State of the Art: Economic Development Through the Lens of Paintings*

Clément Gorin Stephan Heblich Yanos Zylberberg

July 7, 2023

Preliminary draft

Abstract

This paper analyzes paintings to glean information about their context. Relying on a large, open-access repository of art collections, we develop an algorithm to classify the emotions conveyed through paintings. We then project these emotions onto: the characteristics of artists; their influences and styles; and their context (a location within a country, in a given year). Our main object of interest is this context-specific residual of emotions, specifically its variation across locations and over time. In a first step, we use the context-specific vector of emotions to predict economic development and political change where these measures are readily available. In a second step, we extend the prediction to cover most of Europe from the 14th century onward. Our predicted measures of economic change exhibit large fluctuations around key, well-known economic transformations of Medieval Europe, the Renaissance, the later Reformation, and the Enlightenment. The prediction however uncovers interesting, and so far overlooked, geographic variation around more localized historical events inducing significant economic uncertainty.

*Gorin: University of Toronto, email: gorinclem@gmail.com; Heblich: University of Toronto, NBER, email: stephan.heblich@utoronto.ca; Zylberberg: University of Bristol, CEPR, CESifo, email: yanos.zylberberg@bristol.ac.uk. We would like to thank Karol Borowiecki, Michael Hutter, Ömer Özak and participants are the ACEI 2023 for helpful comments. We acknowledge the support of the ORA Grant ES/V013602/1 (MAPHIS). The usual disclaimer applies.

"I never paint dreams or nightmares. I paint my own reality."—Frida Kahlo.

In spite of recent efforts to quantify economic production in history, data on income or welfare along with their spatial distribution is rare across Medieval Europe, the Renaissance, the later Reformation, and the Enlightenment.[1] There exist, however, numerous indirect depictions of these periods through preserved artwork, i.e., essays, novels, musical pieces, or paintings. Notable examples of paintings include: The Garden of Earthly Delights (Hieronymus Bosch, around 1500, see Figure 1, panel a), which depicts novel fruits and imaginary creatures at a time of exploration and discovery of the New World; Guernica (Pablo Picasso, 1937), which famously describes the horror of the Spanish Civil War; or Liberty Leading the People (Eugène Delacroix, 1830, panel b), which captures the short-lived July Revolution in France—an uprising of the middle and lower classes against the revived nobility. At times, the same event may be covered very differently by different artists from different origins. The crossing of the Alps, as a sign of the expansionary ambition of Napoleon, is both depicted by Jacques-Louis David and William Turner (see Figure 1, panels c and d). The former paints Napoleon as an allegorical figure imperiously riding a fiery steed—a strong leader guiding an impetuous country (a French viewpoint in the early years of the French Consulate). The latter shows instead the crossing by Hannibal and represents an over-ambitious leader and his army engulfed in a snow storm (an English viewpoint in the later years of the Napoleonic Wars). Can we systematically rely on such expressions as testament to their times and contexts?

This paper analyzes paintings as expressions of their times and contexts. More specifically, we exploit a large, open-access repository of art collections—Google Arts and Culture, with about 1,000,000 paintings from more than 15,000 artists—where each painting is associated to an artist, a year of production, and a location. We complement this data with the artists' biographies to better understand their movements across locations. Our empirical strategy proceeds in two steps. In a first step, we develop a neural net to classify the emotions conveyed through paintings.[2] The approach relies on a large repository

[1] Prominent contributions to our understanding of historical economic performance across regions come from Angus Maddison (Maddison, 2007) and the subsequent Maddison Project (Bolt and Van Zanden, 2014; Bolt et al., 2018), or work by Bairoch on city populations (Bairoch, 1988). These indicators have been instrumental in understanding the patterns of growth, structural change and urbanization in the very long run (see, e.g., Acemoglu et al., 2005; North, 2010; Fujita et al., 2001; Galor, 2005, 2011, discussing the role of institutions, the spatial distribution of economic activity, or the demographic transition within a unified framework). Additional, recent contributions have provided precise economic indicators covering smaller regions (see, for instance Clark, 2005, 2007, in England from 1200 onward) or a more recent period (see, for instance, Schularick and Taylor, 2012; Jacks et al., 2011, covering various macroeconomic indicators and commodity prices, respectively, across the most advanced economies in the past centuries).

[2] We pre-process our paintings before estimating the model. More specifically, the input goes through: centering; squaring; and a pixelization into 384×384×3 arrays. This pre-processing is necessary; the model cannot accommodate well any heterogeneity in the *structure* of the input.

Figure 1. A few examples.

(a) The Garden of Earthly Delights (around 1500, H. Bosch)

(b) Liberty Leading the People (1830, E. Delacroix)

(c) Napoleon Crossing the Alps (1801–1805, J.L. David)

(d) Snow Storm: Hannibal and his Army Crossing the Alps (1809–1812, W. Turner)

Notes: Panel (a) 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. Panel (b) shows Liberty Leading the People (Eugène Delacroix, 1830), which captures the short-lived July Revolution in France. Panel (c) shows one of the 5 "Napoleon Crossing the Alps" (1801-1805, J.L. David), an allegorical portrait of Bonaparte leading his army (France) throughout the col du Grand Saint-Bernard (the early challenges of the French Consulate) from a French perspective. Panel (d) is a response by William Turner showing the struggle of Hannibal in his earlier crossing (as a possible metaphor for the later struggle of Napoleon, France, and/or the long Napoleonic Wars) from an English perspective.

of annotations (more than 1,500,000 in total, see Achlioptas et al., 2021; Mohamed et al., 2022) based on 80,000 paintings in Wiki-Art (Saleh and Elgammal, 2016) and on the latest advances in image classification through Convolutional Neural Networks (CNNs). Our neural network consists of a pre-estimated encoder model extracting features from images—typically recognizing frequent objects (humans, animals, faces, etc.), shapes and their interactions—and a "regression head" mapping the extracted features into emotion scores. The former is common to most image recognition models. The latter is instrumental to our purpose; the regression head and its interaction with the usual encoder model (which is re-estimated in a sequential manner) help capture the nature of paint-

ings and how they convey emotions through their structure, composition or "texture". An image i produced by an artist in a certain style and at a given time of their career is then associated to a vector of emotions $\mathbf{E}_i$ by the estimated model.[3]

In a second step, we project these emotions onto: the characteristics of artists (e.g., their age and their gender), their influences and styles; and their context (a location l within a country, and in a given year t). Our main object of interest is this context-specific residual—cleaned from the "artist fixed effect"—, which we interpolate to provide a smooth time-series of emotions, $\mathbf{e}_{lt}$, across about 100 locations in about 20 (mostly European) countries from the 14th century onward. A context might however be more than just a location at a given point in time: artists might be confronted to very different living standards, depending on their socio-economic status, their gender or their religion/ethnic group. One possible extension of our method could produce different context-specific residuals along these dimensions to better understand the geography of inequalities and its evolution in the longer run.

The geography described by residual emotions, $\mathbf{e}_{lt}$, sheds light on the economic development, inequalities and structural economic changes experienced by those locations. First, we show the predictive power of the residual $\mathbf{e}_{lt}$ by correlating the vector with measures of political turmoil, but also, and more importantly, measures of income levels, economic uncertainty, structural transformation and inequality where/when readily available (using, among others, data from Schularick and Taylor, 2012; Boix et al., 2013; Alvaredo et al., 2020, covering the nineteenth century onward in a few developed economies). We find that our emotion indices do capture the level of economic development, but also more subtle differences across environments, e.g., the uncertainty induced by rapid, structural changes like new trade opportunities or new technologies with unclear long-term implications.

Second, we (will) use the resulting mapping from paintings to predict political turmoil and economic change across locations of Europe from the 14th century onward. Our predicted economic indicators exhibit large expected fluctuations around key, well-known economic transformations of Europe (e.g., the Napoleonic Wars, the Reformation, or the Russian Revolution) and show the heterogeneous impact of those major economic shocks across space. However, the predictions also uncover major, and so far overlooked,

[3]The annotations/emotions available in Achlioptas et al. (2021) and in Mohamed et al. (2022) are the following: Amusement; Anger; Awe; Contentment; Disgust; Excitement; Fear; Sadness; and Other. We however reduce the dimensionality of the problem by considering the principal components associated with the predicted vectors of emotions. We uncover four principal components with sufficient explanatory power in predicting the observed variation in predicted emotions (in parentheses, we report the emotions with the highest loadings): PC 1 (fear, anger, sadness); PC 2 (excitement, amusement); PC 3 (awe); PC 4 (other). Note that the emotion "Other" is typically used by annotators for abstract paintings or stylistic exercises.

local variation as induced by other, less salient historical events.

Our exercise relates to recent work inferring economic development through indirect data—images in particular. This includes work on nighttime luminosity and other satellite imagery (Overman et al., 2006; Chen and Nordhaus, 2011; Henderson et al., 2012; Jean et al., 2016; Donaldson and Storeygard, 2016); and Google-Street views (Gebru et al., 2017; Naik et al., 2017a). Our procedure relies on the structure of deep convolutional neural networks that is used to extract objects, texture and their relationship, but also to classify images (in the form of paintings). Our main contribution is to provide a technology that infers the measurement of economic activity *in the distant past*, rather than improving the measurement of contemporary economic activity (e.g., at a more disaggregated level, or in environments with poor data coverage).

Our research also relates to the literature using arts and creative work in economics. One strand of research in this vein studies the life of artists in an attempt to understand what drives their productivity and creativity. For instance, Borowiecki (2017) analyzes 1,400 letters exchanged between the composers Beethoven, Mozart and Liszt to generate a measure of well-being which turns out to be a positive driver of work-related engagements and accomplishments. In another paper on composers of Western music, Borowiecki (2022) studies how teachers and mentors influence their students' creative output. A second strand of research employs artwork to better understand the importance of protecting novel ideas, concepts, or artwork (Giorcelli and Moser, 2020; Whitaker and Kräussl, 2020); and finally, there is a strand of literature that studies markets for artwork (Ginsburgh and Jeanfils, 1995; Spaenjers et al., 2015). We contribute to this literature by highlighting another impact of art on our understanding of the economy—through the (preserved) message conveyed by artists themselves.

Finally, our work indirectly relates to the recent literature discussing the measurement of economic activity as recently surveyed in Hulten and Nakamura (2022). The Gross Domestic Product (GDP) became the standard way of measuring economic success, because it is based on a simple accounting and allows policy makers to monitor economic development and benchmark/evaluate policies. Recent discussions have emerged to better account for aggregate welfare, e.g., the Human Development Index. Our contribution to this discussion is small, but we show that the emotions conveyed through paintings predict important economic indicators beyond the level of economic development: uncertainty (Bloom, 2014; Jurado et al., 2015), or economic inequality.

The rest of this preliminary draft is structured as follows. Section 1 introduces the data and the classification method to produce a vector of emotions associated with each painting. Section 2 develops the empirical strategy to map measures of emotions at the image level onto aggregate measures at the context (a location l and a given year t) level.

Section 3 sheds light on the geography of emotions over time.

1 An algorithm to detect emotions

This section describes our data and an algorithm to map paintings into emotions.

1.1 A collection of images and annotations

We exploit: (i) a large, open-access repository of art collections where each painting is associated to an artist, an estimated year of production and a location; and (ii) a smaller repository of paintings with annotated/tagged emotions. We provide some descriptive statistics about these data sources in Appendix A.

A large collection of images Our main dataset is a novel repository of art collections, Google Arts and Culture, with about 1,000,000 paintings from more than 15,000 artists.[4] All pieces come with a description that includes tags and the Wikipedia page of the artist when existing. We retrieve information from Google Arts and Culture using a headless browser which downloads images at the highest precision/zoom level, associates tags to paintings (e.g., production year, style or movement), retrieves Wikipedia pages of the artists and requests information about birth date and location, death, etc.[5] We use a pre-trained Named Entity Recognition model to distinguish location tags from the other tags and the museum meta-data.

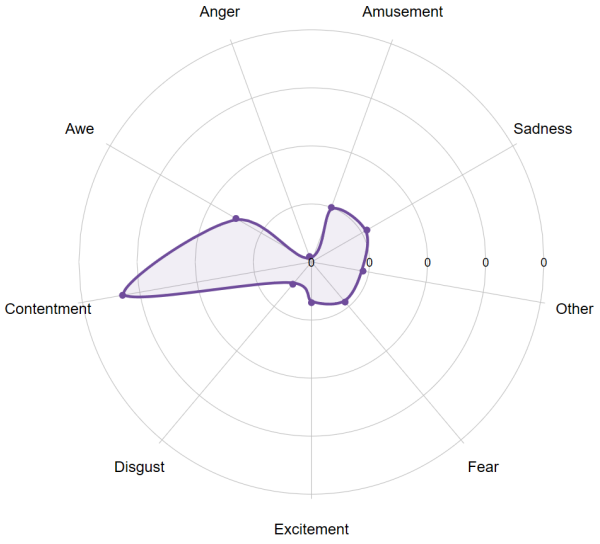
Harmonizing names, periods and places The previous data provides high-quality information about the time of an artist's birth and death, and time intervals for the production of each painting. The location information is however less organized, especially so for all events characterizing their life but their birth and death. We combine information from various sources about their life, collect locations and clean the latter information (city and country in particular) with entity recognition models, allowing us to geocode them with the Google Maps API. One of our specifications indeed hinges on movements across contexts from the same individuals to isolate a context-specific vector of emotions (see Section 2).

[4] See https://artsandculture.google.com/.

[5] We complement this data source with tags associated to more than 500,000 paintings in ART500K (Mao et al., 2017). The tags usually cover the basic information already present in Google Arts and Culture or Wiki-Art (e.g., artist, genre, art movement), but it also provides crucial information about the date and location information associated to paintings (our "context", l, t). See https://deepart.hkust.edu.hk/ART500K/art500k.html.

Annotations and emotions Our two main data sources for annotated emotions both rely on another open-access collection, Wiki-Art (Saleh and Elgammal, 2016), providing reasonably high-quality images of paintings together with some information about the piece itself and the artist—when correctly identified. Our extracted sample from Wiki-Art consists of about 80,000 paintings dating back to the early 12th century for the oldest piece. The website also classifies each piece into 80-90 art movements. We further collect additional information about the paintings (e.g., style, date, location) and the artists (e.g., name, nationality, place and date of birth, death, etc.) from the Wiki-Art website.[6]

Figure 2. The average emotion score.



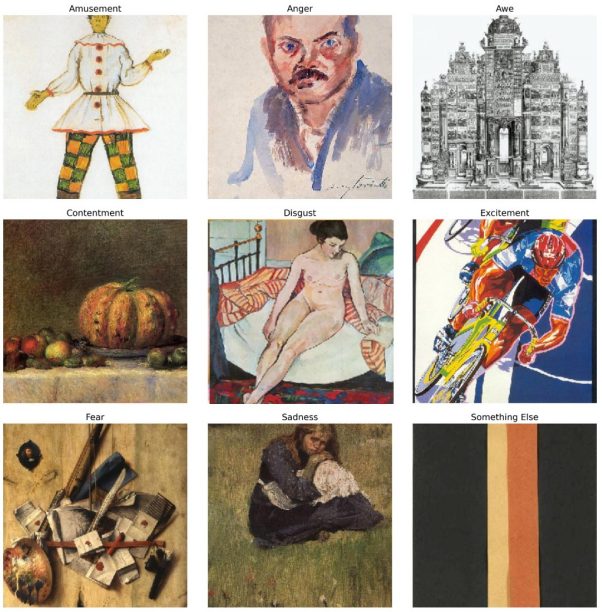
Notes: This Figure shows the distribution of emotion scores over 80,000 labeled paintings. The respective scores are (standard deviations are in parentheses): Amusement, 0.10 (0.11); Anger, 0.01 (0.04); Awe, 0.15 (0.12); Contentment, 0.33 (0.21); Disgust, 0.05 (0.08); Excitement, 0.07 (0.09); Fear, 0.09 (0.14); Sadness: 0.11 (0.15); and Other, 0.09 (0.09).

Our model requires labeled data, i.e., a sub-sample of paintings with associated labels (e.g., a certain emotion intensity), to map a painting i onto a vector of emotion $\mathbf{E}_i$. We rely on ArtEmis (Achlioptas et al., 2021, providing about 450,000 labels for the 80,000 paintings discussed above) and ArtELingo (Mohamed et al., 2022, providing more than 1,000,000 labels for the same 80,000 paintings). These two sources collect emotion scores along 9 dimensions: *amusement*, *anger*, *awe*, *contentment*, *disgust*, *excitement*, *fear*, *sadness* and *other*. In practice, individuals were shown paintings and had to select one

[6]See https://www.wikiart.org/. Note that our final sample excludes a few recent images due to copyright issues.

emotion among these 9 possibilities; ArtEmis has 5.6 annotations per painting on average and ArtELingo has 15.3 labels (thus corresponding to different evaluators). We treat these several labels as a probabilistic evaluation from a unique individual. Each painting will be allocated a vector of scores: the shares of evaluators having mentioned each of the nine emotions, which we normalize such that the scores sum up to 1 for each painting.

Figure 3. The highest emotion scores.



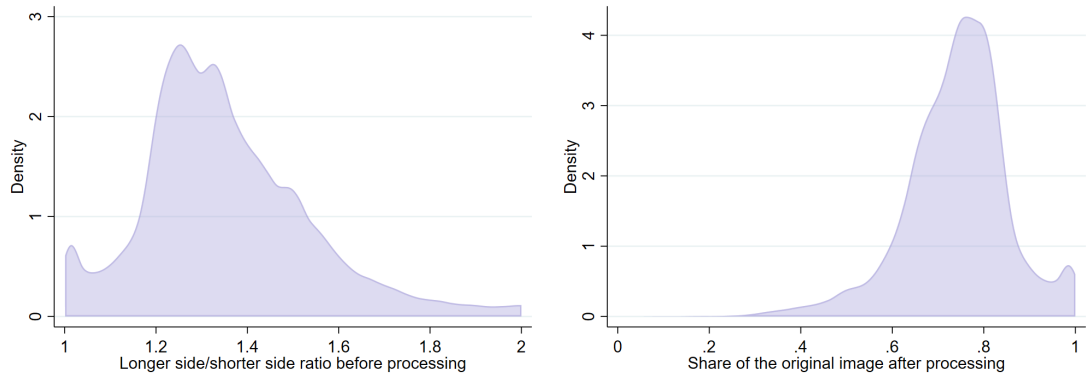
Notes: This Figure shows the highest emotion scores over our 80,000 labeled paintings, by emotion (Amusement; Anger; Awe; Contentment; Disgust; Excitement; Fear; Sadness; and Other). The most frequent emotions are: contentment (typically associated with landscapes, nature morte, and peaceful depictions); awe (typically associated with displays of power or opulence); sadness (typically associated with melancholic portraits, depictions of poverty, or religious imagery); fear (typically associated with chaotic images, either through the subject or through textures—a ship engulfed by waves for instance); excitement; and disgust (often associated with nudity). Sources: ArtEmis (Achlioptas et al., 2021, 450,000 labels), ArtELingo (Mohamed et al., 2022, 1,000,000 labels).

We display in Figure 2 the average emotion score associated to our labeled paintings. The most frequent emotions are: contentment (typically associated with landscapes, nature morte, and peaceful depictions); awe (typically associated with displays of power or opulence); sadness (typically associated with melancholic portraits, depictions of poverty, or religious imagery); fear (typically associated with chaotic images, either through the subject or through textures—a ship engulfed by waves for instance); excitement; and disgust (often associated with nudity). We illustrate this disparity in Figure 3 where we display paintings with the highest score among evaluators for each emotion.

1.2 An image classification algorithm

We develop a neural net to classify the emotions conveyed through paintings.

Figure 4. 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 from Wiki-Art; panel (b) displays the distribution of the average share of these initial images that is kept through our image pre-processing procedure.

Image pre-processing The raw images differ in size, i.e., both in their resolution and in their relative width. The typical canvas would, however, mostly range between a ratio of 1 and 1.5 for the longer side/shorter side ratio (see Figure 4), and the nature of artwork usually implies that the center of the piece is its most important part. The reason why we discuss dimensions is because image classification algorithms typically use squared images and a unique, common size. We thus crop our images using a center-crop, while maintaining the aspect-ratio, and convert them into $384 \times 384 \times 3$ arrays. The paintings displayed in Figure 3 have actually gone through such pre-processing, illustrating that the centering and pixelization of paintings might not alter much our perception in the majority of cases.

A neural net structure We rely on a transfer learning approach to predict emotional scores. Intuitively, our neural network proceeds in two steps: the first step is based on a powerful, pre-trained image classification algorithm which converts the image patterns into objects, textures, and their interaction (the transfer-learning part); and the second step adds layers to map these patterns onto emotions.

More specifically, we use a pre-estimated encoder model to extract multiple generic features from each image. This encoder consists of the convolutional layers of an Efficient-NetV2-S model (Tan and Le, 2021), already estimated on the ImageNet classification problem (Deng et al., 2009, a multinomial classification mapping 1,400,000 photos of objects belonging to 1000 different classes, from which we remove the output layer). The output of such a model is 12×12×1280 feature maps, with 1280 localized variables corresponding to different parts of the input image. We then attach a "regression head" to the encoder

in order to map the extracted features into scores associated with each emotion. This regression head is slightly more sophisticated than a logit and better allows to capture the multi-dimensionality of the Efficient-NetV2-S output, but serves a similar purpose. In effect, the extracted features are used as input to two densely connected hidden layers, where each layer contains 64 units with ReLU activation, and is followed by a batch-normalization (Ioffe and Szegedy, 2015) and dropout layers (Tompson et al., 2014). The output layer adds a multinomial logistic transformation to match the target distribution such that the emotion scores sum up to 1 and can be interpreted as probabilities.

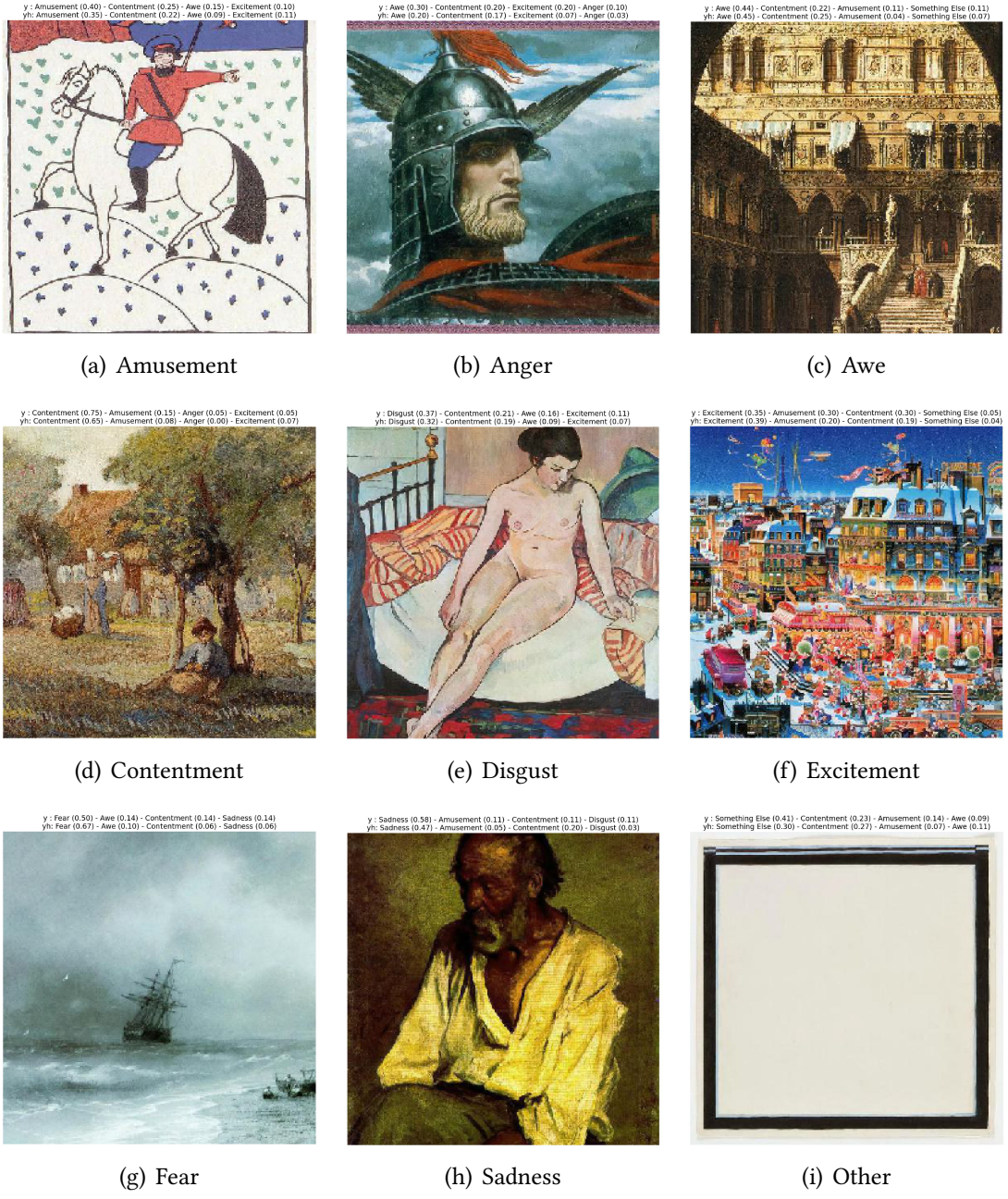
The model minimizes the Kullback-Leibler divergence (or relative entropy) between the predicted and the observed distributions of scores, allowing us to better account for the fact that our target is a "probability distribution". The parameters are estimated by maximum likelihood using the Adam optimizer (i.e., with an adaptive learning rate for each parameter, dynamically estimated learning rates and sub-samples of 64 images at every optimization iteration, see Kingma and Ba, 2015). In practice, the model is written using the GPU implementation of the Tensorflow library and could be run by a standard high-specification laptop with a properly-configured Python library: it does not require access to High Performance Computing (HPC) machines.

Optimization The concrete model optimization works as follows. We first align the parameters of the encoder and the regression head by optimizing the model while freezing the encoder parameters. In other words, we consider the model converting images into objects, texture, etc., as given and adjust the head which maps these multi-dimensional output onto emotions. Optimizing the entire model would indeed cause large gradient updates from the head (initialized with random values), which would break the (inherited) ability of the encoder to extract patterns.

Second, we fine-tune the model by (i) unfreezing encoder parameters (except 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. This last step intuitively allows the image-conversion step to adjust to the targeted output. In particular, our classification does not intend to classify dogs, cats, or taxis, but to capture the emotions conveyed by paintings. The latter might pass through textures, sharp angles, color contrasts, etc., which are carefully considered by artists, but are typically more coincidental in real life depictions.

Validation We validate the model by splitting the annotated sample into three sub-samples: a training sub-sample (70%); a test sub-sample (15%); and a validation sub-sample (15%). The training sample is used to estimate the parameters; the validation

Figure 5. The predicted emotion scores—an illustration.



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

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

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

Notes: This Figure illustrated our model performance on a selection of 9 paintings of the test sub-sample, chosen as representative of each emotion (Amusement; Anger; Awe; Contentment; Disgust; Excitement; Fear; Sadness; and Other).

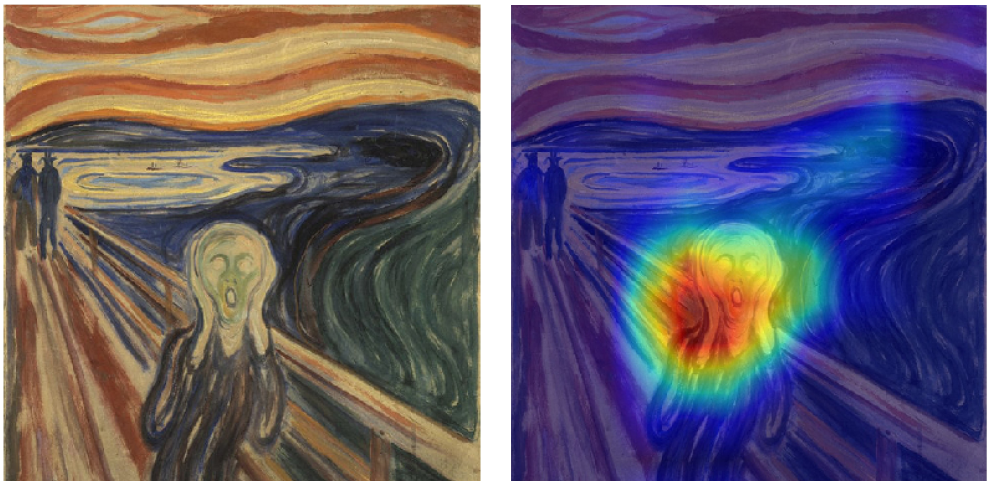
sample is used to select parameters with the best generalization performance;[7] and the test sample is used to evaluate the model performance.

[7]To further limit over-fitting, random transformations (i.e., flipping, zooms, rotations) are applied to each training batch, which also increases the size of the training sample and provides invariance properties to these transformations. Since we also use dropout layers to avoid over-fitting, we can use Monte-Carlo Dropout (Gal and Ghahramani, 2016a) to compute standard errors for our predictions [in progress].

The nature of the prediction, i.e., a probability distribution or a set $\{\mathbf{E}_i, \forall i\}$, and its associated statistics, i.e., the Kullback-Leibler divergence or relative entropy, make it hard to extract telling indicators of performance. To illustrate such performance, we extract 9 images from the test sub-sample, chosen such as to convey one specific emotion and illustrative of the model strengths and weaknesses (see Figure 5). One can see that the model predicts well the most common emotions, e.g., awe (illustrated by opulent buildings), contentment (peaceful landscapes), excitement (through warm colors, lights and human activity), fear (cold colors and indistinct texture evoking uncertainty), or sadness (conveyed through facial expressions or colors). The prediction is weaker for less frequent emotions, e.g., anger, as we discuss below. The prediction also inherits the prejudices of evaluators. For instance, female nudity is typically associated with disgust, reflecting moral codes and beauty standards of our times—maybe more than the artist's original intention.

Identification One drawback of neural networks is that they are very complex predictors combining numerous parameters into connected layers, usually making it hard to understand which variation is most used for identification and prediction. By contrast, regression models allow the statistician to measure the contribution of the different variables to explaining the overall variance of the outcome.

Figure 6. The predicted emotion scores—a heatmap for "identification".



(a) Original picture (b) Heatmap (fear)

Notes: Panel (a) shows The Scream painted by Edvard Munch in 1893; panel (b) displays a heatmap using Class Activation Mapping (Zhou et al., 2016)—which isolates the part of the image that is most used by the network for predicting (in this case, the fear score).

We cannot provide an equivalent approach. Instead, we illustrate the nature of identification in one example: The Scream painted by Edvard Munch in 1893, which is pre-

dicted to inspire fear/angst to annotators. We compare our centered image (panel a of Figure 6) to a heatmap using Class Activation Mapping (Zhou et al., 2016)—which isolates the part of the image that is most used by the network for predicting (in this case, the fear score, panel b). One can see that the main factor for such prediction is the face in the foreground. We provide less sophisticated support for this decomposition in Appendix B where we alter the original painting, its color and the facial expression in the foreground (replacing The Scream by "The Smile").

Issues In general, the vector of emotions is skewed, both in the annotated dataset and in the predictions. Predicting a properly multi-dimensional vector of emotions is challenging. In particular, the more subtle, secondary emotions are typically less precisely captured than the main "in-your-face" message conveyed by each image.

Another issue is neutrality versus distinct emotions. In effect, many images evoke *contentment* to evaluators, and the model tends to over-predict such outcome at the expense of more distinctive emotions (e.g., the infrequent *anger*). With a discrete classification, we could discipline the relative weight of neutral emotions versus more distinct ones using class imbalance tools.

2 Empirical strategy

This section describes our methodology to project measures of emotions at the image level onto their context (a location l and a given year t), thus cleaned from variations as induced by an artist's identity, their influence or their career evolution.

2.1 A reduction of dimensionality

The first stage of our empirical strategy consists in reducing the dimensionality of the problem. We rely on Principal Component Analysis (PCA): we compute the eigenvectors of the covariance matrix (scores) and keep the top 4 components for whom the eigenvalue is above or close to 1. The fifth component has an eigenvalue of about 0.55.

We report the loadings associated to the different emotion scores in Table 1: the first principal component (PC 1) positively loads negative emotions (fear, anger, sadness); the second principal component (PC 2) positively loads positive emotions (excitement, amusement); the third component (PC 3) is really about awe; and the fourth component (PC 4) is capturing the residual category "Other".

Table 1. A reduction of dimensionality.

	PC 1	PC 2	PC 3	PC 4
Eigenvalue	*2.95*	*1.89*	*1.26*	*1.00*
Difference	*1.06*	*0.62*	*0.26*	*0.45*
Awe	-0.2078	0.0014	0.7263	0.2507
Contentment	-0.4976	-0.1799	0.2518	0.1422
Amusement	-0.1496	0.4683	-0.3891	-0.3083
Excitement	-0.2609	0.4040	0.2713	-0.3824
Fear	0.4120	-0.1694	0.3140	-0.3608
Anger	0.4486	0.1815	0.0668	-0.0958
Sadness	0.2922	-0.4824	-0.2622	-0.0198
Disgust	0.3424	0.4199	-0.0937	0.0568
Other	0.2123	0.3384	-0.0585	0.7298

Notes: This Table reports the loadings associated to the different emotion scores by the Principal Component Analysis (PCA). We report the eigenvalue associated with each eigenvector of the variance-covariance matrix in italic, and the difference with the next lower eigenvalue.

2.2 A standard, linear decomposition

The basic framework The output of Section 2.1 is a dataset of images with the following attributes: emotion indices, genre and movement, title, author, and year and location of production/creation.

Let i denote a certain image and $\mathbf{E}_i$ its associated vector of emotions. This image has been produced by a certain artist a in location l and year t. A standard, linear decomposition would correspond to the model,

$$\mathbf{E}_i = \alpha_{a(i)} + \psi_{l(i),t(i)} + \beta \mathbf{X}_i + \varepsilon_i \tag{1}$$

where: $\mathbf{X}_i$ are exogenous covariates such as the age of the artist, the type of exercise, or the adopted style; $\alpha_{a(i)}$ is a fixed, unobserved artist effect; $(l(i), t(i))$ is the context and $\psi_{l(i),t(i)}$ is its associated unobserved effect; and ε_i is the idiosyncratic error term. The previous model is straightforward to estimate but relies on two assumptions: additivity in artist and context effects; and mean independence of the error term, i.e., the error is mean independent of previous (and future) contexts. Note that such decomposition is well-known in the labor literature as an AKM decomposition (see Abowd et al., 1999, our artists correspond to "workers" while our contexts are "firms" in their representation, and their objective is usually to extract firm productivity and isolate it from the selection of workers), and it suffers from various flaws that have been identified by recent contributions (Bonhomme et al., 2020).

Finally, we could consider adding the interaction of context-specific fixed effects with the socio-economic status and characteristics of artists to extract different context-specific residuals. Artists are indeed transcribing an era, but they themselves might be confronted to different living standards, depending on their societal position.

A panel of locations We can use the previous model to identify a set of estimates $\psi_{l(i),t(i)}$ at the year and location level. In practice, however, the number of observations in each year/location cell might be small and we thus aggregate/interpolate our residuals along both the spatial and time dimensions. More specifically, we consider two smoothing procedures: (i) aggregation, in which we aggregate the residual by considering an average within regional administrative units or countries in every decade; and (ii) interpolation, in which we fit a parametric best predictor that is continuous in space and in time. More specifically, we apply a Hodrick-Prescott filter with the same parameter as for business cycles (and annual data) to $\psi_{l,t}$ in order to reduce the volatility of the resulting time series. In stark contrast with the business cycle literature, our object of interest is the smoothed series, rather than the deviation from it.

Issues The previous model would deliver biased context-specific effects $\psi_{l,t}$ if there were complementarities in artist and context effects—the effects would then not be additively separable—or if the error was correlated with previous (and future) contexts. The latter would occur if artists were not moving much across contexts, if mobility was not random, or if artists were influenced by their past contexts—possibly through the production of their peers. It will be challenging to relax separable-additivity; we might however control for endogenous mobility, peer effects or dynamic context effects.

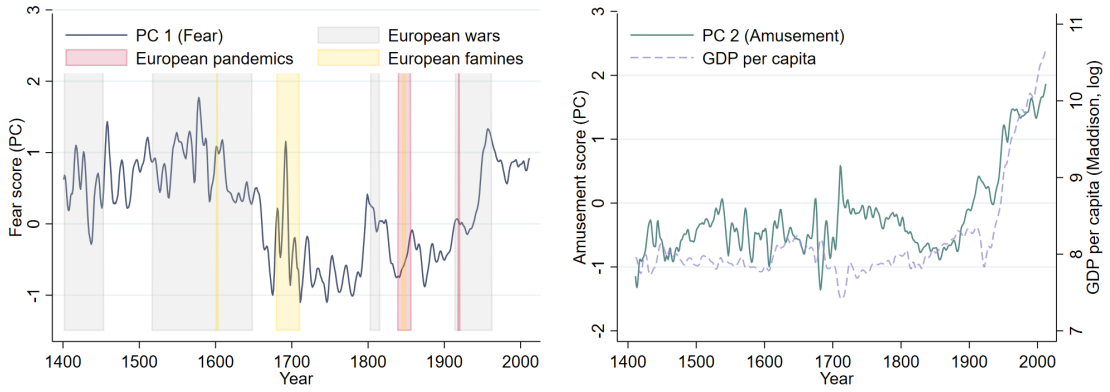
We describe below some of these concerns in greater detail. First, there might be a "limited mobility bias". Artists *do* move across contexts in the sense that t varies across the different productions of a same artist (and could theoretically be differentiated from a life-cycle effect due to age—if the latter was modeled with generic age effects); they might however rarely move across locations l and the mobility patterns might change over eras as apprenticeship is replaced by more formal courses. Mobility patterns might also change over the life-cycle of an artist as apprenticeship gives way to more formal invitations. Second, mobility is not random. An artist would be invited in certain places at certain times as a response to their previous productions, in response to expectations about the local economic environment, or in response to omitted variation. Francois I, who had direct impact on economic outcomes in France from 1515–1547, invited Leonardo da Vinci, Benvenuto Cellini, Andrea del Sarto, Rosso Fiorentino, Francesco Primaticcio, Joos van Cleve, Godefroy le Batave, etc., in part to spread the Italian and Nordic

Renaissance to France. Third, previous contexts might spill over to the future production of artists and introduce spurious time and spatial auto-correlation within the estimated context-specific effects $\psi_{l,t}$.

3 The geography of emotions over time

This section describes the geography of emotions over time.[8]

Figure 7. The emotions over time.



(a) PC 1 (Fear)

(b) PC 2 (Excitement)

Notes: This Figure illustrates the average evolution of PC 1 (Fear) and PC 2 (Excitement) between 1400 and 2020 within the Wiki-Art sample. In Panel (a), we provide shaded areas highlighting the periods of major wars (until 1453 for the Hundred Years War, 1517-1648 for the Religious Wars, 1803-1815 for the Napoleonic Wars, 1914-1962 for WW1, WW2 until the Cuba missile crisis), famines (1601-1603, 1680-1710, 1845-1849) and pandemics (1839-1856, 1918-1920) in Western Europe (where we observe most paintings). In Panel (b), we add a measure of average (log) GDP per capita where we observe most paintings (Source: Maddison, Bolt and Van Zanden, 2014; Bolt et al., 2018).

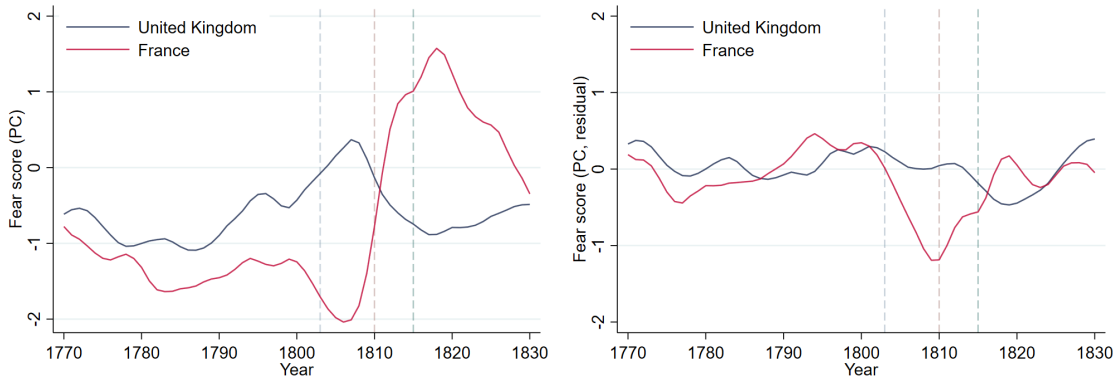
The emotions over time We first provide descriptive statistics about the emotion indices over time and across environments. In Figure 7, we report the average evolution of PC 1 (Fear) and PC 2 (Excitement) between 1400 and 2020 within the Wiki-Art sample. The left panel shows large variation in the depiction of fear through paintings, some of which reflecting long-run, secular trends, some of which being more impermanent. First, one can observe that fear is more prevalent before 1650 and after 1900. This observation is partly related to the movements and genres privileged by artists: Religious/allegorical depictions are typically associated with fear and are over-represented during the 15th-16th centuries; abstract pieces are also more likely to inspire fear and

[8]The validation exercise relies on the following data sources: the Maddison Project (GDP/population from 1400 onward, see Bolt and Van Zanden, 2014; Bolt et al., 2018); the Macro History Database (macroeconomic variables from 1870 onward, including trade openness, for 18 economies, see Schularick and Taylor, 2012); Conflict catalog/Latent Democracy/Polity (from 1800 onward, see Boix et al., 2013; Foldvari et al., 2014); and the World Inequality Database (Alvaredo et al., 2020).

more likely to appear in the 20th century.[9] Second, there are impermanent changes in depicted fear, which appear to correlate with marking historical events: wars, famines or pandemics. We illustrate this relationship by highlighting the periods of major wars (until 1453 for the Hundred Years War, 1517-1648 for the Religious Wars, 1803-1815 for the Napoleonic Wars, 1914-1962 for WW1, WW2 until the Cuban missile crisis), famines (1601-1603, 1680-1710, 1845-1849) and pandemics (1839-1856, 1918-1920) in Western Europe (where we observe most paintings): during these events, we observe a (sometimes sudden) increase in depicted fear.

The right panel (Excitement) also displays some high-frequency fluctuation, but the most striking pattern is a steady increase in excitement from the mid-19th century onward—thus coinciding with the sustained increase in living standards for the average individual living in our sample locations.[10] This evidence is however mixing very different trajectories across countries; we investigate these differential trajectories next.

Figure 8. The Napoleonic Wars (1803–1815).



(a) PC 1 (Fear)

(b) PC 1 (Fear, residual)

Notes: This Figure illustrates the evolution of PC 1 (Fear) between 1770 and 1830 in France and Great Britain. In Panel (a), we use the raw score. In Panel (b), we use a residual score controlling for movements, genres, artist fixed effects and age fixed effects. The vertical lines indicate the formal start of the Napoleonic Wars (1803—note that 1799 is the start of the Consulate), the peak for France (1810), and the end of the Napoleonic Wars (1815).

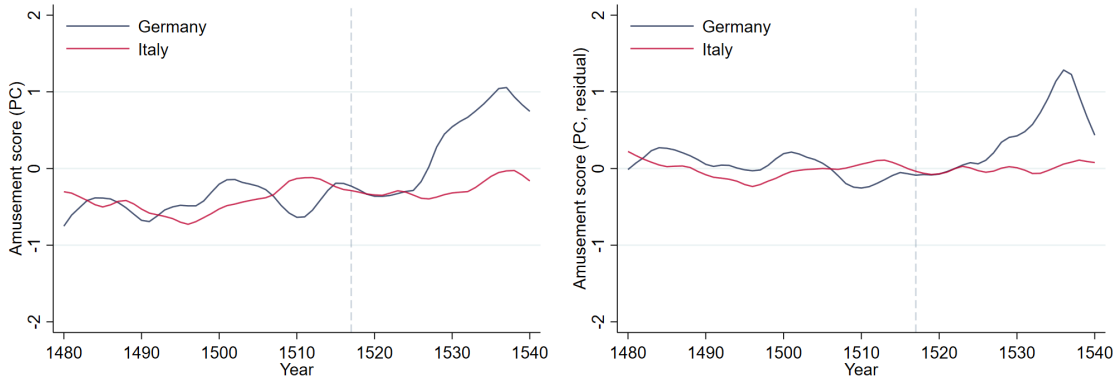
Event studies In this section, we consider a few selected events—a war, the emergence of a new ideology, a new political regime, and a series of societal reforms and an opening to trade and foreign technologies—to illustrate the differential evolutions of depicted emotions across contexts.

[9]We provide descriptive statistics about the prevalence of the different movements/genres over time and their average emotion scores in Appendix B.

[10]We display the cross-sectional variation in those two indices in Appendix C, where we find limited systematic correlation with general economic development (possibly due to fixed, cultural differences across countries).

In a first step, we come back to our motivating discussion about the expansionary ambition of Napoleon, seen from a French perspective in the early, successful years of the Napoleonic Wars and from an English perspective in the later period of stalling advances (see Figure 1, panels c and d). We provide a more systematic comparison of these two contexts (France versus Great Britain) in panel (a) of Figure 8 where we display the average Fear index ($\bar{\mathbf{E}}_{l,t}$) for $l \in \{\text{France, Great Britain}\}$ between 1770 and 1830. We see that: (i) France and Great Britain follow similar trends before the French Revolution; (ii) the two time series diverge subsequently with French paintings being far less likely to convey fear, especially so after the start of the Consulate (1799); (iii) there is a very sharp reversal between 1809 and 1815, coincidental with the stalling of the French empire and the disastrous Russian campaign; and (iv) the two series converge back to their pre-war levels at the end of the 1820s. These fluctuations might however indicate a different selection of genres, movements and artists across the different contexts. For instance, the French Consulate (then Empire) may privilege artists that are more willing to paint the Consul (then Emperor) in a positive light, as was the case with Jacques-Louis David. We shed some light into the extent to which these selection forces matter in panel (b) of Figure 8 where we show the differential dynamics followed by our residualized Fear index ($\psi_{l,t}$). Selection does appear to matter, in that the residualized series exhibit smaller fluctuations. The qualitative insight however remains the same: depicted fear sharply decreases in France before 1800 and strongly reverts to the mean after 1809.

Figure 9. The Reformation (1517).



(a) PC 2 (Excitement)

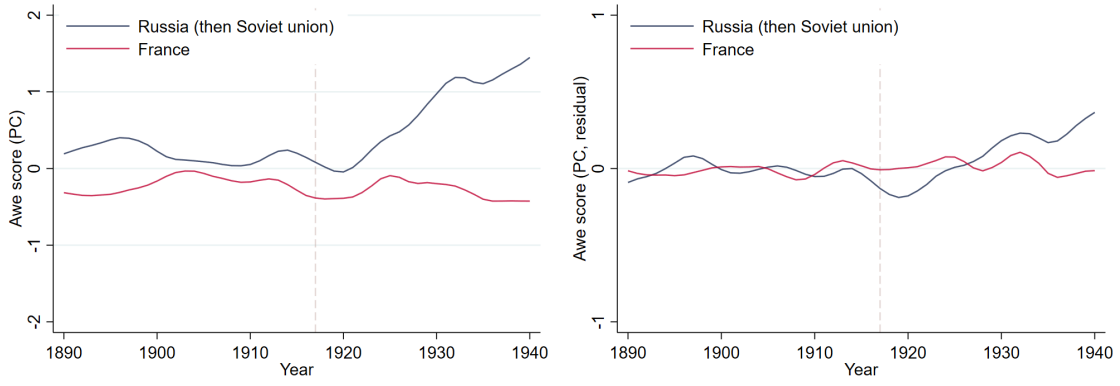
(b) PC 2 (Excitement, residual)

Notes: This Figure illustrates the evolution of PC 2 (Excitement) between 1480 and 1540 in the Holy Empire and Italian cities (as a control). In Panel (a), we use the raw score. In Panel (b), we use a residual score controlling for movements, genres, artist fixed effects and age fixed effects. The vertical line indicates the start of the Reformation (1517).

We replicate this approach with different events and contexts in Figures 9, 10 and 11. Figure 9 shows the differential depiction of excitement in German paintings versus Ital-

ian paintings during the early Reformation where Martin Luther challenged the Catholic Church leading to a propagation of new ideas (based on human capital formation, work ethics and entrepreneurship) throughout Germany and then Europe. Paintings reflect these dynamics with PC 2 (Excitement) growing steadily from 1525 onwards in Germany, irrespective of cleaning for selection or not.

Figure 10. The Russian Revolution (1917).



(a) PC 3 (Awe)

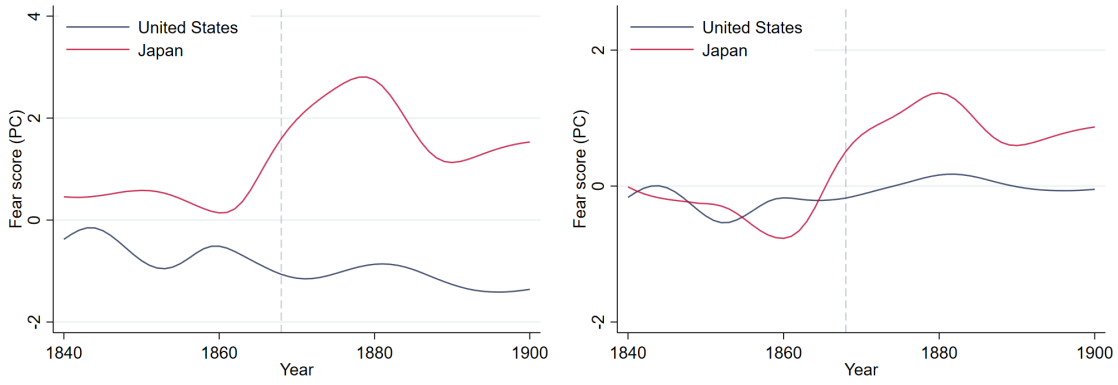
(b) PC 3 (Awe, residual)

Notes: This Figure illustrates the evolution of PC 3 (Awe) between 1890 and 1940 in Russia (then Soviet Union) and France (as a control). In Panel (a), we use the raw score. In Panel (b), we use a residual score controlling for movements, genres, artist fixed effects and age fixed effects. The vertical line indicates the start of the Russian Revolution (1917).

Figure 10 focuses on the sudden emergence of a new political regime in Russia following the Russian Revolution of 1917, the subsequent Civil War, the victory of the Bolsheviks and the creation of the USSR. We use France as a control context, in the absence of better alternatives, and display the relative evolution of emotions from 1890 to 1940. The depiction of Awe in Russia (then Soviet Union) strongly increases after 1922 and the creation of the USSR (panel a of Figure 10). This increase however partly reflects the rise of a new movement, the Soviet art and the advent of the Russian avant-garde, and the selection of certain artists. We do observe, indeed, that the increase is less marked for the residualized Awe index ($\psi_{l,t}$, see panel b of Figure 10).

Figure 11 focuses on a more ambiguous experiment, i.e., the Meiji Restoration in Japan (1868). The Meiji Restoration is both a political change—with a strong centralization of power and a restoration of imperial rule—and an economic transformation—with an opening to trade and foreign technologies, leading to industrialization, rural-urban migration, and human capital formation. While the Meiji Restoration is seen as a positive transformation of Japan with hindsight, those were times of upheaval where the previous order—in place for centuries—would be rapidly replaced with unknown institutions and technologies. Our analysis of paintings shows a sharp increase in fear depicted by

Figure 11. The Meiji Restoration in Japan (1868).



(a) PC 1 (Fear) (b) PC 1 (Fear, residual)

Notes: This Figure illustrates the evolution of PC 1 (Fear) between 1840 and 1900 in the United States (as a control) and Japan. In Panel (a), we use the raw score. In Panel (b), we use a residual score controlling for movements, genres, artist fixed effects and age fixed effects. The vertical line indicates the start of the Meiji Restoration (1868).

artists, whether captured by the raw Fear index or its residualized counterpart.

Table 2. The variation underlying the emotion indices.

	Latent Democracy	Trade openness	GDP p.c. growth
PC 1 (Fear)	-0.1897	0.0141	-0.0029
	(0.0450)	(0.0102)	(0.0019)
PC 2 (Excitement)	0.0693	0.0281	0.0015
	(0.0523)	(0.0110)	(0.0023)
PC 3 (Awe)	-0.0097	0.0477	0.0157
	(0.0647)	(0.0148)	(0.0026)
PC 4 (Other)	0.2692	-0.0535	-0.0020
	(0.0697)	(0.0150)	(0.0030)
Observations	2,447	1,577	3,943

Notes: This Table reports the outcome of specification (2) for three outcomes: a Latent Democracy index constructed for numerous economies from 1850 onward (based on polity, see Foldvari et al., 2014), trade openness for 18 major economies (from the Macro History Database, see Schularick and Taylor, 2012), and annual growth in GDP per capita from the Maddison data (Bolt and Van Zanden, 2014; Bolt et al., 2018). Robust standard errors are reported between parentheses.

Systematic correlations In the last part of this preliminary draft, we provide a more systematic analysis of the variation underlying our emotion indices over time and across countries. Letting l denote a location (a country, as defined using the contemporary

borders) and t a year, we estimate,

$$y_{l,t} = \beta\psi_{l,t} + \nu_l + \eta_l + \varepsilon_{l,t} \tag{2}$$

where $y_{l,t}$ are measures of political and economic development and each unit of observation is weighted by the number of art pieces used for inferring the (standardized) emotion indices $\psi_{l,t}$.

We report the outcome of specification (2) in Table 2 for three outcomes: a Latent Democracy index constructed for numerous economies from 1850 onward (based on polity, see Foldvari et al., 2014), trade openness for 18 major economies (from the Macro History Database, see Schularick and Taylor, 2012), and annual growth in GDP per capita from the Maddison data (Bolt and Van Zanden, 2014; Bolt et al., 2018). We find that a standard deviation in PC 1 (Fear) is associated with a 0.19 decrease in the latent democracy score (from -1.3 to 2), a shift equivalent to 25% of a standard deviation. A standard deviation in PC 4 (Other) is associated with a 0.27 decrease in the same score—about a third of its standard deviation. In short, democracy is associated with fewer fearful depictions and more abstract paintings.

Our emotion scores are also predictive of changes in trade openness. A one standard deviation in PC 3 (Awe) is associated with a 5 percentage point increase in the share of exports and imports as normalized by total output. Trading economies might bring a feeling of opulence that artists transcribe in their work. The third column of Table 2 brings more direct evidence on this opulence effect: the same standard deviation in PC 3 (Awe) is associated with 1.6 percentage point increase in the growth of output per capita, providing similar correlational insight as the right panel of Figure 7 but within the panel of locations.

Preliminary conclusion and next steps The previous evidence provides some support for the use of artistic expression(s) as a way to understand underlying economic forces affecting different societies in the medium and longer run. One possible contribution could be to shed light on subtle dimensions of welfare, e.g., due to uncertain beliefs about technology (automation) or large, external threats (climate change), which are usually poorly captured by measures of output.[11] Another, more direct contribution is to provide measures that correlate with welfare in (otherwise) data-scarce environments and to shed light on geographic variation within countries (when historical data typically misses such within-country variations).

[11] A recent macroeconomic literature discusses the cause and consequences of such uncertainty, usually measuring it by the cross-sectional volatility of expert predictions or by other measures of business confidence.

After properly applying our predictive procedure to the universe of paintings in Google Arts and Culture (and not to a limited sub-sample), the next steps of our analysis would be: (i) to provide a more systematic validation of its predictive power for basic measures of economic welfare in the past (e.g., economic growth); (ii) to provide a more systematic validation of its predictive power for more subtle aspects of welfare (uncertainty, economic inequalities) leveraging more recent data; (iii) to shed better light on the geography of economic development throughout the transformations of Medieval Europe, the Renaissance, the later Reformation, and the Enlightenment; and (iv) to exploit the heterogeneity across artists (e.g., their gender) to discuss inequalities in living standards within a given location.

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.

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

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." PhD dissertation, PSE (Paris School of Economics) 2020.

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

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.

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

Boix, Carles, Michael Miller, and Sebastian Rosato, "A complete data set of political regimes, 1800–2007," *Comparative political studies*, 2013, *46* (12), 1523–1554.

Bolt, Jutta and Jan Luiten Van Zanden, "The Maddison Project: collaborative research on historical national accounts," *The Economic History Review*, 2014, *67* (3), 627–651.

_ , Robert Inklaar, Herman De Jong, and Jan Luiten Van Zanden, "Rebasing Maddison: new income comparisons and the shape of long-run economic development," 2018.

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?," 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.

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.

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.

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

Foldvari, Peter et al., "A latent democracy measure 1850-2000," Technical Report 2014.

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.

_ **and** _ , "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning," in "Proceedings of the 33rd International Conference on Machine Learning," Vol. 48 2016, pp. 1050–1059.

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.

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.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in "2015 IEEE International Conference on Computer Vision (ICCV)" 2015, pp. 1026–1034.

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

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.

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.

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.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature*, 2015, (521), 436–444.

Liu, Z., H. Mao, C. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," in "2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)" IEEE Computer Society 2022, pp. 11966–11976.

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.

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," 2022.

Mohammad, Saif and Svetlana Kiritchenko, "Wikiart emotions: An annotated dataset of emotions evoked by art," in "Proceedings of the eleventh international conference on language resources and evaluation (LREC 2018)" 2018.

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.

_ , _ , _ , 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.

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

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.

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.

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.

Tan, Mingxing and Quoc V. Le, "EfficientNetV2: Smaller Models and Faster Training," *International Conference on Machine Learning*, 2021.

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

van der Maaten, Laurens and Geoffrey Hinton, "Visualizing Data using t-SNE," *Journal of Machine Learning Research*, 2008, *9*, 2579–2605.

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

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.

ONLINE APPENDIX—not for publication

A Data description

A.1 A large collection of images

Google Arts and Culture is a partnership between Google and a large number of cultural institutions (museums, galleries, etc.) which aims at digitizing and publishing art collections for online access. The largest art collections are all partners, such that most well-known paintings that are open to the public are also available as online images through the platform. For instance, the Musee d'Orsay (Paris) shares hundreds of items, of which The Church in Auvers-sur-Oise by Vincent Van Gogh or The Ballet Class by Edgar Degas. The Museum of Modern Art shares its collection, including Hope, II, by Gustav Klimt, or Turning Road at Montgeroult by Paul Cezanne.

Overall, more than 15,000 artists are present on Google Arts and Culture and about 1,000,000 paintings. The most well-known artists, e.g., Albrecht Duerer, Pierre-Auguste Renoir, Rembrandt, Vincent Van Gogh, Edgar Degas, Henry de Toulouse-Lautrec, William Turner, would typically have between 500 and 2000 paintings. While the art collections mostly consist in European pieces, they also cover Chinese calligraphy, Japanese prints and water-based pieces, "primitive" art, landscapes from the Hudson River School, or American realism.

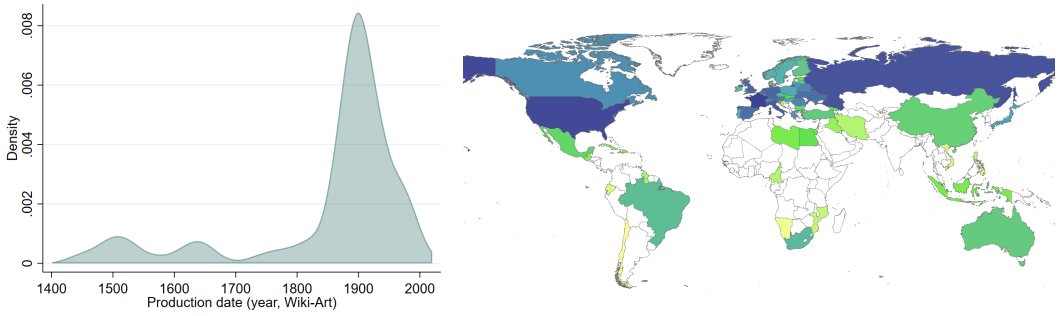
Artists and paintings come with a description and various tags. For instance, the information provided with "The Large Bathers" by Cezanne is the following—where one can see that the location is not consistently documented (Aix-en-Provence is the actual name of the city):

> Title: The Large Bathers
>
> Date: 1900-1906
>
> Location: aix en provance, France
>
> Physical Dimensions: w98.74 x h82.88 in (Overall)
>
> Type: Paintings
>
> External Link: Philadelphia Museum of Art
>
> Medium: Oil on canvas
>
> Provenance: Estate of Paul Cezanne, 1906; purchased by Ambroise Vollard, Paris from Cezanne's son, 1907; Auguste Pellerin (1852-1929), Paris, by 1923; by descent to his son Jean-Victor Pellerin, Paris, 1929-1936 [1]; with Wildenstein & Co., New York, acting as agent for Pellerin, 1936 [2]; purchased by the City of Philadelphia with the W. P. Wilstach Fund, July 6, 1937 [3]. 1.

Lent by M. and Mme. Pellerin to the 1936 exhibition “Cezanne”, MusÃ©e de l’Orangerie, Paris, no. 107. 2. Provenance per John Rewald, The Paintings of Paul Cezanne: A Catalogue Raisonne, New York, 1996, no. 857. See also Joseph Rishel, Cezanne in Philadelphia Collections, Philadelphia, 1983, p. xvi. 3. Copy of dated receipt in registrar file., Purchased with the W. P. Wilstach Fund, 1937

Artist/Maker: Paul Cezanne, French, 1839 - 1906

Figure A1. Wiki-Art and its coverage.



(a) Time of production (b) Location

Notes: This Figure shows the distribution of production years and geographic locations for the 80,000 paintings in Wiki-Art.

A.2 Wiki-Art and annotations

Wiki-Art Wiki-Art is another online repository of reasonably high-quality images of paintings (Saleh and Elgammal, 2016). The collection is smaller than Google Arts and Culture (about 80,000 paintings), but (i) provides well-organized information about the painting itself and the artist (see example below) and (ii) is augmented with two annotation projects whereby individuals were asked to associate 9 emotions with each painting.

Paul Cezanne

Born: January 19, 1839; Aix-en-Provence, France

Died: October 22, 1906; Aix-en-Provence, France

Nationality: French

Art Movement: Post-Impressionism

Field: painting

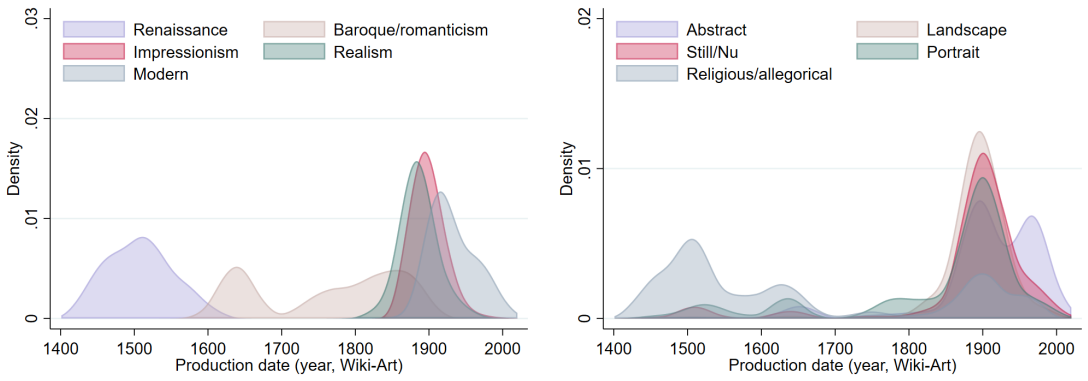
Influenced by: Gustave Courbet, El Greco, Charles-Francois Daubigny, Nicolas Poussin, Pierre-Auguste Renoir, Eugene Delacroix, Jean-Baptiste-Simeon Chardin

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

Friends and Co-workers: Paul Gauguin

As shown in Figures A1 and A2, its coverage is heavily biased towards European and American work, and towards the nineteenth century (partly because copyrights prevent later pieces to be shared online). One corollary is that the training sample will be better suited, in principle, to capturing emotions for work between 1800 and 2000.

Figure A2. Wiki-Art and its coverage (movements and genre).



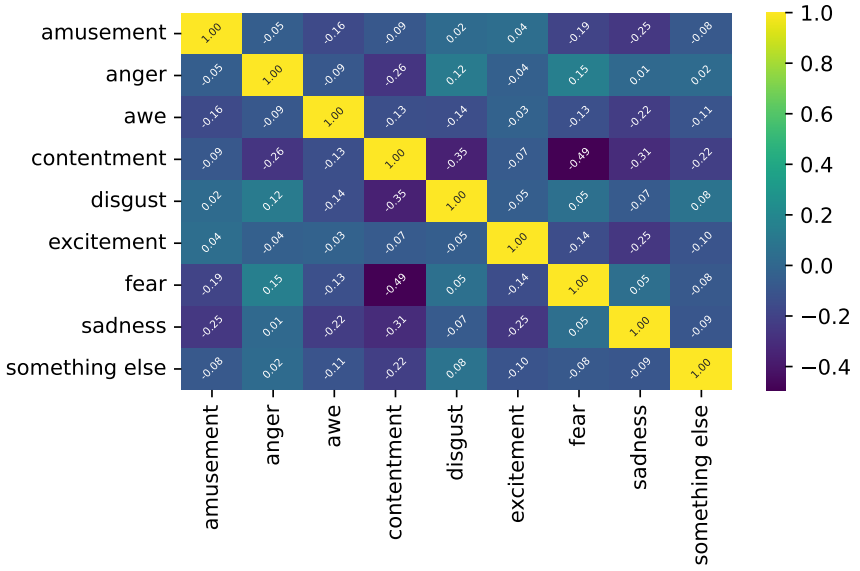
(a) Movements (b) Genres

Notes: Panel a shows the distribution of production years across different movements (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). Panel b shows the distribution of production years across different exercises (abstract: abstract, genre painting; landscape: landscape, cityscape; still/nu: animal painting, flower painting, marina, sketch and study, still life, nude painting; portrait: portrait, self-portrait; religious/allegorical: mythological painting, allegorical painting, religious painting).

Annotations Our training sample consists of two sources of annotations associated with Wiki-Art images: ArtEmis (Achlioptas et al., 2021, providing about 450,000 labels for the 80,000 paintings discussed above); and ArtELingo (Mohamed et al., 2022, providing more than 1,000,000 labels for the same 80,000 paintings). The former dataset

contains a verbal explanation to justify the choice of emotion; the latter dataset exploits differences across possible annotators in language and culture (thereby covering many different countries). We however ignore these dimensions and create a unique score associated with each painting.

Figure A3. Correlation across emotions.



Notes: This Figure shows the correlation between each dimension of the emotion score associated with our 80,000 annotated paintings (Amusement; Anger; Awe; Contentment; Disgust; Excitement; Fear; Sadness; and Other).

More specifically, we treat these several labels as a probabilistic evaluation from a unique individual. Each painting will be allocated a vector of scores: the shares of evaluators having mentioned each of the nine emotions (*amusement*, *anger*, *awe*, *contentment*, *disgust*, *excitement*, *fear*, *sadness* and *other*), which we normalize such that the scores sum up to 1 for each painting. Figure A3 shows the correlation between the resulting scores: fear, anger and disgust appear to be substitutes within a well-defined set of negative emotions; at the other end of the spectrum, contentment is often thought as a separate emotion and is rarely mentioned with other emotions in a systematic manner.

B An image classification algorithm

We provide in Section 1.2 a description of our image classification algorithm. In this Appendix, we provide intuitions about our model structure, its identification of emotions, and an illustration of its predictive power across different genres and movements.

Figure A4. Texture and the difference between photographs and paintings.

Notes: This Figure illustrates the difference between a photograph and a painting, justifying our methodology—which re-optimizes the encoder model that extracts multiple generic features from each image. The example is "The Church at Auvers" (Vincent van Gogh, 1890) which we compare with the actual church, still in a comparable state (Auvers-sur-Oise).

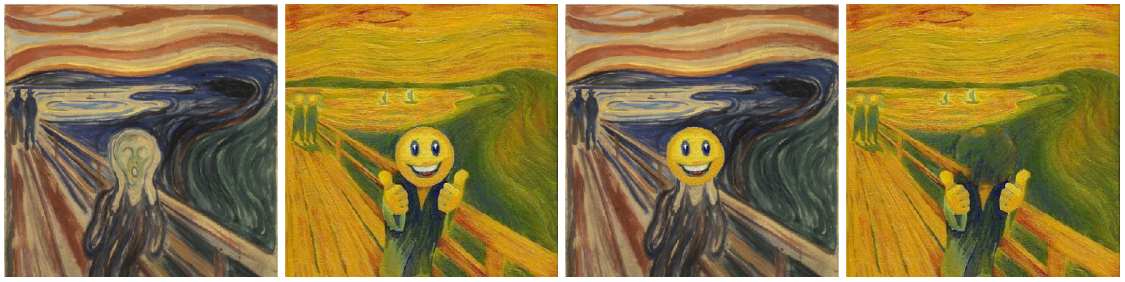
B.1 Texture and the difference between photographs and paintings

Our neural net structure re-optimizes the encoder model that extracts multiple generic features from each image, even though with a slow learning rate. The motive for doing so is that paintings differ from photographs in their composition, structure and texture—all possibly used by artists in order to convey emotions. A well-known example is Impressionism (and Post-Impressionism); we provide an illustration in Figure A4 with "The Church at Auvers" (Vincent van Gogh, 1890) which we compare with the actual church, still in a comparable state (Auvers-sur-Oise). The textures and deformations create a focal point towards the church and makes it a threatening figure.

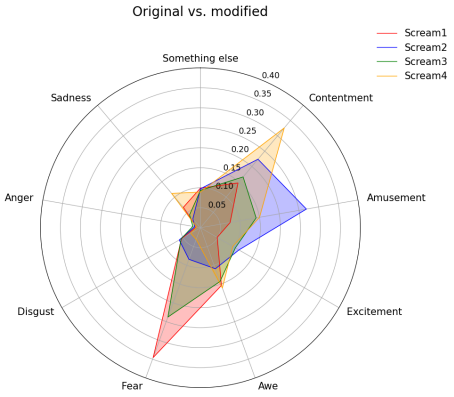
B.2 An illustration of "identification"

In Section 1.2, we isolate the part of "The Scream", painted by Edvard Munch in 1893, which inspires fear/angst to annotators. Our approach exploits the Class Activation

Figure A5. The predicted emotion scores—an illustration of "identification".



(a) Original (1) (b) Parody (2) (c) Original (m, 3) (d) Parody (m, 4)



(e) Scores

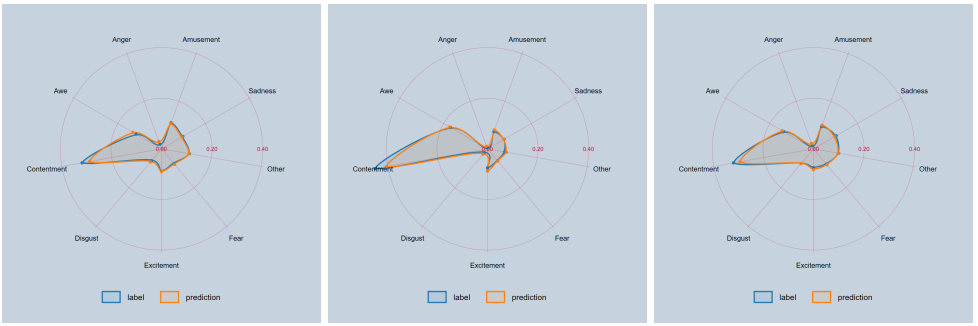
Mapping (Zhou et al., 2016). An alternative, less theoretical approach consists in manually overwriting parts of the original image, and predicting emotion scores for these alternative paintings. We consider such a exercise in Figure A5 where we provide four variations of "The Scream" and their predicted emotional scores through our estimated model.

One can see that the facial expression of the central/foreground character is the main factor explaining "fear". We do see however that the dominant colors and background texture are subtle factors that do end up mattering from the viewpoint of our model. For instance, the modified image in panel (d) evokes "contentment" because the foreground then fades away and the model focuses instead on the background which appears to depict a peaceful scene.

B.3 Predictive power across different genres and movements

In this subsection, we discuss the predictive power of our model across different genres and movements. More specifically, we report in Figure A6 the predicted emotion scores across different exercises (abstract: abstract, genre painting; landscape: landscape,

Figure A6. The predicted emotion scores across genres.



(a) Abstract (b) Landscape (c) Still

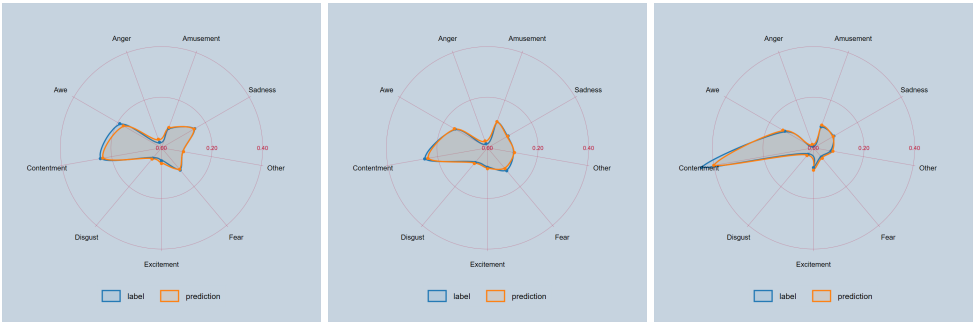
(d) Portrait (e) Religious/allegorical

Notes: This Figure shows the distribution of emotions—annotated and predicted—across different exercises (abstract: abstract, genre painting; landscape: landscape, cityscape; still/nu: animal painting, flower painting, marina, sketch and study, still life, nude painting; portrait: portrait, self-portrait; religious/allegorical: mythological painting, allegorical painting, religious painting).

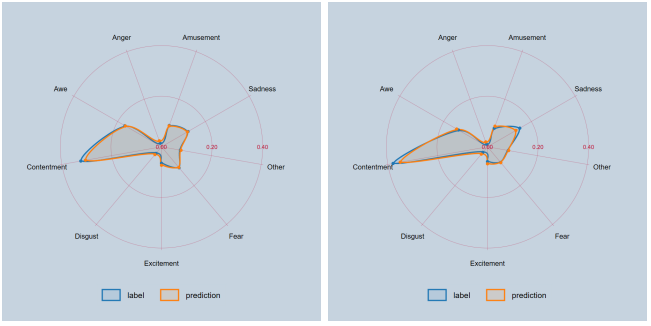
cityscape; still/nu: animal painting, flower painting, marina, sketch and study, still life, nude painting; portrait: portrait, self-portrait; religious/allegorical: mythological painting, allegorical painting, religious painting) and we report in Figure A7 the predicted emotion scores across different movements (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).

Different genres evoke different emotions to annotators: landscapes evoke "contentment" or "awe"; religious/allegorical pieces evoke "awe", "fear" or "sadness". Different movements also evoke different emotions to annotators. One can see however that our model performs well across the board and does capture these subtle differences.

Figure A7. The predicted emotion scores across movements.



(a) Renaissance (b) Modern (c) Impressionism

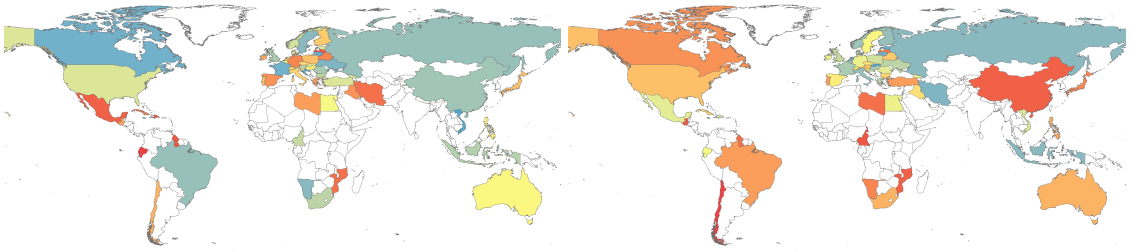


(d) Baroque (e) Realism

Notes: This Figure shows the distribution of emotions—annotated and predicted—across different movements (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).

C The geography of emotions over time

Figure A8. The geography of emotions.



(a) PC 1 (Fear) (b) PC 2 (Excitement)

Notes: This Figure illustrates the cross-sectional variation in PC 1 (Fear) and PC 2 (Excitement) between 1400 and 2020 within the Wiki-Art sample.