divergence is defined as:

$$\sum_{e\in\mathcal{E}} \left(q^e \ln(q^e) - q^e \ln(p^e)\right).$$

where q^e is the actual probability that one of the 15-25 labelers chooses emotion e as the dominant emotion, and p^e is the predicted probability. In practice, we estimate the model using maximum likelihood with an adaptive learning rate for each parameter, dynamically estimated learning rates, and sub-samples of 64 images at each optimization iteration (as in Kingma and Ba, 2015). The optimization process begins by freezing the encoder parameters, taking the pre-trained feature extraction as given and aligning the regression head to predict emotions. Once the regression head parameters are well initialized, we fine-tune the model by (i) unfreezing the encoder parameters (except for the batch-normalization layers, which are updated using an exponential moving average) and (ii) training the model with a small learning rate to avoid over-fitting. The rationale is straightforward: the pre-trained encoder may not extract features optimally for our purposes, but its initial parametrization is valuable for initializing the regression head (justifying step i). Optimizing the entire model simultaneously would likely cause large gradient updates from the head, initialized with random values, potentially disrupting the encoder's ability to extract meaningful patterns. However, as the process of conveying emotions through paintings differs from real-life pictures—as illustrated by Figure 3 and Figure 4—the encoder model requires updating (justifying step ii).

Figure C3. Feature projections.



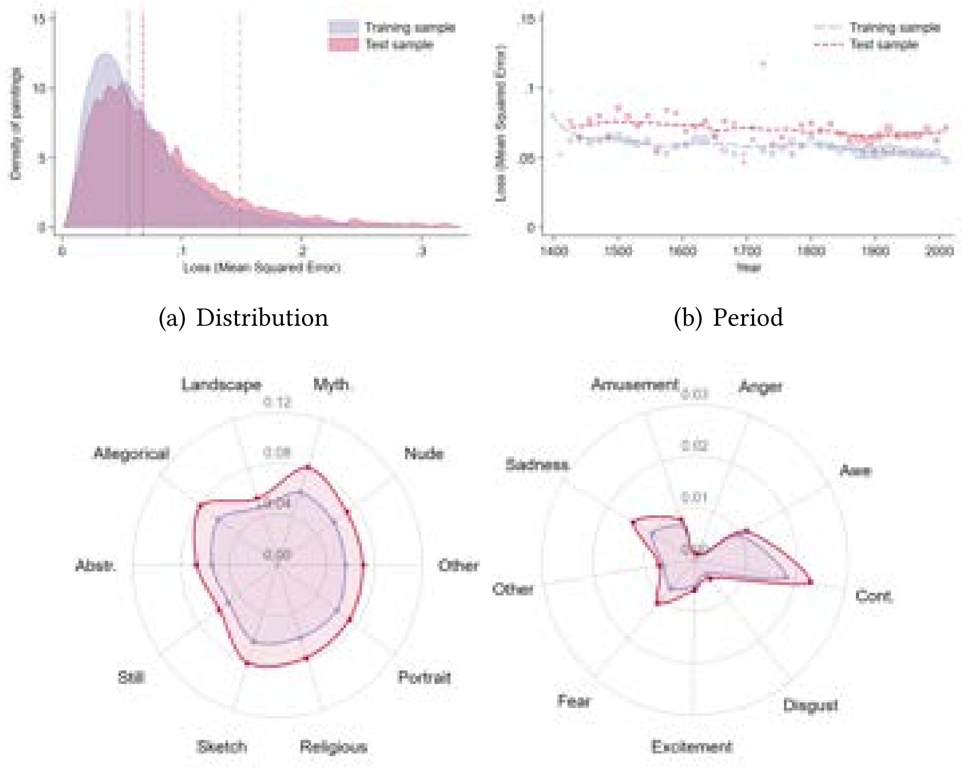
(a) Input image (b) Early layer (c) Intermediate layer (d) Deep layer

Notes: This figure shows visualizations of activations from selected filters at different depths in the network, projected onto the original image of *A Sunday on La Grande Jatte* by Georges Pierre Seurat (1884), panel **(a)**. Panel **(b)** displays activations from early layers, panel **(c)** from intermediate layers, and panel **(d)** from deep layers. In general, different filters specialize in detecting distinct patterns *within* layers, while between-layer hierarchy allows for the identification of increasingly abstract features at larger spatial scales. Specifically, early layers respond to fine-grained features such as colors, edges, and textures; middle layers capture motifs and patterns at intermediate spatial scales; and deep layers extract large-scale, abstract features, often giving rise to surreal or composite visual elements.

Convolutional layers and composition The use of multiple convolutional layers is fundamental to convolutional networks, as it enables the model to capture hierarchical interactions and intensity patterns at different spatial scales. Each layer builds on the previous one, generating increasingly abstract features and spatial relationships within the image. For example, initial layers typically capture simple features such as edges or textures, while deeper layers combine these elements into more complex patterns, ultimately yielding high-level representations. This hierarchical structure allows convolutional networks to efficiently represent the intricate visual structure of images, and of paintings in particular.

One way to illustrate this functionality is through filter visualization at different stages of the network (Zeiler and Fergus, 2014). In Figure C3, we consider the painting *A Sunday on La Grande Jatte* by Georges Pierre Seurat (1884) and display the activations of successive hierarchical layers projected back onto the original image (Szegedy et al., 2015). These projections help illustrate how convolutions at different depths extract information related to texture, contrast, and composition. For instance, panel (d)—corresponding to the deepest layer(s)—highlights the symmetry in the foreground (between the two girls/women, and between the women and the dog), the overall color contrasts, and textural patterns.

Figure C4. Model validation—mean squared error.



(a) Distribution (b) Period

(c) Genre (d) Emotion

Notes: These figures provide statistics on the Mean Squared Error (MSE) metric between the emotion ratings and the predicted ratings within the training sample and the test sample. The metric is defined as $\sum_{e\in\mathcal{E}} (q^e - p^e)^2$, where $\{q^e\}_{e\in\mathcal{E}}$ is the actual probability distribution and $\{p^e\}_{e\in\mathcal{E}}$ is the predicted probability distribution. Panel **(a)** shows the distribution of this metric within the training sample (light blue) and the test sample (red); the median metric is displayed as dashed lines. For interpretation purposes, we report the median distance between predictions and randomly reshuffled labels to provide a fully uninformed benchmark. Panels **(b)** and **(c)** show the median distance metrics across sub-samples: (i) production periods, and (ii) genre or type of exercise—where *Abstr.* stands for abstract, genre painting; *Landscape* stands for cityscape, landscape, marina; *Still* stands for animal painting, flower painting, still life; *Sketch* stands for sketch and study; *Nude* stands for nude painting; *Portrait* stands for portrait, self-portrait; *Religious* stands for religious painting; *Myth.* stands for mythological painting; and *Allegorical* stands for allegorical painting. Finally, panel **(d)** displays the emotion-specific error, i.e., the median of $\left[(q^e - p^e)^2\right]$ across all $e \in \mathcal{E}$. Relative to their importance, *sadness* and *fear* appear to have a slightly lower fit within the test sample.

Model validation In Section 3.1, we discuss the fit of the model across the test and training samples. The analysis then relies on the (targeted) Kullback-Leibler divergence.

To further evaluate the properties of the estimated model—specifically its ability to predict emotion scores on an out-of-sample set of paintings—we calculate the mean squared error (MSE), $E\left[\sum_{e\in\mathcal{E}} (q^e - p^e)^2\right]$, between the actual emotion ratings, $\{q^e\}_{e\in\mathcal{E}}$, and the predicted ratings, $\{p^e\}_{e\in\mathcal{E}}$, within the training sample (70% of the annotated dataset) and the test sample (15% of the annotated dataset, with the remaining 15% used for validation). The mean squared error provides an intuitive metric, allowing us to decompose the error by emotion. Figure C4 reveals several key insights: (i) the MSE is generally low (around 0.056 within the training sample and 0.067 within the test sample, indicating a minor prediction error, such as predicting 0.83 for contentment and 0.17 for excitement instead of 1 for contentment); (ii) the error is slightly lower for more recent paintings but the difference is not substantial (panel b); (iii) higher errors are observed for sketches, religious, and mythological paintings, likely due to their limited representation in the annotated sample (panel c); and (iv) the emotion-specific MSE shows larger errors for emotions with higher probabilities of being mentioned, but, relative to their importance, sadness and fear appear to have a slightly lower fit within the test sample (panel d).

Figure C5. Examples of predictions (*co*: contentment, *am*: amusement, *ex*: excitement, *aw*: awe, *fe*: fear, *an*: anger, *sa*: sadness, *di*: disgust).

(a) p^{co}: 0.70, q^{co}: 0.66 (b) p^{am}: 0.20, q^{am}: 0.21 (c) p^{ex}: 0.24, q^{ex}: 0.26 (d) p^{aw}: 0.65, q^{aw}: 0.55

(e) p^{fe}: 0.50, q^{fe}: 0.52 (f) p^{an}: 0.08, q^{an}: 0.08 (g) p^{sa}: 0.23, q^{sa}: 0.25 (h) p^{di}: 0.20, q^{di}: 0.22

Notes: This figure illustrates our model performance on selected paintings from the test sub-sample, where the latter are chosen as representative of each emotion. We display the dominant emotion for the target probability distribution, $\{q^e\}_{e\in\mathcal{E}}$, and the predicted probability distribution, $\{p^e\}_{e\in\mathcal{E}}$. Panel **(a)** shows *House with Red Roof* by Paul Cézanne (1890); the prediction and target for *contentment* are (p^{co}: 0.70, q^{co}: 0.66). Panel **(b)** displays *The Beauty and the Beast* by Leopold Survage (1939); the prediction and target for *amusement* are (p^{am}: 0.20, q^{am}: 0.21). Panel **(c)** shows *Street Decked with Flags* by Raoul Dufy (1906); the prediction and target for *excitement* are (p^{ex}: 0.24, q^{ex}: 0.26). Panel **(d)** displays *The Grand Canyon of the Yellowstone* by Thomas Moran (1872); the prediction and target for *awe* are (p^{aw}: 0.65, q^{aw}: 0.55). Panel **(e)** shows *Rough Sea* by Ivan Aivazovsky (1844); the prediction and target for *fear* are (p^{fe}: 0.50, q^{fe}: 0.52). Panel **(f)** displays *The Four Horsemen of the Apocalypse* by Viktor Vasnetsov (1887); the prediction and target for *anger* are (p^{an}: 0.08, q^{an}: 0.08). Panel **(g)** shows *During Haying* by Efim Volkov (1883); the prediction and target for *sadness* are (p^{sa}: 0.23, q^{sa}: 0.25). Panel **(h)** displays *Sculpture Drawing, SF64-571* by Sam Francis (1964); the prediction and target for *disgust* are (p^{di}: 0.20, q^{di}: 0.22).

In Figure C5, we provide a few illustrations of prediction quality by selecting eight artworks with a "dominant" emotion for which we benchmark the target dominant probability, q^e, against the predicted dominant probability, p^e (*co*: contentment, *am*: amusement, *ex*: excitement, *aw*: awe, *fe*: fear, *an*: anger, *sa*: sadness, and *di*: disgust).[25] In general, the model tends to accurately predict the most common emotions associated with paintings, i.e., awe (illustrated by opulent buildings), contentment (peaceful landscapes), excitement (through warm colors, lights and human activity), or fear (cold colors and indistinct texture evoking uncertainty).

Figure C6. The identification of emotion in paintings—additional evidence (1/2).

(a) Contentment (b) Amusement (c) Excitement (d) Awe

(e) Fear (f) Anger (g) Sadness (h) Disgust

Notes: This figure displays heatmaps generated using Class Activation Mapping (Zhou et al., 2016)—which isolate the parts of the image most used by the network to predict a latent emotional probability. We apply this methodology to a single painting—*The Promenade Woman with a Parasol* by Claude Monet, 1875—and display eight different mappings, each corresponding to one emotion (excluding "*other*"). The colors, from blue to green, yellow, and red, indicate the zones of activation for each dominant emotion.

Illustrations In Section 3.1, we introduce a convenient dichotomy to validate our general procedure, based on two hypotheses: (H1) the model predicts the emotions perceived by annotators (discussed in the previous section); and (H2) annotator labels—and the resulting predictions—capture the artists' original intentions. In Figure 6, we provide anecdotal support for hypothesis (H1) using an inversion procedure. More specifically, our neural network combines a large number of parameters into connected layers that operate on localized regions of the image, then aggregates these signals across the entire image to produce a vector of emotional scores. A technique known as Class Activation Mapping partially inverts this process and allows us to measure the contribution of

[25] Anger is rarely mentioned by annotators such that it is rarely allocated a high probability, either in the annotated sample or in the prediction set.

different image regions to the variance in the predicted emotions (Zhou et al., 2016).

Figure C7. The identification of emotion in paintings—additional evidence (2/2).

(a) Contentment (b) Amusement (c) Excitement (d) Awe

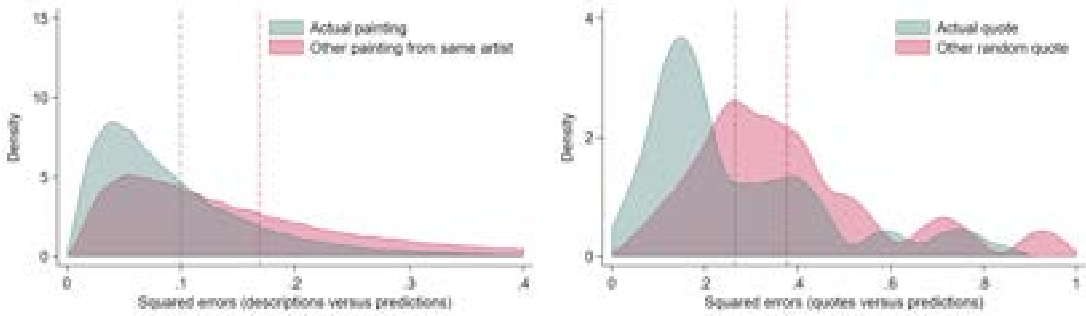
(e) Fear (f) Anger (g) Sadness (h) Disgust

Notes: This figure displays heatmaps generated using Class Activation Mapping (Zhou et al., 2016)—which isolate the parts of the image most used by the network to predict a latent emotional probability. We apply this methodology to a single painting—*The Raft of the Medusa* by Théodore Géricault, 1818–1819—and display eight different mappings corresponding to each emotion (excluding "*other*"). The colors, from blue to green, yellow, and red, indicate the zones of activation for each dominant emotion.

In Section 3.1, we use this methodology to produce a mapping for each dominant emotion across the eight artworks displayed in Figure 3. In Figures C6 and C7, we instead focus on a single artwork at a time and display eight distinct mappings corresponding to each emotion (excluding "other"). Figure C6 shows that the two characters in *The Promenade Woman with a Parasol* (Claude Monet, 1875) are predicted to convey contentment, while the woman's robe—and its intricate folds—weakly suggests awe. Figure C7 considers a negatively toned painting, *The Raft of the Medusa* by Théodore Géricault (1818–1819), which depicts a real historical event (the wreck of a French frigate). The central figures—crying for help—are predicted to convey fear, while the recumbent (possibly dead) bodies strongly suggest sadness. Importantly, image regions are not exclusive in their emotional content: there is overlap between the heatmaps for fear and sadness, and, conversely, large areas of the image are not strongly associated with any specific emotion.

Predictions and the intentions of artists The preceding examples provide visual support for hypothesis (H1). Testing hypothesis (H2) is considerably more challenging,

Figure C8. The internal and external validity of annotated and predicted emotions.



(a) Descriptions against labels

(b) Predictions against quotes

Notes: Panel **(a)** displays the distribution of squared errors, $\sum_e (q^e - \vartheta^e)^2$, between the probabilistic vectors of emotions, as averaged across the emotions provided by annotators, $\{\mathbf{q}\}_i$, and the probabilistic vectors of emotions associated with their justifications, $\{\vartheta\}_i$. The red distribution provides a benchmark in which these justifications are randomly allocated across paintings within the same artist. Panel **(b)** displays the distribution of squared errors, $\sum_{e \in \mathcal{E}} (q^e - \rho^e)^2$, between the probabilistic vectors of emotions, as averaged across the emotions provided by annotators, $\{q^e\}$, and the probabilistic vectors of emotions associated with quotes from the artists, $\{\rho^e\}$. The red distribution provides a benchmark in which these quotes are randomly allocated across paintings. The textual classification relies on a transformer model (DistilBERT).

as we have limited observable information about (i) the reasoning underlying annotators' choices and (ii) the intentions and mental states of artists at the time of production. In panel (a) of Figure C8, we show that the textual descriptions provided by annotators are more closely matched to their assigned painting than to a randomly selected painting by the same artist (panel a). In panel (b), we further quantify the correspondence between our predicted emotion vectors, $\{q^e\}$, and probabilistic emotion vectors derived from selected artists' statements about their own work, denoted $\{\rho^e\}$.

C.2 The identification of material representation(s)

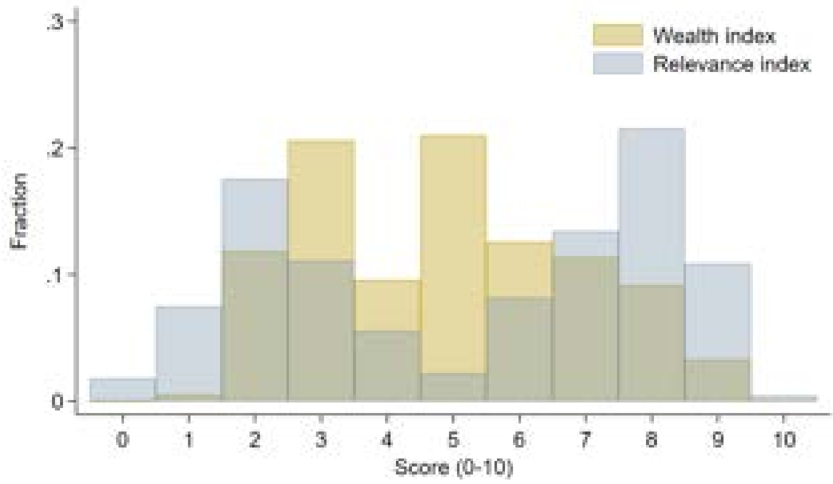
The calibration through a large language model The identification of material representation(s) in paintings relies on a calibration step by a state-of-the-art large language model (GPT-5-mini). We use a prompt, listing the different requirements as follows:

- task: assess whether the painting contains visual indicators of the era's living standards; specifically signs of wealth or poverty.
- requirements: output only a valid Python dictionary literal. No explanations, no code fences, no extra text; dictionary keys: 'wealth', 'relevance', 'elements'.
- 'wealth': int in $\{0, ..., 10\}$ where 0 is extreme poverty and 10 is extreme wealth.
- 'relevance': int in $\{0, ..., 10\}$ assessing the informativeness of such an assessment.

- 'elements': simple, concise and plain-text list of depicted elements indicating living standards.

The output of this procedure is a distribution of 627,369 wealth indices w_j (average: 4.7, standard deviation: 2.1), relevance scores (typically bi-modally distributed with many uninformative paintings, see Figure C9), and lists of elements (comprising between 4 and 10 items, where an item is "well-groomed hair and complexion" for instance).

Figure C9. Indicators of material representation(s).



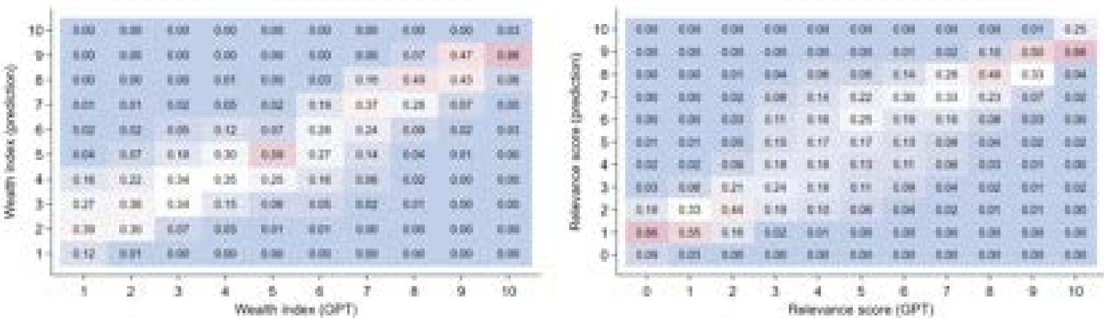
Notes: This figure displays the wealth score (in gold) and relevance score (in blue), as provided by the state-of-the-art large language model (GPT-5-mini). One can see that relevance is bi-modal: many paintings are not informative about living standards, but when they do, they are perceived as very relevant.

We analyze these elements systematically and find that the large language model identifies poverty through: the simplicity of a setting, attire, or interior (simple, sparse, bare, plain, barren, rough, basic, empty, lack, absence); the wear of objects, clothes, or houses (worn, tattered, patched, ragged, ruined, broken, debris, shelter); the depiction of death, suffering, or illness; and hardship related to labor (manual, rural). The model identifies wealth through: the refinement of a setting, attire, or interior (ornate, fine, elaborate, carved, refined, patterned, intricate, polished, delicate); the display of social status (wig, posture); the representation of personal signs of wealth (gold, lace, jewel, necklace, crown, silk, brooch, drapery, fur, embroidery, satin); and furniture or architecture. This characterization allows us to construct a corpus of features and adjectives, associated with weights, that we use to interpret the model's output.

The semantic categories used to map visual features into living-standards measures are not imposed ex ante, but refined using the empirical frequency distribution of descriptive terms generated by the model itself. In practice, we consider four groupings, two domains within each grouping, and two categorizations for each domain: a poverty-leaning corpus and a wealth-leaning corpus. The four groupings are: 0. Uninformative, 1. Objects, 2. People, and 3. Context. The eight domains are: intrinsically neutral words (0.a); words requiring additional context to be informative (0.b); decorative intensity, ornamentation, material quality, and abundance (1.a); luxury goods, portable wealth, furniture, and interior (1.b); presentation, grooming, expression, and posture (2.a); physical markers of social status (2.b); scene-level interpretation and social positioning (3.a); and background environment and general setting (3.b).[26]

In practice, we construct a sub-index ι_j^d for each domain $d \in \mathcal{D}$ by summing positive connotations—the number of terms in the wealth-leaning corpus—and subtracting negative connotations—the number of terms in the poverty-leaning corpus—associated with painting j. Our main description-based indices aggregate these sub-indices: without differential weights, $w_j^u = \sum_{d \in \mathcal{D}} \iota_j^d$; and with weights α_d derived from regressing the large-language-model wealth index (w_j) on the sub-indices, $w_j^w = \sum_{d \in \mathcal{D}} \alpha_d \iota_j^d$.

Figure C10. Predicting material representation(s).



(a) Wealth (b) Relevance

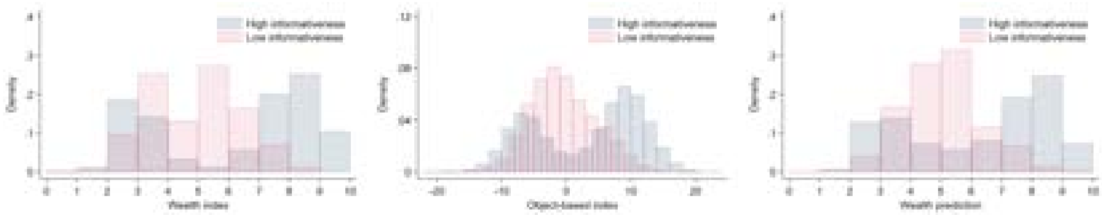
Notes: Panel **(a)** reports a correspondence matrix between the targeted wealth index w_j and the predicted index $\widehat{w}_j$ within the test sample (15% of the overall sample). Panel **(b)** replicates the exercise for the targeted and predicted relevance indices ($z_j, \widehat{z}_j$).

[26] Below a small selection of terms associated to each informative domain: (+) wealth-leaning corpus, (-) poverty-leaning corpus. **1.a** decorative intensity, ornamentation, material quality, abundance: (+) decorative, ornate, patterned, luxurious, details, embroidered, fine, polished, well-maintained, white, clean, grand, multiple, rich; (-) plain, simple, modest, muted, rough, worn, weathered, unadorned, tattered, sparse, bare, minimal, utilitarian, absence, lack, empty. **1.b** luxury goods, portable wealth, furniture, and interior: (+) silk, drapery, gold, jewelry, lace, flowers; (-) baskets, bundles, candle, box, earthenware, bowl, stool, straw, brick. **2.a** presentation, grooming, expression, and posture: (+) well-groomed/dress/kept, tailored, manicured, fashionable, elegant, calm, confident, relaxed, reclining, seated, healthy; (-) emaciated, gaunt, thin, barefoot, distressed, naked, exposed, bent, tired. **2.b** physical markers of social status: (+) wig, tie, collar, cravat, gown, uniform, sash, epaulette, regalia, scepter; (-) headscarf, cap, cane, apron. **3.a** scene-level interpretation and social positioning: (+) leisure, entertainment, attendants, servants, hunting, birds, cherubs, formal; (-) domestic, peasants, manual, hardship, fishing, carrying, crowded, cramped, dense. **3.b** background environment and general setting: (+) architecture, castle, manor, church, estate, garden, spacious; (-) rural, fields, unpaved, agricultural, communal, earthy, rocky, industrial, shelter, smoke.

A convolutional neural network architecture To predict material representation in artwork, we use the *structure* of our sentiment prediction architecture (Section 3.1) and adapt it to jointly target the wealth index (a), w_j, and the relevance score (b), z_j, on the WikiArt sample of artworks.

While predicting living standards may seem challenging relative to the prediction of emotional scores (for instance, given the role of contextual clues), the task is made easier by: the lower dimensionality of the targeted vector, and its distributional properties (with fewer "rare" occurrences). We develop and optimize the convolutional neural network in a similar manner as our sentiment-extraction algorithm and present in Figure C10 the mapping between our predictions and the targeted indices in the test sample (15% of the overall dataset). The algorithm performs well on average, with the caveat that it underestimates the likelihood of very high and very low wealth indices and relevance scores—an issue commonly observed when predicting low-probability events. To quantify the model fit, the probability for the wealth prediction, $\widehat{w}_j$, to be between 8 and 9, conditional on $w_j = 9$, is 0.90; and the probability for the wealth prediction, $\widehat{w}_j$, to be between 2 and 4, conditional on $w_j = 2$, is 0.88. Along the same lines, the probability for the relevance prediction, $\widehat{z}_j$, to be between 8 and 9, conditional on $z_j = 9$, is 0.83; and the probability for the relevance prediction, $\widehat{z}_j$, to be between 1 and 2, conditional on $z_j = 1$, is 0.88. The full set of such conditional probabilities is provided in Figure C10.

Figure C11. Indicators of material representation(s).



(a) Wealth index (b) Description-based index (c) Wealth prediction

Notes: Panel **(a)** displays the distribution of the wealth index w_j (from 0 to 10) across all our paintings. The blue distribution is computed for the 30% of artworks identified as being highly informative (score above 8); the red distribution is computed across the less informative paintings (score below 7). Panel **(b)** displays the distribution of the unweighted description-based index, w_j^u. Panel **(c)** displays the distribution of the wealth prediction, $\widehat{w}_j$.

A set of wealth indices Overall, the outcome of our procedure is a set of wealth indices (see Figure C11): (i) the wealth index produced by the state-of-the-art large language model, which will be used for validation purposes in Section 3.3 and Appendix C.5; (ii) description-based indices constructed from wealth- and poverty-leaning corpora, which will be used for interpretive purposes in Appendix C.5; and (iii) an in-house wealth prediction calibrated to replicate the output of the large language model, which will be used to produce representations of material living standards across the largest sample (see Section 4.4).

C.3 Attrition, commercialization, and selection into conservation

Our main data sources, described in Appendix B.1, may be subject to attrition bias, driven not only by selection into these digital collections, but also by selection into preservation or conservation. For instance, one might expect more valuable artworks to be more likely to survive aggregate shocks and idiosyncratic events (e.g., inheritance). A potential concern for our analysis is that such selection might be systematically correlated with the emotional or material signals conveyed through paintings, such that the average "color" of preserved artwork reflects the preferences of a later audience. To explore this possibility, this section draws on digitized sales catalogs from the Getty Research Institute to shed light on the role of attrition and commercialization.

Figure C12. Selection of artwork and selling prices.

Notes: This figure is based only on artwork that can be associated with an artist present in the sales catalogs and displays the average emotion probabilities for artwork attributed to an above-median artist (in lighter green, with an average selling price above the median of their contemporaries) versus a below-median artist (in darker green, with an average selling price below the median of their contemporaries).

The selection of preserved artwork While both our collection and the sales catalogs contain information on the artwork (e.g., title and physical characteristics), the data in the sales catalogs are too unstructured and incomplete to allow for reliable matching *at the artwork level*. For instance, a year of production is rarely available, and artworks are often described informally rather than by a formal title—titles that are often assigned retrospectively by curators or owners. Given these limitations, we instead link the two data sources using artist identity. We clean artist names in both datasets and apply a fuzzy matching procedure. Approximately 45% of artworks featured in the sales catalogs were produced by artists also present in our sample—a share that rises to 51% for transactions with a disclosed price.

This bridge across the two data sources allows us to examine two sources of selection: (i) the selection of artwork into commercialization, and (ii) the selection of artwork into preservation or conservation. For (i), we compare artworks in our dataset based on whether the artist also appears in the sales catalogs. For (ii), we examine differences in transaction values between "preserved" artists (those included in our data) and "omitted" artists (those not covered).

Another, possibly more important, question concerns selection into preservation or conservation: which artworks, produced at a given time and with a given "value," are preserved to the present day and ultimately selected for inclusion in digital collections? To address this, we proceed in two steps.

Table C1. Selection of artwork into preservation/conservation: estimating a value gap.

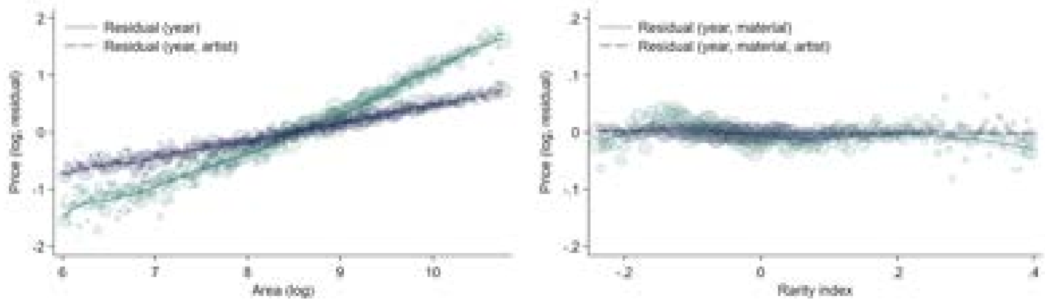
Price (log)	(1)	(2)	(3)
In our sample	0.2062	0.2027	0.1060
	(0.0457)	(0.0453)	(0.0276)
Fixed effects (month)	No	Yes	No
Fixed effects (auction)	No	No	Yes
Observations	509,424	509,424	509,424

Notes: A unit of observation is a transaction within the sales catalogs digitized by the Getty Research Institute (part of the so-called *Getty Provenance Index*), for which we observe: (i) a price and currency, (ii) the artist and their country of origin, and (iii) a consistent link between pieces by the same artist within the sales catalogs. The explaining variable is a dummy equal to 1 if the artist is present in our main sample of artwork (described in Appendix B.1), and standard errors are clustered at the artist level. In column **(1)**, we add fixed effects at the level of the currency, the country (of the artist), and the year. In column **(2)**, we further include a set of month fixed effects. In column **(3)**, we additionally include a set of auction fixed effects to identify a premium from pieces sold during the same auction.

In a first step, we focus on preserved artworks whose creators appear in the sales catalogs and ask: do highly valued artists produce different kinds of works than their peers? Figure C12 displays the average emotion probabilities for artworks attributed to artists with an average selling price above (lighter green) or below (darker green) the median price for their contemporaries. The differences in conveyed emotions between these groups are minimal. Although not entirely comparable—sales may feature the same artwork multiple times and may occur long after the time of production—the number of auctions within a given context is roughly four times higher than the number of preserved paintings in our collection.

In a second step (see Appendix Table C1), to quantify the selection of artwork into preservation or conservation, we ask: how much more valuable were preserved artists at the time of production? We focus on a subsample of transactions from our digitized sales catalogs that include consistent information on price, currency, and artist, and compare prices for artists represented in our main digitized, preserved collection of paintings versus "omitted" artists. In the most stringent specification (column 3), we identify the preservation premium by comparing pieces sold during the same auction and find it to be around 10%. This moderate premium may suggest that the disproportionate value present-day societies now place on notable artists was not fully reflected in their market value at the time of production.

Figure C13. Pricing of artwork, material, and emotional signals—sensitivity analysis (1/2).



(a) Material

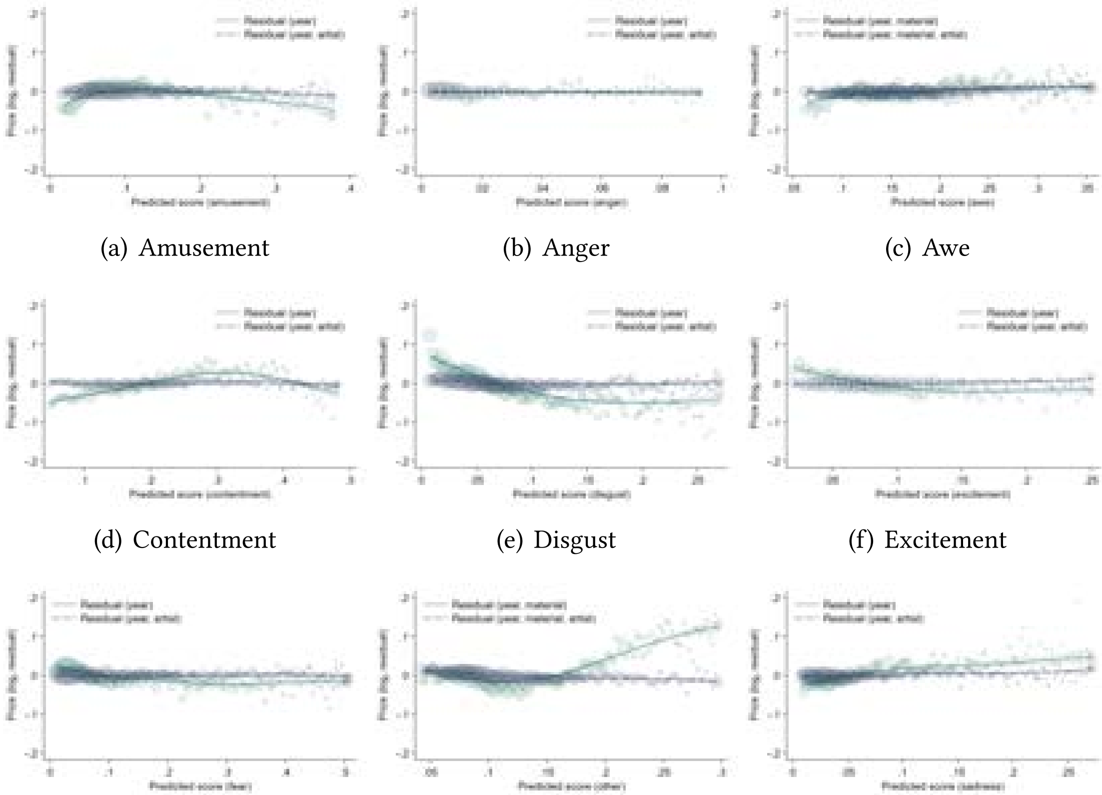
(b) Rarity index (sentiment)

Notes: These figures provide statistics about the valuation of artwork within *Artsy*. Panel **(a)** displays the correlation between the (log, residualized) price and the (log) dimension of the object—height multiplied by width in centimeters. The green plain curve residualizes the (log) price by yearly dummies; the blue dashed curve residualizes the (log) price by yearly dummies and artist fixed effects. Panel **(b)** displays the correlation between the (log, residualized) price and a sentiment rarity index derived from predicted emotional scores, $1/E \sum_e p^e/\bar{p}^e - 1$—equal to 0 if the artwork has the same emotional profile as the average artwork. In both cases, the green plain curve residualizes the (log) price by yearly dummies and object characteristics (height, width, material); the blue dashed curve residualizes the (log) price by yearly dummies, object characteristics, and artist fixed effects.

Pricing of artwork, material, and emotional signals In Section 3.3, we briefly discuss how the pricing of artwork may correlate with their average emotional signal or with the display of extreme emotions. We do so exploiting our modern dataset of paintings from *Artsy*—a sample for which we can match images with auction prices.

Figures C13 and C14 complement the evidence provided in Section 3.3 by showing the correlation between artwork prices and: (i) the dimensions of the artwork (Figure C13, panel a); (ii) an emotion rarity index, $1/E \sum_e p^e/\bar{p}^e - 1$, equal to 0 if the artwork has the same emotional profile as the average artwork (Figure C13, panel b); and (iii) all separate emotional scores (Figure C14). We do find that 10% larger paintings command a 2-4% higher price; the emotional content of a painting does not seem, however, to be

Figure C14. Pricing of artwork, material, and emotional signals—sensitivity analysis (2/2).



(a) Amusement (b) Anger (c) Awe

(d) Contentment (e) Disgust (f) Excitement

(g) Fear (h) Other (i) Sadness

Notes: All panels display the correlation between the (log, residualized) price and predicted emotional scores. The green plain curve residualizes the (log) price by yearly dummies and object characteristics (height, width, medium such as oil or gouache, support such as canvas or paper); the blue dashed curve residualizes the (log) price by yearly dummies, object characteristics, and artist fixed effects.

systematically correlated with its price. The latter observation is especially true when conditioning on the artist identity.

C.4 The mapping between well-being and conveyed emotions

Section 3.3 discusses the *general* relationship between societal sentiment and emotions expressed in artwork. In this Appendix, we complement that analysis by mapping satisfaction measures to each of the nine conveyed emotions, considering alternative welfare measures, and providing a within-context identification of this relationship—leveraging artist characteristics and reported happiness among their peers.

The mapping with the nine emotion scores Specifically, we reproduce Table 1 and replace the positivity index—constructed by summing the positive emotion scores associated with artwork j (contentment, excitement, amusement, awe) and subtracting

Table C2. The mapping between well-being and *all* conveyed emotions.

Panel A: Negative emotions	Fear	Sadness	Anger
Level of satisfaction (0-1, Gallup)	-0.11986	-0.00661	-0.01114
	(0.02864)	(0.01516)	(0.00454)
	[-0.260]	[-0.032]	[-0.181]
Observations	90,451	90,451	90,451
Panel B: Positive emotions	Contentment	Amusement	Excitement
Level of satisfaction (0-1, Gallup)	0.07368	0.03470	0.03075
	(0.02434)	(0.01814)	(0.01286)
	[0.197]	[0.127]	[0.162]
Observations	90,451	90,451	90,451
Panel C: Other emotions	Disgust	Other	Awe
Level of satisfaction (0-1, Gallup)	-0.03212	-0.00977	0.04038
	(0.01481)	(0.01116)	(0.01537)
	[-0.153]	[-0.046]	[0.175]
Observations	90,451	90,451	90,451

Notes: Standard errors are reported between parentheses and are clustered at the level of a country/year. Standardized effects are reported in square brackets. Each artwork is weighted to account for the probability that the artist is from this country—this probability is high with an average of 93.5%—and to ensure that each country/year contributes the same—thus down-weighting countries with many producers. The dependent variables are the emotional probability scores, and the explaining variable is the average satisfaction of the artist cohort—gender and age "decade"—within the country of the artist in the year of production (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025). The satisfaction variable in Gallup, originally between 0 and 10, is divided by 10 to be on a 0-1 scale. The set of controls include (log) GDP per capita, the (log) price of the artwork, and dummies for the material (e.g., oil gouache, ink) and support (e.g., canvas, paper, silk). The set of fixed effects include year and country fixed effects.

the negative ones (fear, anger, sadness, disgust), with the full set of individual emotion probability scores (Table C3). Fear, contentment, anger, and awe emerge as the most responsive emotions, followed by excitement and disgust—all with standardized effects above 0.15 in absolute value. While all estimated correlations have the expected sign, sadness appears surprisingly muted—possibly reflecting the fact that modern artwork is substantially less likely to convey this emotion than the baseline corpus, which spans periods in which the expression of grief was more socially accepted (and poverty was more prevalent across all countries).

Table C3. Alternative measures of satisfaction.

Panel A: Baseline	0-1	Bottom-25%	Top-25%	Median
Satisfaction with life	0.34924	-0.11151	0.20707	0.13368
	(0.08215)	(0.04455)	(0.05545)	(0.04342)
	[0.271]	[-0.068]	[0.110]	[0.115]
Observations	90,451	90,451	90,451	90,451
	Expected in	Living	Satisfaction with ...	
Panel B: Other measures	5 years	Standards	income	food
Satisfaction	0.18166	0.09507	0.15760	0.09573
	(0.06277)	(0.04405)	(0.08532)	(0.04998)
	[0.124]	[0.122]	[0.120]	[0.139]
Observations	90,454	90,346	90,078	90,372
Panel C: Experienced	Joy	Pain	Sadness	Worry
Experienced	0.13781	-0.08717	-0.03569	-0.05914
	(0.06013)	(0.04857)	(0.05213)	(0.04848)
	[0.122]	[-0.070]	[-0.025]	[-0.051]
Observations	90,453	90,453	90,453	90,453

Notes: Standard errors are reported between parentheses and are clustered at the level of a country/year. Standardized effects are reported in square brackets. Each artwork is weighted to account for the probability that the artist is from this country—this probability is high with an average of 93.5%—and to ensure that each country/year contributes the same—thus down-weighting countries with many producers. The dependent variable is the sentiment positivity index derived from predicted emotional scores, and the explaining variables are average satisfaction measures reported by the artist cohort—gender and age "decade"—within the country of the artist in the year of production (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025). In Panel **A**, these measures are the baseline measure (life satisfaction between 0 and 1), the share of respondents in the bottom-25% of all respondents within a given country across all years, the share of respondents in the top-25% of all respondents within a given country across all years, and the share of respondents above median. In Panel **B**, the satisfaction measures are: expected satisfaction in 5 years, satisfaction about current living standards, satisfaction about current household income, and ability to afford basic food. In Panel **C**, the satisfaction measures are: whether respondents experienced joy, pain, sadness or stress and worry the day prior to the interview. The set of controls include (log) GDP per capita, the (log) price of the artwork, and dummies for the material (e.g., oil gouache, ink) and support (e.g., canvas, paper, silk). The set of fixed effects include year and country fixed effects.

Alternative measures of well-being Table C3 reproduces Table 1 using alternative measures of reported welfare. These include transformations of our baseline life satisfaction measure—specifically, the share of respondents falling within a given percentile of the country-specific distribution across all years (panel A); expected life satisfaction in five years, satisfaction with current living standards, satisfaction with current household income, and the ability to afford basic food (panel B); and indicators for whether respondents experienced joy, pain, sadness, stress, or worry on the day prior to the interview (panel C). Our baseline measure—life satisfaction on a 0-1 scale—exhibits the strongest correspondence with conveyed positivity in artwork.

Table C4. The mapping between conveyed and reported emotions: the examples of worry and joy.

Panel A: Worry	Fear	Sadness	Anger
Worried	0.02524	-0.00850	0.00281
	(0.01599)	(0.00737)	(0.00270)
	[0.060]	[-0.045]	[0.050]
Observations	90,453	90,453	90,453
Panel B: Joy	Contentment	Amusement	Excitement
Joyful	0.04611	0.02122	0.01076
	(0.01591)	(0.01519)	(0.00793)
	[0.141]	[0.089]	[0.065]
Observations	90,453	90,453	90,453

Notes: Standard errors are reported between parentheses and are clustered at the level of a country/year. Standardized effects are reported in square brackets. Each artwork is weighted to account for the probability that the artist is from this country—this probability is high with an average of 93.5%—and to ensure that each country/year contributes the same—thus down-weighting countries with many producers. The dependent variables are the emotional probability scores (negative ones in panel **A**, positive ones in panel **B**), and the explaining variable is the probability to have experienced worry (**A**) or joy (**B**) for the artist cohort—gender and age "decade"—within the country of the artist in the year of production (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025). The set of controls include (log) GDP per capita, the (log) price of the artwork, and dummies for the material (e.g., oil gouache, ink) and support (e.g., canvas, paper, silk). The set of fixed effects include year and country fixed effects.

Types of reported and conveyed emotions In Appendix Table C4, we instead report correlations between reported emotions of different types—worry in panel A and joy in panel B—and our predicted emotion scores (negative emotions for worry and positive emotions for joy). We find, in particular, that reported worry is most strongly correlated with fear, whereas reported joy primarily correlates with contentment. Note, however, that these correlations are weaker than those obtained with life satisfaction on a 0-1 scale: reported worry or joy is indeed inferred in Gallup from very recent experience ("did you experience [emotion] yesterday?").

Relationship within a given context Finally, Appendix Table C5 exploits our ability to match Gallup respondents and artists within the same country, year, gender, and age cohort. Note that we define a cohort as an age decade, grouping, for instance, all individuals between 20 and 29 years old. In this setting, we identify a correlation between reported happiness and artwork-extracted emotional signals within a given country-year, relying on variation across cohorts and gender. The relationship remains evident

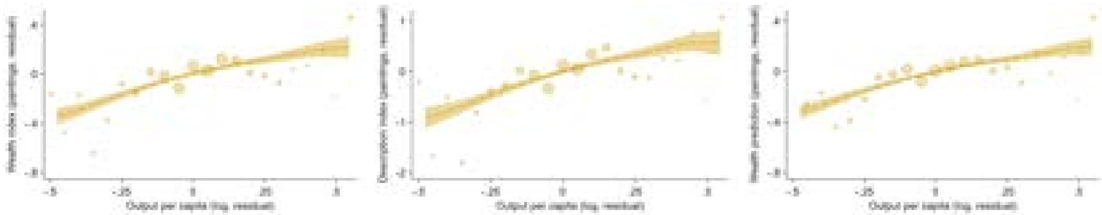
Table C5. The mapping between conveyed and reported emotions within context.

Positivity index	(1)	(2)
Level of satisfaction (0-1, Gallup)	0.2639	0.2700
	(0.1287)	(0.1288)
	[0.205]	[0.210]
Observations	90,421	90,421

Notes: Standard errors are reported between parentheses and are clustered at the level of a country/year. Standardized effects are reported in square brackets. Each artwork is weighted to account for the probability that the artist is from this country—this probability is high with an average of 93.5%—and to ensure that each country/year contributes the same—thus down-weighting countries with many producers. The dependent variable is the sentiment positivity index derived from predicted emotional scores, and the explaining variable is the average satisfaction of the artist cohort—gender and age "decade"—within the country of the artist in the year of production (2,753,803 interviews, about 15,000 per country, see Helliwell et al., 2025). The satisfaction variable in Gallup, originally between 0 and 10, is divided by 10 to be on a 0-1 scale. The set of controls include (log) GDP per capita, the (log) price of the artwork, and dummies for the material (e.g., oil gouache, ink) and support (e.g., canvas, paper, silk). The set of fixed effects include context fixed effects and fixed effects accounting for whether the artist's age and/or their gender were identified (column **2**).

within contexts: for the sake of illustration, it means that countries in which a given cohort (e.g., women in their thirties) reports lower happiness also tend to produce female artists in their thirties whose work conveys more negative emotions. The standardized effects remain large, around 0.20–0.21, albeit less precisely estimated than in our baseline specification of Section 3.3.

Figure C15. Visual correlation(s) between material representation(s) and actual living standards.



(a) Wealth index (b) Descriptions (c) Wealth prediction

Notes: Panels **(a)**, **(b)** and **(c)** display the correlations reported in Table 2 (panels A, B, and D, column 2). We residualize the wealth- and description-based indices using year, country and artist age fixed effects; group observations by bins of residualized (log) GDP per capita (Bolt and Van Zanden, 2025); and report the average indices within bins.

C.5 Material representations as a measure of living standards

Section 3.3 and Table 2 document the pass-through between actual living standards and their material representations in paintings. Panels (a), (b), and (c) of Figure C15 provide a visual counterpart to the correlations reported in Table 2 (panels A, B, and D, column 2), based respectively on the wealth-based index w_j, the description-based index w_j^u, and our in-house prediction $\widehat{w}_j$.

Living standards and specific visual elements depicted in artworks In this Appendix, we provide a more explicit link between living standards and the specific visual elements depicted in artworks. To this end, we construct corpora based on the most frequent words identified by the algorithm as describing poverty or wealth, and classify them into the following categories: (1.a Objects) decorative intensity, ornamentation, material quality, and abundance; (1.b Objects) luxury goods, portable wealth, furniture, and interiors; (2.a People) presentation, grooming, expression, and posture; (2.b People) physical markers of social status; (3.a Context) scene-level interpretation and social positioning; and (3.b Context) background environment.

Table C6. Material representations as a measure of living standards—description-based sub-indices ι_j^d.

	0. Neutral		1. Objects	
Panel A	a. Unrelated	b. No directions	a. Quality	b. Luxury
Output per capita	0.029	0.063	0.584	0.172
	(0.030)	(0.083)	(0.207)	(0.110)
	[0.047]	[0.045]	[0.234]	[0.128]
Observations	81,393	81,393	81,393	81,393
	2. People		**3. Context**	
Panel B	a. Appearance	b. Status	a. Role	b. Setting
Output per capita	0.170	0.125	0.283	0.065
	(0.042)	(0.023)	(0.040)	(0.049)
	[0.288]	[0.194]	[0.290]	[0.098]
Observations	81,393	81,393	81,393	81,393

Notes: Standard errors are reported between parentheses and are clustered at the level of a period/country, where a period spans 50 years. Standardized effects are reported in square brackets. Each specification is weighted by a score from 0 to 1 capturing how informative a painting is of living standards. The explained variables are sub-indices ι_j^d, for each of the nine domains $d \in \mathcal{D}$, obtained by summing all positive connotations—the number of terms within the wealth-leaning corpus—and subtracting all negative connotations—the number of terms within the poverty-leaning corpus—associated to painting j. These domains are: (1.a Objects) decorative intensity, ornamentation, material quality, abundance; (1.b Objects) luxury goods, portable wealth, furniture, and interior; (2.a People) presentation, grooming, expression, and posture; (2.b People) physical markers of social status; (3.a Context) scene-level interpretation and social positioning; (3.b Context) background environment. The dependent variable is a recently revised estimate of (log) GDP per capita (Maddison Project Database, see Bolt and Van Zanden, 2025). The set of baseline controls include year, country and artist age fixed effects.

Table C6 shows that the elements most predictive of living standards in paintings include:

- displays of simplicity, low-quality objects, ragged clothing, and broken items—and, conversely, refined objects, clothing, or furniture (1.a, with a standardized effect around 0.23);

- depictions of physical weakness, hardship, or exposed body parts—and, conversely, healthy, confident, and well-groomed individuals (2.a, with a standardized effect around 0.29);
- physical markers of social status (2.b, with a standardized effect around 0.19); and
- contextual clues about social positioning, such as a character begging or a figure surrounded by servants (3.a, with a standardized effect around 0.29).

For instance, a 10% higher output per capita is associated with a 6 percentage point higher probability to detect object quality in artwork descriptions (such as "refined", "delicate" or "patterned"). By contrast, the presence of luxury items (1.b) or features of the general environment or background (3.b) are less predictive.

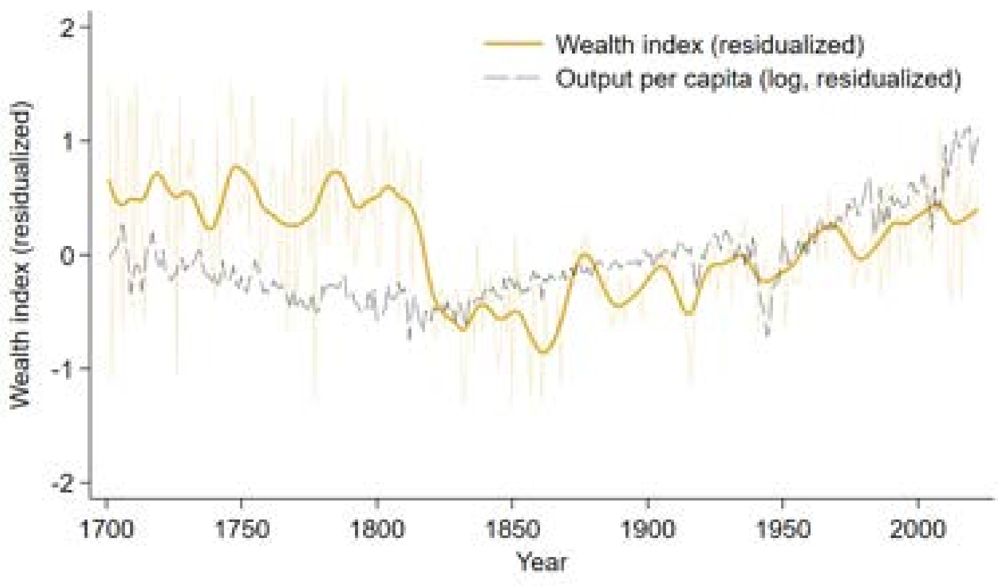
Table C7. Material representations as a measure of living standards—the role of painting genre/exercise.

Sample selection	Abstract	Relig./Myth.	Portrait	Genre
Output per capita	0.500	1.123	0.586	0.605
	(0.362)	(0.386)	(0.221)	(0.159)
	[0.326]	[0.733]	[0.382]	[0.395]
Observations	10,336	4,960	18,297	11,503
Sample selection	Still	Landscape	Sketch	Other
Output per capita	0.052	0.209	-0.447	1.105
	(0.249)	(0.329)	(0.542)	(0.432)
	[0.034]	[0.136]	[-0.292]	[0.721]
Observations	5,209	14,654	2,066	9,313

Notes: Standard errors are reported between parentheses and are clustered at the level of a period/country, where a period spans 50 years. Standardized effects are reported in square brackets. Each specification is weighted by a score from 0 to 1 capturing (i) how informative a painting is of living standards and (ii) possible duplicates across data sources. The dependent variable is (log) GDP per capita (Maddison Project Database, see Bolt and Van Zanden, 2025), and the explaining variables is the wealth index w_j from 0 to 10. The set of baseline controls include all emotion probabilities and year, country, artist age, and exercise fixed effects. The different columns restrict the sample to: abstract art; religious, mythological, symbolic, allegorical art; portrait, self-portrait, figurative, nude, caricature; genre painting, history painting, literary painting; still life, flower painting, animal painting; landscape, cityscape, marina; sketch and study; and the rest (e.g., miniature, calligraphy)

The role of painting genre As acknowledged in our derivation of relevance scores and weighting choices, not all artworks are equally informative about material living standards. In Table C7, we replicate the analysis of Section 3.3 and Table 2 within specific painting genres or exercises: abstract art; religious, mythological, symbolic, and allegorical art; portraiture (including self-portraits, figurative works, nudes, and caricatures); genre, history, and literary painting; still life, flower, and animal painting; landscape, cityscape, and marina; sketches and studies; and a residual category (e.g., miniatures, calligraphy). We find that several genres are informative—religious art, mythological art, portraits, and genre scenes—whereas others, such as landscapes, abstract art, or sketches, are substantially less so. The fact that representations of material living standards in religious, allegorical, or mythological art do correlate with contemporaneous living standards may illustrate several mechanisms: anachronisms—in which limited knowledge of past societies implies that artists depict scenes that are representative of their own environment; a translation exercise, in order to facilitate understanding by the public or patrons; or a voluntary social critique, in which painters would use biblical or mythological scenes to comment on inequality or hardship of their own societies.

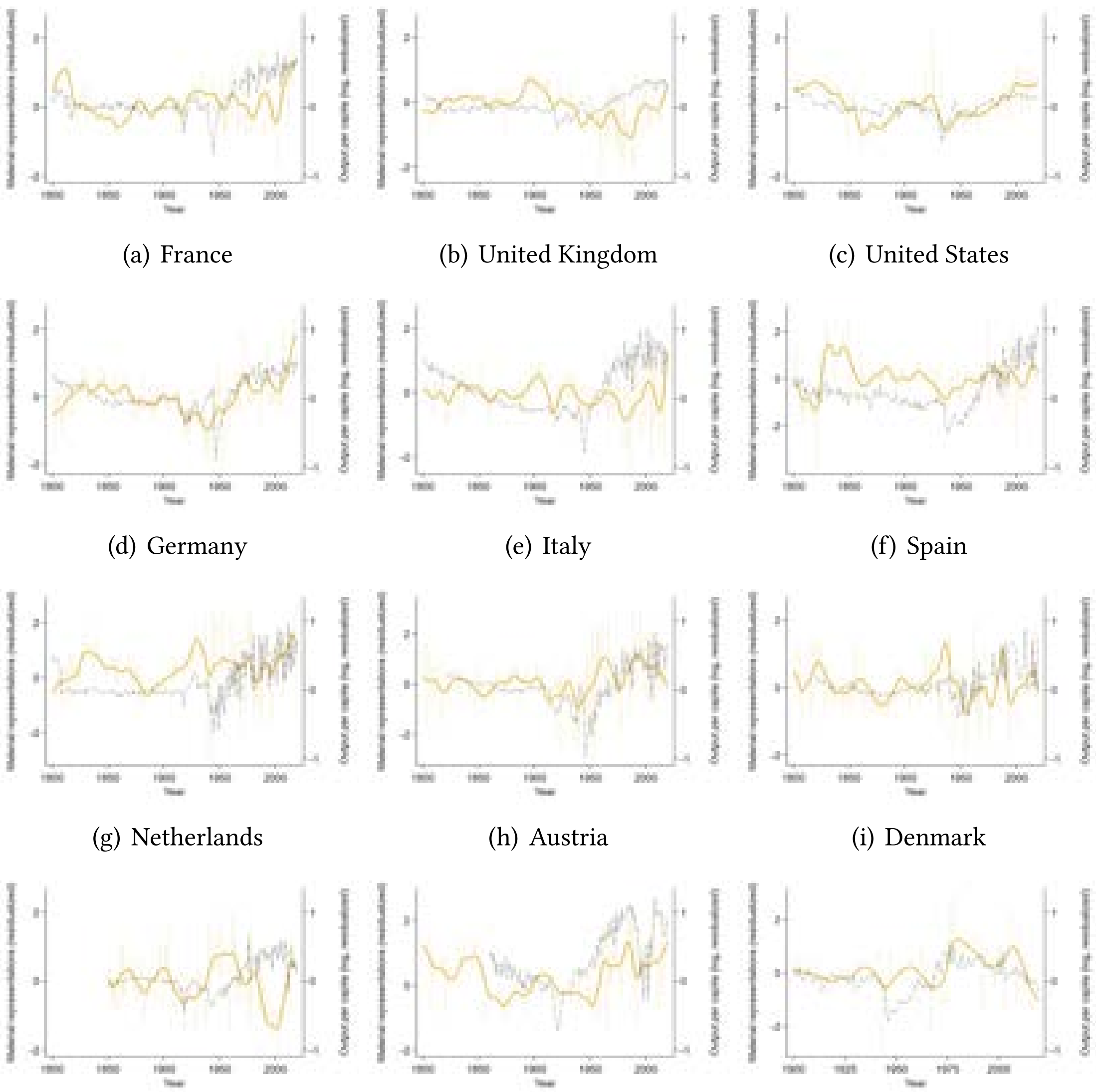
Figure C16. The dynamics of predicted living standards in France.



Notes: This figure aggregates the artwork-level measures at the yearly level to show the dynamics of the wealth-based index, w_j, in artwork over time. For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a smoothing parameter of 100—as a thicker line.

The dynamics of predicted living standards Finally, Figure C16 and Figure C17 show the dynamics of material representations in artwork. Specifically, Figure C16 aggregates artwork-level measures at the context level to illustrate the evolution of these extracted indices over time in France over the past two centuries. One can see that there are significant short- to medium-run fluctuations, and that there are longer-horizon trends. One such trend is the prominence of wealth in earlier art, possibly due to patronage, to artist selection, and to their much more exclusive audience. This observation justifies the inclusion of artist fixed-effects in most of our analysis, e.g., within our main specification (2).

Figure C17. The dynamics of predicted living standards over time.



(a) France (b) United Kingdom (c) United States

(d) Germany (e) Italy (f) Spain

(g) Netherlands (h) Austria (i) Denmark

(j) Belgium (k) Russia (l) Japan

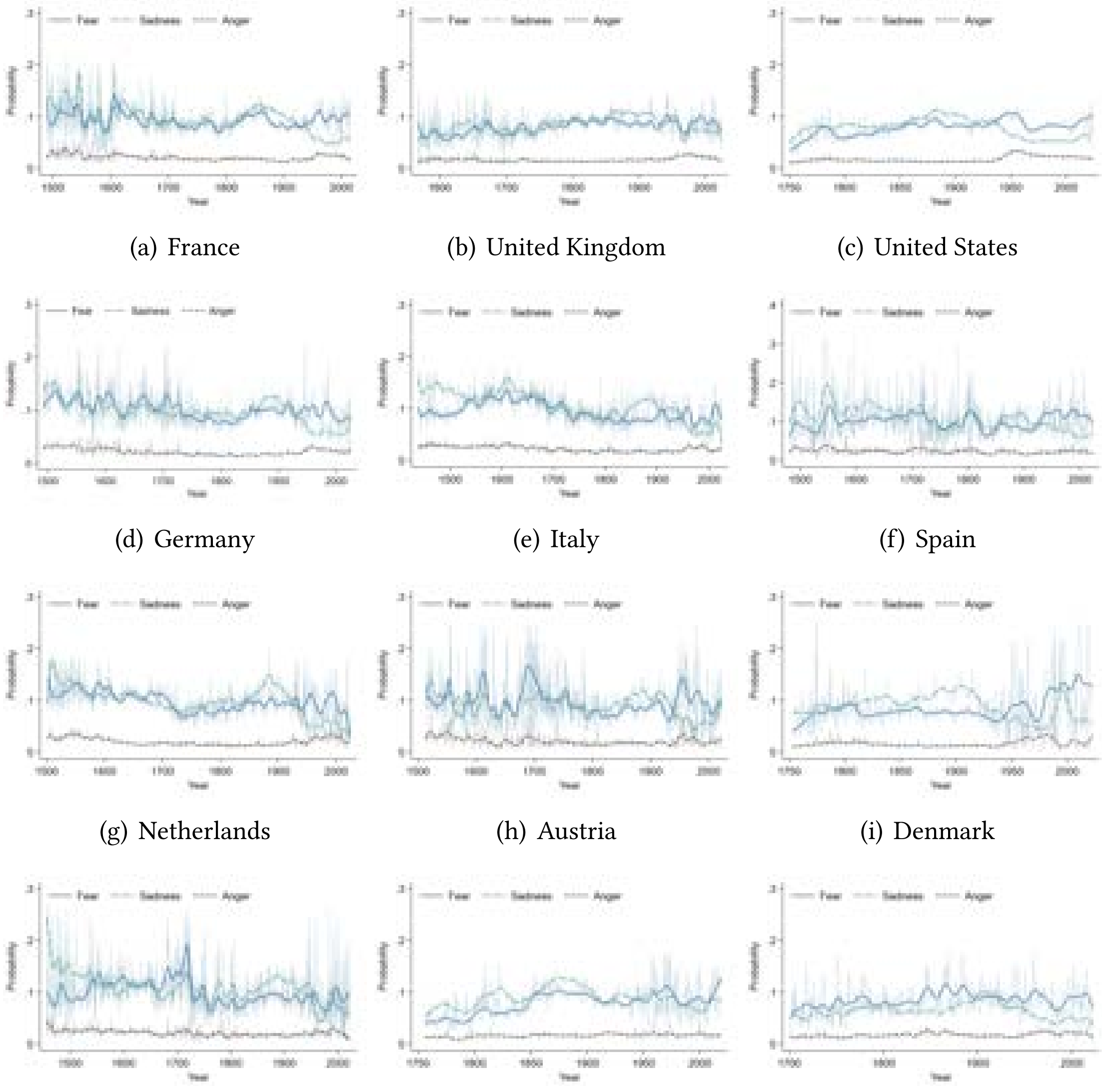
Notes: This figure illustrates the dynamics of the wealth prediction, $\widehat{w}_{\ell,t}$, in France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a coefficient of 40—as a thicker line. The thin dark line represents the (log) GDP per capita (Bolt and Van Zanden, 2025). Both measures are residualized—and detrended—by genre and artist age fixed effects.

In Figure C17, we reproduce this exercise within the present-day borders of twelve countries ℓ that have contributed more than 10,000 artworks: France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. A few countries show very tight relationships between depicted and actual living standards (e.g., Germany, the United States, Austria, or Japan); the relationship is weaker in other settings (e.g., the Netherlands).

D Emotional and material welfare through socioeconomic transformation

This section provides complements to Section 4.

Figure D1. The dynamics of negative emotions (fear, sadness, and anger) over time.



(a) France (b) United Kingdom (c) United States

(d) Germany (e) Italy (f) Spain

(g) Netherlands (h) Austria (i) Denmark

(j) Belgium (k) Russia (l) Japan

Notes: This figure illustrates the dynamics of the predicted probability to evoke emotion e, $p^{e}_{\ell,t}$, for negative emotions in France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. Negative emotions are: *fear* (blue, plain), *sadness* (teal, long dash), and *anger* (black, shorter dash). For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a coefficient of 100—as a thicker line. Note that the respective starting dates are set to exclude the first percentile of artworks in terms of production time, and each observation within a year is weighted to account for duplicates and the overall production of each artist.

D.1 The dynamics of extracted signals across countries

Negative emotions over time Figure D1 depicts the evolution of negative emotions within the present-day borders of twelve countries, each contributing more than 10,000 artworks: France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. The selected negative emotions are fear (blue, plain), sadness (teal, long dash), and anger (black, shorter dash). As discussed in Section 4.1, the dynamics of fear display considerable short-run volatility, which appears to be context-specific—except during major global conflicts, such as the World Wars. A few countries, e.g., as France, Germany, Austria, Spain, and Belgium, exhibit more impermanent emotional patterns, in contrast to the more stable trajectories observed in the United States or the United Kingdom.

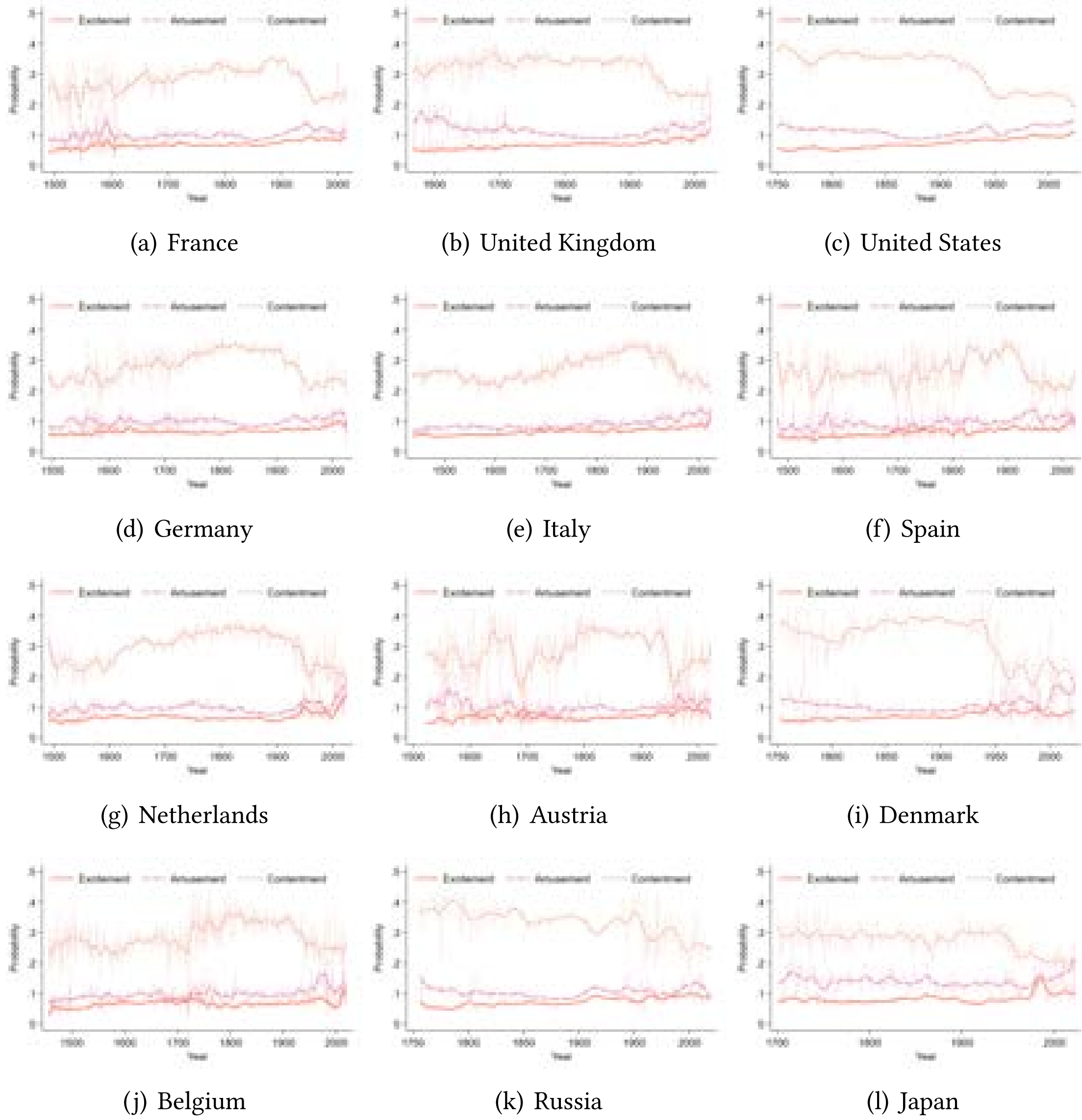
For example, Austria experienced large fluctuations in fear corresponding to the numerous conflicts affecting its territory prior to 1800—a strong correlation between the depiction of fear and the outbreak of war that appears across many societies. The evolution of sadness sometimes mirrors or follows that of fear, but not consistently. The end of the nineteenth century and the turn of the twentieth century are "sad" periods for most countries—a pattern echoed in the rise of nineteenth-century European realism—while the remainder of the twentieth century is not. One possible explanation is that the sustained rise in living standards, which began around 1800, was only broadly felt by the average individual in our sample locations after 1900, coinciding with the expansion of leisure, labor protections, social welfare, and a general decline in inequality.

Finally, anger emerges as a more residual emotion, with limited variation prior to the second half of the twentieth century.

Positive emotions over time Figure D2 depicts the evolution of positive emotions within the present-day borders of twelve countries. The selected positive emotions are excitement (orange, plain), amusement (pink, long dash), and contentment (salmon, shorter dash). Contentment and amusement display non-negligible volatility, the former also showing significant secular trends: predicted contentment steadily rises (or remains stable) from 1500–1900 and sharply decreases afterward before stabilizing from 1950–2020 (yearly variations across countries explain about 22% of the overall variance in depicted contentment, $p^c_{\ell,t}$). By contrast, excitement is less often evoked and apparently less volatile.

Other emotions over time and a positivity index Figure D3 focuses on the three remaining emotions: awe (gold, plain), disgust (green, long dash), and other (gray, shorter dash). The last two emotions follow similar dynamics across all societies: they are flat

Figure D2. The dynamics of positive emotions (excitement, amusement, and contentment) over time.



(a) France (b) United Kingdom (c) United States

(d) Germany (e) Italy (f) Spain

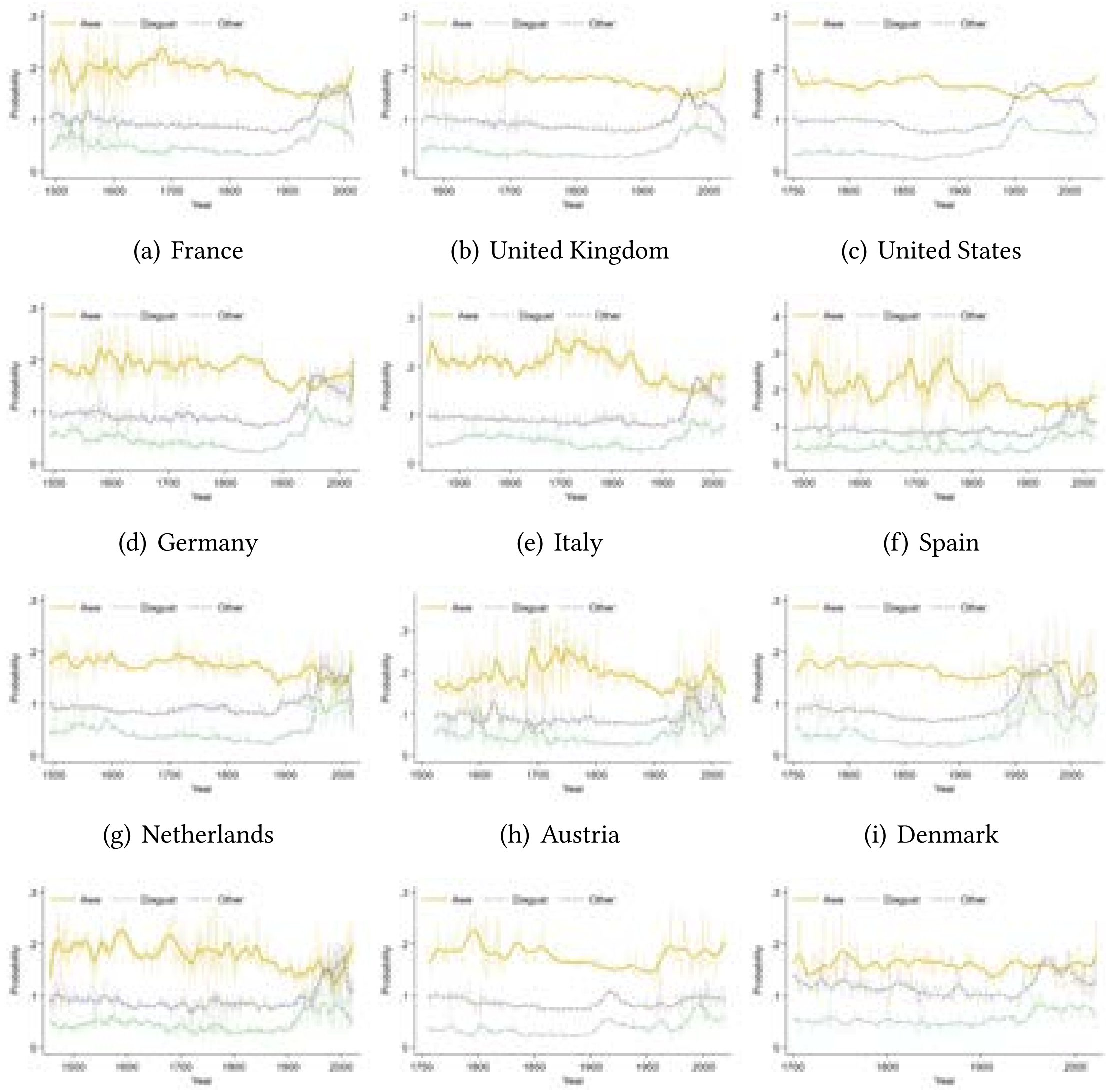
(g) Netherlands (h) Austria (i) Denmark

(j) Belgium (k) Russia (l) Japan

Notes: This figure illustrates the dynamics of the predicted probability to evoke emotion e, $p^e_{\ell,t}$, for positive emotions in France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. Positive emotions are: *excitement* (orange, plain), *amusement* (pink, long dash), and *contentment* (salmon, shorter dash). For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a coefficient of 100—as a thicker line. Note that the respective starting dates are set to exclude the first percentile of artworks in terms of production time, and each observation within a year is weighted to account for duplicates and the overall production of each artist.

and with limited volatility between 1500 and 1900 before they rise swiftly during the twentieth century. Awe, in a similar way as fear or contentment, is more volatile and with fluctuations that seem less correlated across economies. For instance, it reaches a peak during the reign of Louis XIV in France and a low point before the French Wars of Religion; it varies markedly in Spain, with local maxima around 1520 (under the rule of Charles V), 1660 (before the end of the Habsburg dynasty), and 1750 (under the Bourbon

Figure D3. The dynamics of other emotions (awe, disgust, and other) over time.



(a) France (b) United Kingdom (c) United States

(d) Germany (e) Italy (f) Spain

(g) Netherlands (h) Austria (i) Denmark

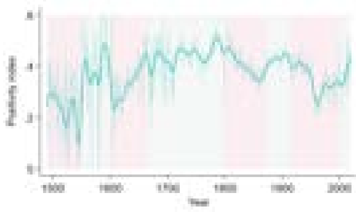
(j) Belgium (k) Russia (l) Japan

Notes: This figure illustrates the dynamics of the predicted probability to evoke emotion e, $p^{e}_{\ell,t}$, for other emotions in France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. Other emotions are: *awe* (gold, plain), *disgust* (green, long dash), and *other* (gray, shorter dash). For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a coefficient of 100—as a thicker line. Note that the respective starting dates are set to exclude the first percentile of artworks in terms of production time, and each observation within a year is weighted to account for duplicates and the overall production of each artist.

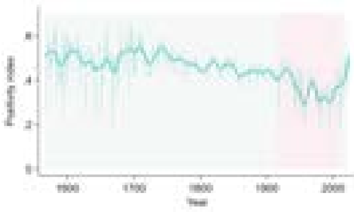
dynasty); and it increases significantly after the Second World War in the (then) Soviet Union, possibly reflecting propaganda efforts and the regime's use of cultural production for political support.

Finally, we collapse all emotion indices into a positivity index ($I_{\ell,t}$) in Figure D4 and show that: (i) there are differences in levels, with Great Britain, the United States, and Denmark being persistently more positive in earlier periods, while Italy, Spain, Germany,

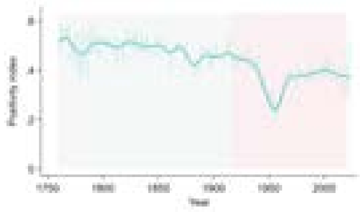
Figure D4. The dynamics of positivity ($I_{\ell,t}$) over time.



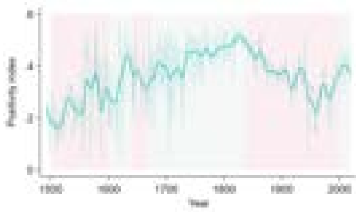
(a) France



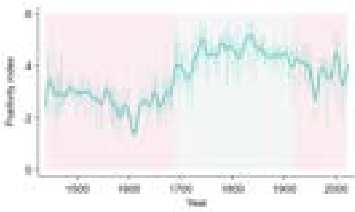
(b) United Kingdom



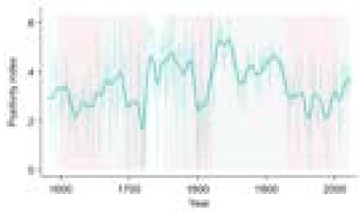
(c) United States



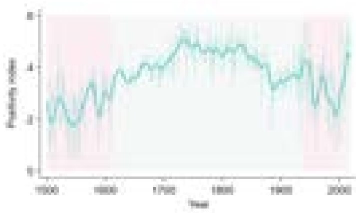
(d) Germany



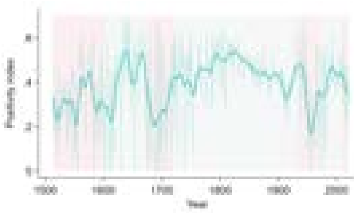
(e) Italy



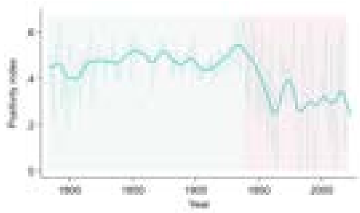
(f) Spain



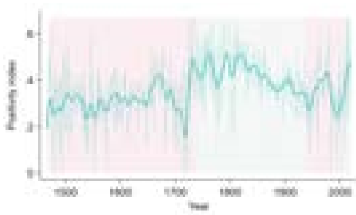
(g) Netherlands



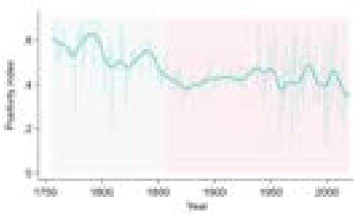
(h) Austria



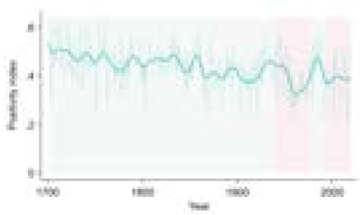
(i) Denmark



(j) Belgium



(k) Russia



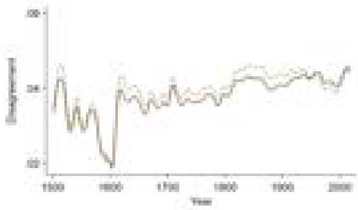
(l) Japan

Notes: This figure illustrates the average evolution of the *positivity* index ($I_{\ell,t}$) in France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. For illustration purposes, we represent the raw time series as a thin line and a smoothed time series—using a Hodrick-Prescott filter with a coefficient of 100—as a thicker line. Note that the respective starting dates are set to exclude the first percentile of artworks in terms of production time, and each observation within a year is weighted to account for duplicates and the overall production of each artist. The red/green areas indicate general periods of negativity/positivity.

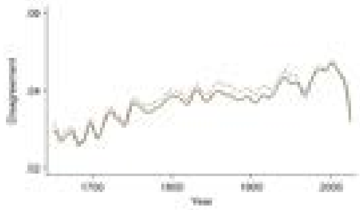
and the Low Countries are persistently more negative; (ii) there are differences in volatility, with France, Germany, Spain, and Austria displaying the most variation over time; (iii) many continental European countries experience a prolonged negative episode, possibly coinciding with the Little Ice Age; (iv) the eighteenth and nineteenth centuries are marked by a sustained period of positivity; and (v) the geopolitical context of 1914–1963 appears to correspond with a widespread decline in positivity across all countries.

Interestingly, positivity reaches its highest point around 1850—coinciding with the German revolutions of 1848–1849—within the present-day borders of Germany, followed by a long decline that only reverses after World War II. Additional spikes in positivity are also observable: in France during la Belle Époque, in the early twentieth century or after the liberation from Ottoman rule in Austria, during the Dutch Golden Age in the Netherlands, and under the reign of Nicholas I in Russia.

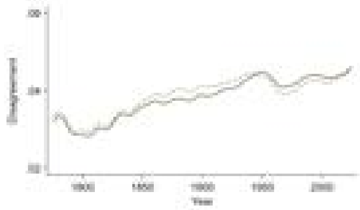
Figure D5. The dynamics of disagreement ($d_{\ell,t}$) across countries and over time.



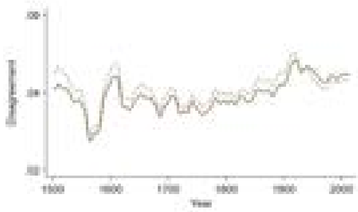
(a) France



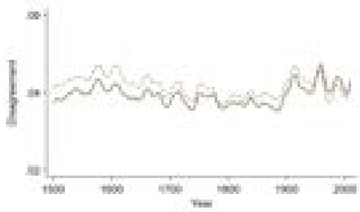
(b) United Kingdom



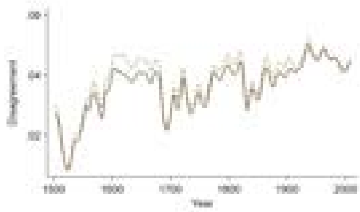
(c) United States



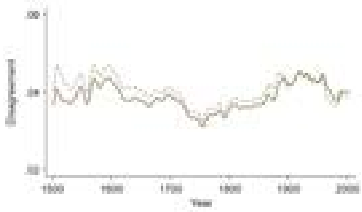
(d) Germany



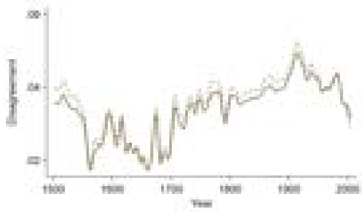
(e) Italy



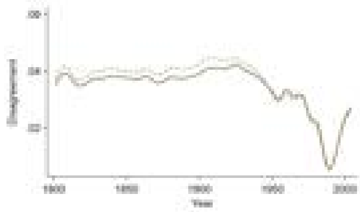
(f) Spain



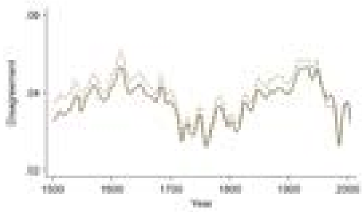
(g) Netherlands



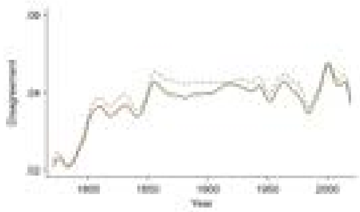
(h) Austria



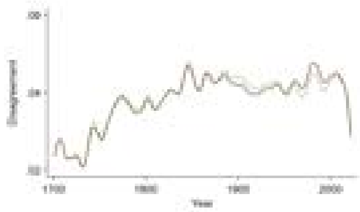
(i) Denmark



(j) Belgium



(k) Russia



(l) Japan

Notes: This figure displays the dynamics of *disagreement*, $d_{\ell,t}$, in France, the United Kingdom, the United States, Germany, Italy, Spain, the Netherlands, Austria, Denmark, Belgium, Russia, and Japan. For illustration purposes, the time series are smoothed using a Hodrick-Prescott filter with a coefficient of 100, and the respective starting dates are set to include the bulk of artworks in terms of production time. The context-level disagreement index is computed as the average level of disagreement across all paintings produced in a given year t and location ℓ, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}}|p_j^e - \bar{p}_{\ell,t}^e|$. The dashed line is based on measures that are residualized by genre, movement, age, and artist fixed effects.

The disparity in emotions within a context The previous evidence sheds light on the "commonality" of different artworks produced in a given context. In Figure D5, we illustrate the "disparity" in expressed emotions within a context, $\delta_{\ell,t}$, as captured by the distance to the average depiction within a context, $1/|\mathcal{E}| \sum_{e \in \mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|$, and its average across all paintings produced in any given year t and location ℓ, and by a similar index but based on emotion scores residualized by the identity of artists, their age, the harmonized movements, and the harmonized exercises.

One striking shift in the disparity of displayed emotions over time appears in the second half of the sixteenth century in France, corresponding to the French Wars of Religion—during which many artists fled to Great Britain, Saxony, Prussia, the Low Countries, and Switzerland.

D.2 Empirical strategy

Section 4.2 outlines our empirical strategy, which isolates variation in conveyed emotions within an artist's production cycle. In this Appendix, we first examine this residual variation and then describe a Principal Component Analysis (PCA) with two objectives: (i) to construct simplified indices associated with the probabilistic emotion vector, and (ii) to explore the latent structure underlying the probabilistic emotion vector $\{p^e\}_{e \in \mathcal{E}}$.

Explanatory power of artist and context fixed-effects Figure D6 quantifies the respective explanatory power of artist and context fixed-effects. Let j index an image produced by artist $a(j)$ in location $\ell(j)$ and year $t(j)$, and let $p_{j,\ell,t}^e$ denote the predicted probability of evoking emotion e. We consider specifications of the form,

$$p_{j,\ell,t}^e = \psi_{\ell,t} + \alpha_{a(j)} + \varepsilon_{j,\ell,t}$$

where $\psi_{\ell,t}$ are context fixed effects and $\alpha_{a(j)}$ represents unobserved artist effects. The inner black line in Figure D6 shows the adjusted R-squared from a linear regression with 100 country fixed effects; the green line corresponds to a specification with 600 year fixed effects; the blue line reflects a model with 14,384 context fixed effects; and the red line shows results from a regression with 32,000 artist fixed effects. The outer purple line conditions on both artist and context fixed effects.

The figure reveals that country differences explain less than 2-3% of the overall variance in depicted emotions, and common yearly dynamics across countries explain less than 8% for most emotions. In contrast, unobserved artist-level factors account for 40-50% of the variance. Conditioning additionally on context fixed effects (ℓ, t) increases the explained variance by a further 3-4 percentage points across our emotion indicators.

Figure D6. Explanatory power of artist and context fixed-effects.

Notes: This figure reports a variance decomposition of our probabilistic emotion scores, quantifying the respective explanatory power of artist and context fixed effects. For each regression, the unit of observation is a painting j, and the left-hand side variable is the predicted probability to evoke emotion e, p_j^e. The inner, black line corresponds to the adjusted R-squared of linear regressions with 100 country fixed effects; the green line corresponds to the adjusted R-squared of linear regressions with 600 year fixed effects; the blue line corresponds to the adjusted R-squared of linear regressions with 14,384 context fixed effects; and the red line corresponds to the adjusted R-squared of linear regressions with 32,000 artist fixed effects. The purple, outer line conditions on the two previous sets of fixed effects.

Table D1. A reduction of dimensionality.

	p^{c_1}	p^{c_2}	p^{c_3}	p^{c_4}
Eigenvalue	*2.56*	*1.91*	*1.35*	*1.10*
Difference	*0.65*	*0.56*	*0.25*	*0.48*
Awe	-0.234	-0.114	**0.613**	**0.444**
Contentment	**-0.503**	-0.174	-0.366	0.025
Amusement	-0.167	**0.502**	-0.211	-0.382
Excitement	-0.190	0.351	**0.421**	**-0.412**
Fear	**0.424**	-0.204	0.346	-0.221
Anger	**0.456**	0.143	0.033	-0.027
Sadness	0.340	**-0.418**	-0.320	-0.140
Disgust	0.323	**0.448**	-0.058	0.109
Other	0.146	0.381	-0.205	**0.637**

Notes: This table reports the loadings of the different emotion scores from the Principal Component Analysis (PCA). We present the eigenvalue associated with each eigenvector of the variance-covariance matrix in italics, along with the difference from the next lower eigenvalue.

A reduction of dimensionality In Table D1, we instead explore the *cross-emotion* variation and identify latent factors underlying our probabilistic emotion vector $\{p^e\}_{e\in\mathcal{E}}$ through a Principal Component Analysis. The first principal component ($p_j^{c_1}$) loads positively on negative emotions (fear, anger, sadness, and disgust) and is associated with a very high eigenvalue. As suggested by Figure 10, which focuses on fear, this principal component captures meaningful fluctuations around salient, often geopolitical, events.

The second component ($p_j^{c_2}$) also loads on non-neutral emotions, including amusement, excitement, as well as disgust and other, and is associated with a reasonably high eigenvalue. As shown in Appendix Figure D2, its fluctuations only partly reflect secular trends in the depiction of positive emotions, coinciding with rising living standards across developed economies.

The third component ($p_j^{c_3}$) loads predominantly on awe, while the fourth component ($p_j^{c_4}$) is primarily associated with the residual category other. An interesting feature of the data is that contentment, the most frequently expressed emotion, is not positively loaded on any of these principal components: it is negatively loaded on the first component, suggesting that artists may substitute contentment with emotions such as fear, anger, and sadness during turbulent times.

An illustration of the empirical strategy based on well-known artists In this section, we discuss our empirical strategy through the lens of the following 30 well-known artists, representative of their respective eras (in parentheses):

- Jan van Eyck (1400–1450);
- Hieronymus Bosch (1450–1500);
- Giovanni Bellini (1450–1500);
- Sandro Botticelli (1450–1500);
- Albrecht Dürer (1500–1550);
- Leonardo Da Vinci (1500–1550);
- Michelangelo Buonarroti, known as Michelangelo (1500–1550);
- Raphael Sanzio da Urbino, known as Raphael (1500–1550);
- Doménikos Theotokópoulos, known as El Greco (1550–1600);
- Pieter Bruegel The Elder (1550–1600);
- Tiziano Vecellio, known as Titian (1550–1600);
- Michelangelo Merisi da Caravaggio (1600–1650);
- Diego Velazquez (1600–1650);

- Peter Paul Rubens (1600–1650);
- Johannes Vermeer (1650–1700);
- Rembrandt Harmenszoon van Rijn, known as Rembrandt (1650–1700);
- Giovanni Battista Tiepolo (1700–1750);
- Jean-Antoine Watteau (1700–1750);
- Jacques-Louis David (1750–1800);
- Francisco Goya (1750–1800);
- Eugene Delacroix (1800–1850);
- Joseph Mallord William Turner (1800–1850);
- Claude Monet (1850–1900);
- Edouard Manet (1850–1900);
- Vincent Van Gogh (1850–1900);
- Henri Matisse (1900–1950);
- Pablo Picasso (1900–1950);
- Wassily Kandinsky (1900–1950);
- Andy Warhol (1950–2000);
- Francis Bacon (1950–2000);
- Jackson Pollock (1950–2000)

Analyzing the emotional content of these artists' work yields the following insights. First, they differ markedly in their emotional expression. Table D2 presents a systematic selection of the most expressive artists by emotion.

Appendix Figure D7 then illustrates these differences for three prominent artists, Wassily Kandinsky, Jan van Eyck, and Jackson Pollock. Panel (a) shows the average emotion scores for each artist across nine emotion categories, capturing distinct emotional profiles: Kandinsky expresses relatively high levels of amusement and excitement, van Eyck emphasizes awe and contentment, while Pollock features more anger, fear, and disgust. Panel (b) highlights the substantial variation *within* each artist's body of work

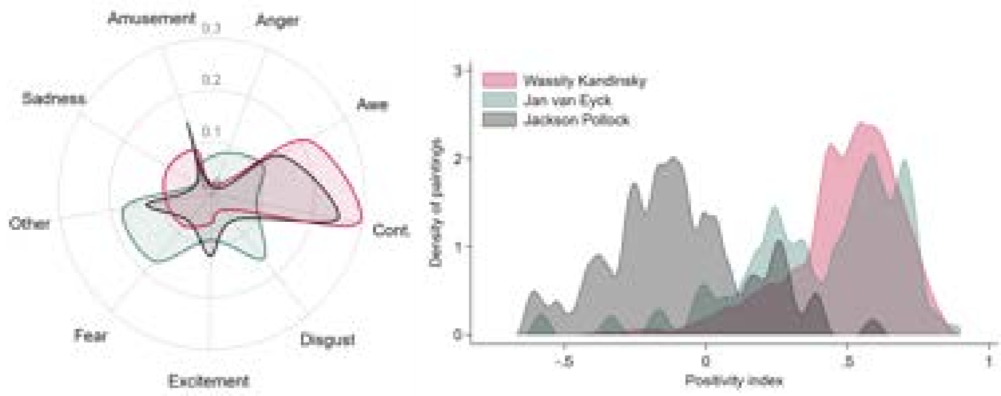
Table D2. Most expressive artists by emotion.

Emotion	Artists
Contentment	Monet, Vermeer
Amusement	Picasso, Kandinsky
Excitement	Kandinsky, Warhol
Awe	Tiepolo, van Eyck, Rubens
Fear	Bruegel the Elder, Bosch, Bacon, Pollock
Anger	Pollock
Sadness	Caravaggio, Bellini, Rembrandt, El Greco
Disgust	Pollock
Other	Pollock

Notes: This table presents a systematic selection of the most expressive artists by our nine emotion categories: *contentment*, *amusement*, *excitement*, *awe*, *fear*, *anger*, *sadness*, *disgust*, and *other*.

by plotting the distribution of the positivity index across their respective paintings. The spread of each distribution indicates that, even within a single artist's oeuvre, emotional tone can vary widely from one painting to another.

Figure D7. Emotional expressions of different artists (Kandinsky, van Eyck, and Pollock).

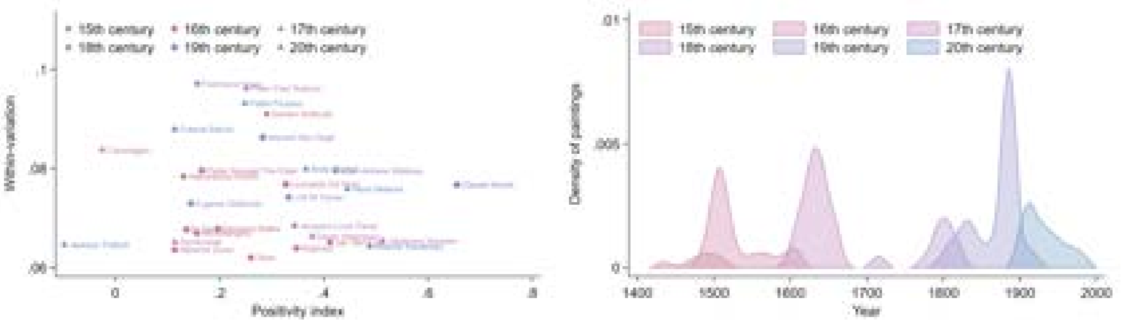


(a) Emotions (b) Within-variation

Notes: Panel **(a)** displays the distribution of predicted emotion scores for 1,025 paintings produced by Wassily Kandinsky (red), Jan van Eyck (green), and Jackson Pollock (black). Panel **(b)** displays the distribution of the *positivity* index, I_i, across the same set of 1,025 paintings.

Second, there is substantial variation in emotional expression within each artist's production cycle. Figure D8 extends the insight from the previous Figure D7 to the full sample of 30 selected artists. Panel (a) places each artist in "emotion space," using the average positivity index to summarize overall emotional tone and a within-artist disagreement index to capture the extent of emotional variability across works. Panel (b) shows the temporal coverage of the sample by century, revealing distinct clusters in artistic production over time.

Figure D8. Emotional expressions of different artists (30 selected artists).



(a) Emotions

(b) Coverage

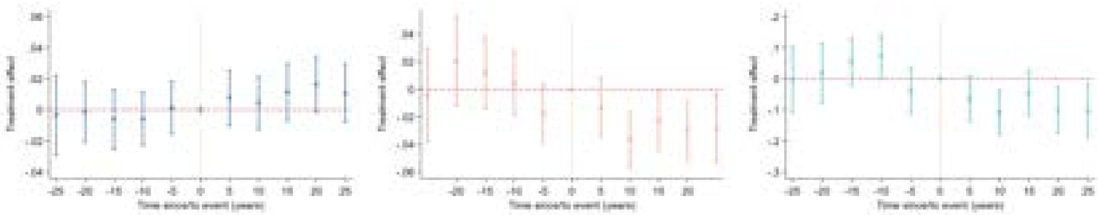
Notes: Panel **(a)** displays the degree of disagreement within artist against the average positivity index, I_i—the figure is based on 27,387 paintings produced by 30 well-known artists: Dürer, Warhol, Caravaggio, Monet, Velazquez, Manet, El Greco, Delacroix, Bacon, Goya, Tiepolo, Bellini, Matisse, Bosch, Turner, Pollock, David, van Eyck, Watteau, Vermeer, Da Vinci, Michelangelo, Picasso, Rubens, Bruegel the Elder, Raphael, Rembrandt, Botticelli, Titian, van Gogh, Kandinsky. The disagreement measure is the average of the sum (across emotions) of absolute deviations between each painting's score and the artist's average score. Note that the disagreement measure would be distributed as follows if emotion scores were randomly assigned: 0.035, 0.043, 0.064. Panel **(b)** displays the distribution over time of the paintings covered in the previous exercise.

The results highlight notable differences in average emotional tone across artists. Some, like Jan van Eyck and Johannes Vermeer, exhibit consistently high positivity, while others, such as Jackson Pollock and Caravaggio, display much lower average scores. The within-artist disagreement index further reveals meaningful disparities in emotional range across artists. While painters like van Eyck, Vermeer, and Kandinsky demonstrate relatively focused emotional expression, others, including Francisco Goya and Pablo Picasso, exhibit wide variation. In the case of Goya, this range reflects the contrast between his early, decorative works and his later, darker period marked by political violence and psychological unrest. For Picasso, the emotional diversity corresponds to his successive stylistic periods, from the Blue Period to Cubism and beyond.

Additionally, certain individual paintings diverge sharply from their creators' typical emotional profiles. For example, Claude Monet's *Still Life with Pheasants and Plovers* (1879) is unusually somber, reflecting the personal grief following the death of Camille Doncieux. Similarly, Sandro Botticelli's *Lamentation over the Dead Christ with the Saints Girolamo, Pietro, and Paolo* carries the weight of Counter-Reformation austerity, likely influenced by the spiritual severity of Girolamo Savonarola's rule in Florence.

Third, variation in emotional expression sometimes reflects broader secular changes in the artists' environments. Sandro Botticelli, Caravaggio, Titian, Pieter Bruegel the Elder, and El Greco all experienced the chastening effects of the Counter-Reformation in Italy or the rule of Philip II in the Low Countries and Spain. To quantify the impact of these contextual "shocks" on conveyed emotions, we embed them within a two-way fixed

Figure D9. The dynamics of emotions after counter-reformations (Sandro Botticelli, Pieter Bruegel the Elder, Titian, El Greco, Caravaggio).

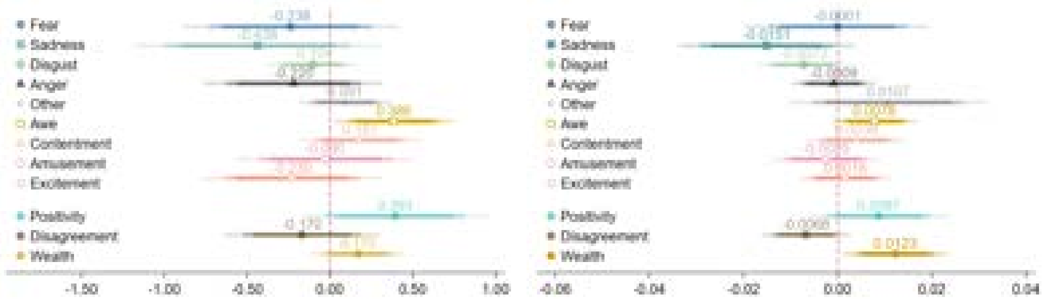


(a) Fear (b) Contentment (c) Positivity

Notes: This figure presents a collapsed two-way fixed effects event study, controlling for artist fixed effects, time fixed effects, and country fixed effects. The specification considers 5 events: Sandro Botticelli, religious extremism, 1490; Pieter Bruegel the Elder, religious persecution, 1560; Titian, Counter-Reformation, 1545; El Greco, Spanish Inquisition, 1582; Caravaggio, Counter-Reformation, 1595. For visualization purposes, we group years into the closest five-year intervals to identify the estimates β_τ.

effects event study, controlling for artist, time, and country fixed effects (see Figure D9). We find that fear increases by approximately 2 percentage points, contentment decreases by 3 percentage points, and the positivity index declines by around 0.10.

Figure D10. Alternative indicators of economic development and the emotions conveyed through paintings.



(a) A valence index (b) Life expectancy

Notes: These Figures display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to alternative indicators of economic development. We consider Equation (2), with location, year, artist, and genre fixed-effects, and we replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with: **(a)** a "valence index" using sentiment analysis on published books (Hills et al., 2019); and **(b)** a measurement of life expectancy at birth (Zijdeman and Ribeira da Silva, 2015). The coefficients represent separate regressions with the following left-hand side variables: the predicted probability of expressing emotion e, p_j^e for painting j, where e is *fear* (dark-blue, plain circle), *sadness* (light-blue, plain square), *disgust* (teal, plain diamond), *anger* (black, plain triangle), *other* (gray, cross), *awe* (gold, hollow square), *contentment* (salmon, hollow triangle), *amusement* (pink, hollow diamond), and *excitement* (orange, hollow circle); the *positivity* index, $l_j = \sum_{e\in\mathcal{E}^+} p_j^e - \sum_{e\in\mathcal{E}^-} p_j^e$, in light blue; a sum of deviations from the mean emotion in a given context, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|$; and the wealth prediction, $\widehat{w}_j$. All left-hand side variables are standardized, so the reported estimates can be interpreted in terms of standard deviations across country/year observations. The bands represent 10%, 5%, and 1% confidence intervals, and standard errors are clustered at the level of country/half-century.

Alternative indicators of economic development In this Appendix, we also contrast our own measures of welfare to alternative, sometimes more standard, measures of economic development in economic history, where and when such data are available (using data from Hills et al., 2019; Zijdeman and Ribeira da Silva, 2015). While our baseline approach focuses on measures of GDP per capita and wealth concentration (see Figure 11), we correlate the predicted probability of evoking each emotion $e \in \mathcal{E}$, p_j^e, the positivity index, I_j, the disagreement index, d_j, and the wealth prediction, $\widehat{w}_j$ with: a sentiment analysis of books written in a few countries (Hills et al., 2019); and a measure of life expectancy at birth (Zijdeman and Ribeira da Silva, 2015).

In Figure D10, one can see that sadness, the aggregate positivity index, and the wealth prediction correlate strongly with these alternative measures of living standards, in much the same way as with the baseline measures. However, compared to Figure 11, these alternative indicators appear to give more prominence to awe than to other positive emotions, possibly because their variation is concentrated in the nineteenth century.

D.3 The effect of large socio-economic transformations

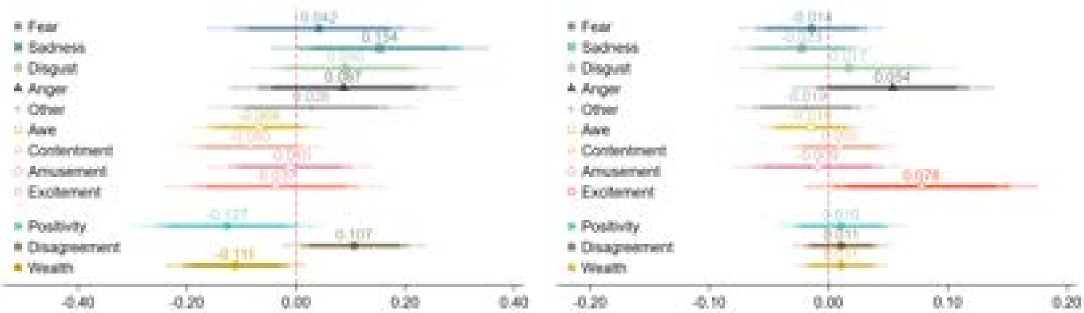
This section provides complements to Section 4.4.

Basic subsistence costs In panel (a) of Figure D11, we provide a complement to Figure 12: climatic fluctuations have very marked effect on the emotional expression of artists, possibly due to their impact on basic subsistence costs. To support this hypothesis, we regress our emotion indicators on a residual measure of bread prices in Figure D11 (a): A 50% increase in the price of bread raises the probability of evoking anger and sadness by 0.045–0.075 standard deviations, at the expense of awe and contentment—mirroring the effects of temperatures.

Luxury trade before the first wave of globalization Panel (b) of Figure D11 complements Figure 13: before the first wave of globalization, trade in luxury goods did not benefit a broad share of the population, and we observe different effects compared to the large trade expansion that followed. Specifically, an increase in the price of luxury goods is associated with heightened depictions of anger and excitement.

Universities, the decline of religiosity, and the scientific revolution The evidence reported in Section 4.4 focuses on academic openness—measured by the (log) number of foreign scholars—and the share of scientific research in total academic production. Figure D12 (panels a and b) further illustrates the large variation in the extent and nature of academic production within the present-day borders of five countries (France, Germany, Italy, Spain, and the United Kingdom) between 1400 and 1800, based on two

Figure D11. Basic subsistence costs, and luxury trade.



(a) Price of bread

(b) Price of luxury

Notes: These figures display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to local prices. We consider Equation (2), with location, year, artist, and genre fixed-effects, and we replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with: **(a)** the residualized (log) price of bread as inferred from the market prices reported in the Allen-Unger Global Commodity Prices Database (Allen and Unger, 2019, see Appendix B.3); and **(b)** the price of luxury goods. The coefficients represent separate regressions with the following left-hand side variables: the predicted probability of expressing emotion e, p_j^e for painting j, where e is *fear* (dark-blue, plain circle), *sadness* (light-blue, plain square), *disgust* (teal, plain diamond), *anger* (black, plain triangle), *other* (gray, cross), *awe* (gold, hollow square), *contentment* (salmon, hollow triangle), *amusement* (pink, hollow diamond), and *excitement* (orange, hollow circle); the *positivity* index, $l_j = \sum_{e\in\mathcal{E}^+} p_j^e - \sum_{e\in\mathcal{E}^-} p_j^e$, in light blue; a sum of deviations from the mean emotion in a given context, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|$; and the wealth prediction, $\widehat{w}_j$. All left-hand side variables are standardized, so the reported estimates can be interpreted in terms of standard deviations across country/year observations. The bands represent 10%, 5%, and 1% confidence intervals, and standard errors are clustered at the level of country/half-century.

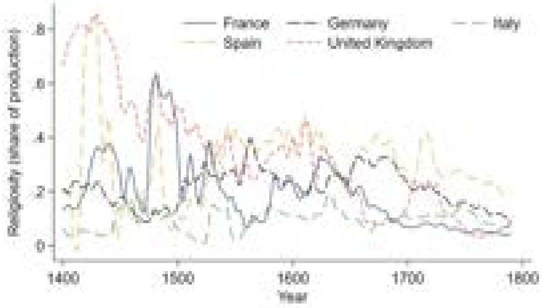
additional characteristics: research "quality"—measured by the (log) number of influential scholars—and religiosity—the share of religious research in total academic production. Across these economies, academic output increases steadily over time, though less markedly in Italy and more prominently in Great Britain. Spain is a notable exception, where academic output stagnates from 1550 onward, following the reign of Charles V, Holy Roman Emperor.

Appendix Figure D12 (c) shows that academic output is associated with a more positive emotional tone of contemporary artistic output across all countries and periods. An increase in academic output equivalent to the secular trend in the United Kingdom between 1400–1800 would be associated with -0.18 standard deviations in the depiction of fear and increases of 0.17 and 0.08 standard deviations in the depiction of contentment and amusement, respectively. Appendix Figure D12 (d) shows that the relative dominance of religiosity against scientific research in academic production yields negative emotions—particularly sadness and fear (at the expense of contentment and amusement). This emotional pattern is observed in spite of limited effects on material representation.

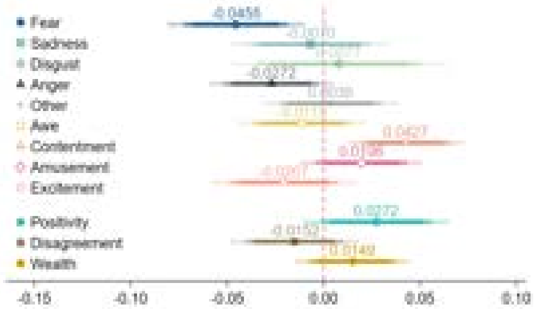
Labor strikes, emancipation, and revolutions Section D.3 of the paper focuses on objective indicators of regime characteristics (see, e.g., Figure 16). In Figure D13, we turn to a list of key events and define indicators equal to 1 in the ten years following a

Figure D12. Universities, the decline of religiosity, and the scientific revolution.

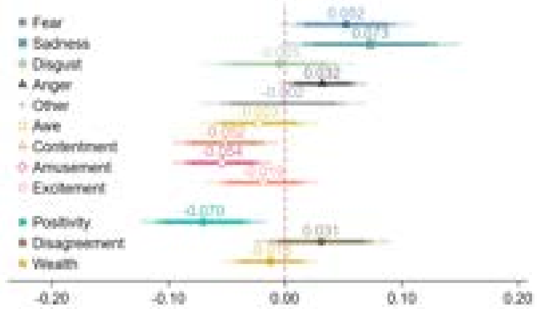
(a) Influential scholars (evol.)



(b) Religiosity (evol.)



(c) Influential scholars (effect)

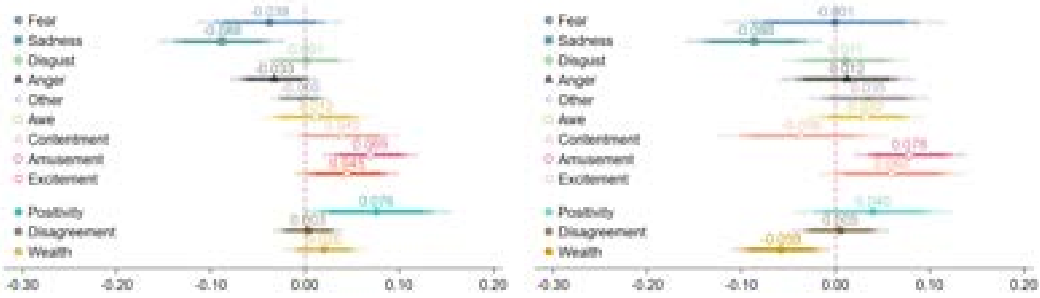


(d) Religiosity (effect)

Notes: The top panels illustrate the variation in the intensity and nature of knowledge production in France, Germany, Italy, Spain, and the United Kingdom between 1400 and 1800 (De La Croix et al., 2024). Panel **(a)** reports a measure of "quality" or influence, as measured by the (log) number of scholars with more than 10 Wikipedia pages (under different languages); and panel **(b)** reports the weighted share of knowledge production in theology. The bottom panels display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to these indicators of knowledge production. We consider Equation (2), with location, year, artist, and genre fixed-effects, and we replace the right-hand side variable $\mathbf{s}_{\ell,t}$ with: **(c)** the (log) number of scholars with more than 10 Wikipedia pages and **(d)** the weighted share of knowledge production in theology. The coefficients represent separate regressions with the following left-hand side variables: the predicted probability of expressing emotion e, p_j^e for painting j, where e is *fear* (dark-blue, plain circle), *sadness* (light-blue, plain square), *disgust* (teal, plain diamond), *anger* (black, plain triangle), *other* (gray, cross), *awe* (gold, hollow square), *contentment* (salmon, hollow triangle), *amusement* (pink, hollow diamond), and *excitement* (orange, hollow circle); the *positivity* index, $l_j = \sum_{e\in\mathcal{E}^+} p_j^e - \sum_{e\in\mathcal{E}^-} p_j^e$, in light blue; a sum of deviations from the mean emotion in a given context, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|$; and the wealth prediction, $\widehat{w}_j$. All left-hand side variables are standardized, so the reported estimates can be interpreted in terms of standard deviations across country/year observations. The bands represent 10%, 5%, and 1% confidence intervals, and standard errors are clustered at the level of country/half-century.

major reform of labor laws and slavery (panel a) or a revolution/transition to a democratic regime (panel b). The insights from Figure D13 suggest that reforms reducing labor coercion and political revolutions significantly influence the emotions conveyed in artworks—and in a generally positive direction. Following labor emancipations, the positivity index increases by 0.076 standard deviations, primarily driven by a more prevalent depiction of amusement at the expense of sadness. Instead, revolutions generate a subtle, novel emotional outlook, with more frequent depictions of amusement or excitement, and less frequent depictions of sadness (in spite of lower representations of material standards) *and* contentment.

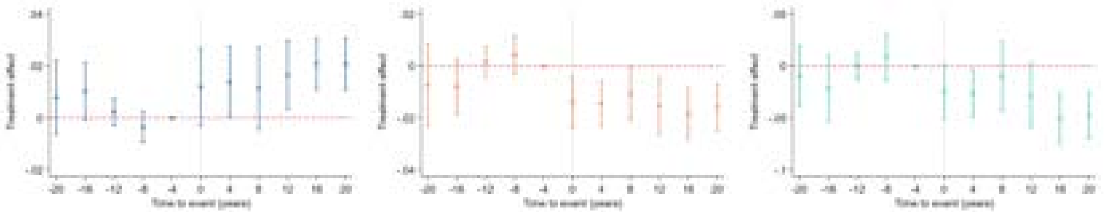
Figure D13. Labor strikes, emancipation, revolutions, and the emotions conveyed through paintings.



(a) Labor unions/emancipation (b) Revolutions

Notes: These Figures display the estimates of regressions relating the predicted emotions of a given painting, $\{p^e\}_{e\in\mathcal{E}}$, to indicators of political reforms and revolutions. We consider Equation (2), with location, year, artist, and genre fixed-effects, and we replace the right-hand side variable $s_{\ell,t}$ with: **(a)** an indicator equal to 1 in the ten years following a major reform of labor laws and slavery (e.g., in France, the Waldeck-Rousseau law recognizing the right to organize labor unions in France, May 1968, or the abolition of slavery in its colonies); and **(b)** an indicator equal to 1 in the ten years following a transition to democracy. The coefficients represent separate regressions with the following left-hand side variables: the predicted probability of expressing emotion e, p_j^e for painting j, where e is *fear* (dark-blue, plain circle), *sadness* (light-blue, plain square), *disgust* (teal, plain diamond), *anger* (black, plain triangle), *other* (gray, cross), *awe* (gold, hollow square), *contentment* (salmon, hollow triangle), *amusement* (pink, hollow diamond), and *excitement* (orange, hollow circle); the *positivity* index, $l_j = \sum_{e\in\mathcal{E}^+} p_j^e - \sum_{e\in\mathcal{E}^-} p_j^e$, in light blue; a sum of deviations from the mean emotion in a given context, $\frac{1}{|\mathcal{E}|}\sum_{e\in\mathcal{E}} |p_j^e - \bar{p}_{\ell,t}^e|$; and the wealth prediction, $\widehat{w}_j$. All left-hand side variables are standardized, so the reported estimates can be interpreted in terms of standard deviations across country/year observations. The bands represent 10%, 5%, and 1% confidence intervals, and standard errors are clustered at the level of country/half-century.

Figure D14. The dynamics of emotions during conflicts.



(a) Fear (b) Contentment (c) Positivity

Notes: This figure presents a collapsed two-way fixed effects event study, controlling for artist fixed effects, time fixed effects, and country fixed effects. Standard errors are clustered at the country level. The specification considers 24 events affecting France, Germany, Italy, the Netherlands, Spain, the United Kingdom, the United States, China, Austria, Japan, Portugal, Russia, Sweden, Switzerland, and Turkey (e.g., the French Wars of Religion, the Thirty Years' War, the War of the Austrian Succession, the Anglo-Dutch Wars, the Russo-Japanese Wars, the Russo-Turkish War (1768–1774), the English Civil War, the American Civil War, and the Sino-Japanese War). For visualization purposes, we group years into the closest five-year intervals to identify the estimates β_τ.

Event studies While our baseline approach in Equation (2) considers correlations without addressing potential dynamic effects, this appendix presents results from event studies that control for artist, time, and country fixed effects (as in Equation 2). To do so, we focus on abrupt events with clearly defined start dates and analyze 62 event-country treatments associated with 24 distinct wars. The analysis includes the following countries or current political entities: France, Germany, Italy, the Netherlands, Spain, the United Kingdom, the United States, China, Japan, Portugal, Russia, Sweden, Switzerland, and Turkey.

We use this exercise as a proof of concept: can we detect a large negative shock through shifts in the emotions conveyed by artists before and after an event? Figure D14 reports the results for three main indices: fear, contentment, and positivity. We find that fear increases by approximately 2 percentage points, contentment moves in the opposite direction, and positivity declines by about 0.05. On average, the effect is long-lasting: emotion indices do not return to baseline levels even 20 years after the event—despite a median conflict duration of around 10 years in our sample.