Values of China: Toward a Cultural Map*

Melanie Meng Xue†

LSE

March 20, 2025

Abstract

This paper constructs a cultural map of China by analyzing nearly 20,000 folktales collected in the 1980s, illustrating how cultural traits can shape economic outcomes. By leveraging advanced text-analysis techniques, the study identifies themes such as trust, honesty, and effort within these oral traditions. It then connects these historical narratives to contemporary attitudes and beliefs from the Chinese General Social Survey, thereby linking longstanding cultural values to modern social and economic behaviors. The findings show that traits like honesty and effort are positively associated with GDP per capita, whereas clan-based narratives correlate negatively. These results highlight the enduring influence of cultural traits on economic performance in present-day China. Moreover, the paper's approach offers a transferable framework for studying how cultural narratives influence development, with implications for other regions featuring rich historical legacies.

Keywords: Culture, History, Growth, China

*I thank Andrew Bernard, Andrei Markevich, Joel Mokyr, Moses Shayo, Guido Tabellini and Xiaodong Zhu for their valuable comments. I also thank seminar audiences at Aix-Marseille School of Economics, Fudan University, Helsinki GSE and PKU INSE, as well as audiences at 2023 Harvard Workshop on Historical Kinship, Psychology, and Economic Prosperity, 2024 HKU Conference on Asia and the Global Economy, 2024 KCL Workshop in Comparative Politics and Political Economy, and 2024 Oxford-Tsinghua Conference on Chinese Economic History from the Comparative Perspective for their feedback. I am especially grateful to Runping Zhao and Congshan Shen for their excellent research assistance, and to Jinghan Chang, Shanshan Li, and Dong Xue for their help with an early version of the project, which was conducted before the availability of large language models.

†Assistant Professor at London School of Economics and Political Science (LSE) and Research Affiliate at the Centre for Economic Policy Research (CEPR). Email: m.m.xue@lse.ac.uk

I Introduction

Culture significantly shapes decision-making processes and societal structures, influencing economic performance, governance, and institutional development (Boyd and Richerson, 1985; Putnam, 1993; Boyd and Richerson, 2005; Nunn, 2012). Research has shown that cultural values—such as trust, social capital, individualism, and reciprocal norms—are particularly important in promoting economic growth and political stability, as demonstrated in historical contexts like the Industrial Revolution and in modern development (Tabellini, 2008; Gorodnichenko and Roland, 2017; Besley, 2020).

This perspective aligns with prior research emphasizing how societal beliefs and narratives shape long-run economic choices, social norms, and institutional structures (Bénabou and Tirole, 2006; Fernández, 2011). In particular, internalized cultural values—whether about honesty, effort, or reciprocity—can persist through intergenerational transmission and influence policy preferences, labor-market participation, and collective behaviors.

Comparative analyses of cultural and institutional evolution in Europe and China have also provided critical perspectives on divergent developmental pathways (Greif and Tabellini, 2017; Mokyr and Tabellini, 2024). For instance, Greif and Tabellini (2017) analyze how China's clan-based social structures fostered collective decision-making and hierarchical organization, contrasting with Europe's individualistic frameworks that supported the emergence of democratic governance and market-based economies. Similarly, Mokyr and Tabellini (2024) analyze how differences in social organization shaped technological innovation and economic progress, highlighting the role of culture in driving development.

Despite substantial advancements, systematically measuring cultural traits across time and space remains a significant challenge. This is especially true in the context of China, where the vast geographic and temporal scale, combined with limited data on localized cultural practices, creates a gap in understanding how these traits are distributed and how they relate to modern outcomes. Although a nascent literature has started to examine cultural narratives as a means of capturing enduring beliefs and practices (Michalopoulos and Xue, 2021), efforts to systematically quantify these traits at the local level, particularly within China, remain scarce.

China offers favorable conditions for studying historical cultural values, including extensive written records, diverse local traditions, and relatively recent urbanization. However, much of this potential remains unrealized due to historical barriers to anthropo-

logical research during the socialist period and the lack of systematic efforts to measure cultural traits across regions. Modern datasets, such as the CGSS, provide only limited geographic coverage and offer an incomplete picture of the country's historical cultural traits and their regional variation. As a result, much of China's historical and regional variation in cultural traits remains underexamined.

These gaps are particularly striking given the historical continuity of cultural traits in China. Despite rapid modernization, much of the country's traditional social fabric remains intact due to relatively recent urbanization. Scholars such as Perkins et al. (1975) and Richardson (1999) have highlighted the importance of these enduring cultural traits in shaping local governance, economic behavior, and social interactions. Historical narratives, captured in folktales and other cultural artifacts, provide a crucial lens for understanding these continuities and their modern relevance.

To address these challenges, this study analyzes a dataset of nearly 20,000 folktales collected in the 1980s. These narratives were digitized and standardized, creating a structured dataset suitable for systematic analysis. Using this dataset, the study employs a methodological framework to quantify cultural themes embedded in the folktales. Synthetic archetypes representing key cultural topics—such as honesty, trust, charity, Confucianism, and gift exchange—serve as benchmarks for evaluating thematic similarity.

The thematic analysis leverages BERT embeddings to calculate similarity scores between the collected folktales and these synthetic benchmarks. Additionally, a large language model (LLM) classifies the thematic content of each tale into predefined cultural categories, offering a scalable and efficient alternative method for analyzing and categorizing thematic content. Together, these approaches provide complementary insights into how historical narratives reflect cultural values across regions.

By aggregating the similarity scores at the county level, the study links historical cultural traits to modern attitudes and behaviors using datasets such as the CGSS. Furthermore, it examines the relationship between these cultural traits and regional economic performance, specifically exploring whether historical narratives align with contemporary differences in GDP per capita.

This study makes several contributions. First, it provides a systematic cultural map of China, offering quantified measures of cultural traits across regions using historical narratives. By aggregating cultural themes derived from nearly 20,000 folktales, this research creates a geographically and thematically detailed representation of China's cultural characteristics. This cultural map fills a significant gap in the literature, where within-country

analyses of cultural variation often lack systematic and consistent measures across multiple dimensions of culture. The framework established here enables future studies to explore local cultural traits and their relationships to economic and political phenomena.

Second, the research introduces historical cultural traits as an explanatory factor for regional economic disparities in China. While much of the existing literature attributes these disparities to institutional reforms, infrastructure development, and human capital investment (Lin, 1995; Brandt and Rawski, 2008; Zhu, 2012), this study demonstrates that cultural traits—such as honesty and effort—are also closely associated with contemporary governance practices and economic performance. Large-scale experimental evidence from Falk et al. (2018) highlights the broader significance of such preferences, showing that trust, patience, and reciprocity systematically vary across societies and are strongly linked to long-term economic and institutional outcomes. By incorporating cultural factors into the analysis, the study expands the scope of explanations for China's uneven growth and provides evidence of historical cultural traits' relevance to modern outcomes.

Third, the methodological framework used in this research contributes to the growing field of cultural quantification. While survey-based measures of cultural traits, such as trust or reciprocity, are well-documented, this study demonstrates how qualitative data from historical narratives can be systematically analyzed to quantify cultural themes. By combining machine learning methods—such as BERT embeddings and LLM-based classification—with large-scale historical datasets, the research offers a replicable approach for analyzing qualitative cultural data. This methodology can be applied to other regions, enabling comparative studies of cultural persistence and its economic implications.

Fourth, this study advances research on traditional religions and superstitions by analyzing how these beliefs were embedded in historical narratives and their relevance to modern economic and social outcomes. While much existing literature focuses on monotheistic traditions, this study highlights the role of indigenous Chinese belief systems—including ancestor worship, Buddhist influences, and supernatural themes in folklore—in shaping cultural attitudes. Prior research has shown how religious beliefs influence economic incentives and social structures (Iannaccone, 1998; Barro and McCleary, 2003; Becker, Rubin, and Woessmann, 2024) and how supernatural beliefs can function as informal institutions that regulate behavior and cooperation (Nunn and Sierra, 2017). Von Glahn (1991) and Von Glahn (2004) examine Wutong, a morally ambivalent wealth god, who granted fortune but often at ethically troubling costs, reflecting anxieties over commerce and economic risk in Song China. By quantifying these cultural traits, this study pro-

vides new evidence on their historical persistence and their association with contemporary beliefs and regional economic development.

Finally, the study contributes to broader discussions on how historical cultural traits shape long-term development. Comparative research has examined how cultural and institutional evolution diverged between Europe and China (Greif and Tabellini, 2017; Mokyr and Tabellini, 2024), shaping distinct developmental trajectories. This study shifts the focus to China's regional variation, showing that cultural traits documented in historical narratives remain relevant for understanding contemporary economic differences. By providing systematic evidence of regional cultural variation within a single country, it complements existing cross-country analyses and highlights the enduring influence of historical beliefs and values.

The remainder of this paper is organized as follows. Section II describes the data and methodologies employed in this study. Section III analyzes the values embedded in historical folktales and their connections to contemporary societal attitudes and behaviors. Section IV explores the relationship between these cultural traits and economic growth. Section V discusses other historical sources, such as local gazetteers and poetry, which capture cultural variation over time and could be explored in future research. Section VI concludes by summarizing the findings, discussing their contributions to the literature, and outlining directions for future research. Additional methodological details and supporting materials are included in the Appendix.

II Data and Methodology

This study analyzes cultural traits across regions in China using historical narratives—specifically, folktales. Folklore scholars have long emphasized the utility of oral traditions for uncovering deeply rooted social norms, moral frameworks, and communal values (Dundes, 1980; Bascom, 1965). Folktales, in particular, offer insights into long-standing cultural beliefs that may be less visible in modern surveys, which often capture contemporary attitudes influenced by recent societal shifts. Instead, folktales—transmitted across generations through oral storytelling—tend to preserve older layers of cultural meaning, illuminating how communities once articulated notions of morality, kinship, and collective identity (Lévi-Strauss, 1962; Malinowski, 1926).

Unlike official written histories, which frequently showcase elite or state perspectives, folktales emerge from everyday social contexts. They reflect local concerns, experiences,

and social relations, providing a more grassroots vantage point on how values and norms were negotiated in daily life (Finnegan, 1992). This makes folktales particularly valuable for analyzing persistent cultural traits that cut across time, as well as for examining the historical underpinnings of today's regional differences in social and economic outcomes. Moreover, the regional specificity of Chinese folktale collections—each shaped by local dialects, historical events, and communal identities—allows for granular, county-level comparisons of cultural patterns.

By systematically quantifying the thematic content of folktales, this study addresses a broader need for data that illuminates the historical foundations of cultural traits while facilitating cross-regional analysis. Such an approach is especially relevant in the Chinese context, where large-scale anthropological fieldwork was limited during much of the twentieth century, leaving significant gaps in our understanding of regional cultural variation (Eberhard, 1952). Folktales thus offer a relatively untapped resource for examining the continuity and evolution of cultural values, connecting historical themes—such as reciprocity, cooperation, and hierarchy—to modern social, economic, and political indicators. Through this lens, the study contributes to broader conversations on how cultural legacies shape contemporary outcomes, aligning folklore studies with quantitative research in economics and political science.

II.A Background of the Folktale Collection Project

In the 1980s, China launched a large-scale initiative to systematically document folktales that had been orally transmitted for generations. This project, part of the "Three Collections of Chinese Folk Literature" initiative launched in 1984. It involved extensive fieldwork and collaboration among local storytellers, scholars, and recorders, resulting in a rich archive of narratives that were at risk of being lost amidst rapid modernization.

This period was particularly well-suited for such an effort: China was predominantly rural, with a GDP per capita of less than $350, and much of its population still lived in close connection with traditional lifestyles and practices. The limited urbanization at the time meant that many local customs, belief systems, and oral traditions remained relatively intact. This provided an ideal setting for capturing these stories before the rapid modernization and urbanization of subsequent decades risked eroding these cultural traditions.

The project was driven by a top-down directive and relied on a collaboration between local storytellers, who shared the narratives, and recorders, who transcribed them. This

collaboration ensured both authenticity and standardization in the documentation process. The consistent use of a uniform writing system throughout Chinese history further facilitated the task of transcription and analysis.

Despite the richness and scale of this archive, it has remained largely untapped in academic research. To date, no systematic effort has been made to examine these folktales on a national scale or to explore their potential for understanding historical cultural values, belief systems, or their connections to contemporary societal patterns.

Tales Recorded The folktale collection project spanned all provinces in China, documenting a total of 19,168 tales. Each province contributed a dedicated volume, forming a geographically organized archive. The number of tales contributed by each province ranges from 335 to 1,128.

Most of these tales can be traced to specific counties. Out of the 19,618 total tales, 17,070 are linked to particular counties, while some tales from certain ethnic minority groups cannot be precisely mapped. Since these regions are not included in the current sample, this distinction does not affect the analysis.

The folktales cover diverse themes, including myths, legends, and local stories. Examples range from *Pangu Separates Heaven* and *Earth to The Butterfly Lovers*. On average, each folktale comprises about 981 words in length.

Storytellers and Collectors The collection of folktales relied on contributions from local storytellers, many of whom were peasants living in rural China. Among them, 2,166 were illiterate, and 2,931 had only elementary education. Peasants accounted for a significant portion, contributing 4,149 tales. The average age of the storytellers was 59.5 years.[1] Their predominantly rural backgrounds and illiteracy suggest that the narratives they shared were deeply rooted in oral traditions and largely uninfluenced by modern societal changes or written texts, preserving stories passed down through generations.

In contrast, the recorders of these tales were generally younger and more educated. Of the recorders, 5,412 had attained at least secondary education, with an average age of 36.9 years. This combination of traditional oral storytelling and systematic transcription ensured the preservation of these narratives during a period of rapid modernization.

[1]According to the 1982 census, over 70% of individuals aged 59 were illiterate. This high illiteracy rate makes the limited formal education of most storytellers unsurprising.

II.B From Tales to Values: Mapping China's Cultural Map

The process of analyzing cultural traits embedded in Chinese folktales begins with the digitization and preprocessing of the collected texts. Scanned folktales are converted into machine-readable formats using OCR technology, followed by a rigorous cleaning procedure to standardize terms, resolve inconsistencies, and ensure consistency in linguistic structure. These steps are essential for enabling systematic and scalable analysis.

Studying Chinese folklore presents unique challenges. Unlike Western folklore research, which benefits from standardized indices such as the Aarne-Thompson-Uther (ATU) system (Uther, 2004) or the Berezkin Catalogue (Berezkin, 2015), Chinese folklore lacks comparable classification resources. Although Michalopoulos and Xue (2021) employed the Berezkin Catalogue and ConceptNet (Speer, Chin, and Havasi, 2018) to examine cross-cultural folktale patterns,[2] its scope is insufficient for large-scale, region-specific research. Furthermore, the breadth and complexity of the data demand methods capable of consistently classifying and quantifying cultural traits across thousands of narratives.

To address these gaps, this paper introduces a methodological framework integrating natural language processing (NLP) and machine learning techniques for systematically classifying thematic content. At its core, the methodological framework integrates natural language processing (NLP) and machine learning techniques to systematically classify thematic content. The Synthetic Archetype approach is used to represent predefined cultural themes (e.g., honesty, charity, effort) while reflecting the language and narrative conventions typical of Chinese folklore. Sentence-BERT embeddings (Reimers and Gurevych, 2019), a fine-tuned variant of BERT (Devlin et al., 2019), are then used to calculate semantic similarity scores between each folktale and the corresponding Synthetic Archetypes, capturing both semantic and contextual information.

To account for random fluctuations in text generation and embedding similarity, the Synthetic Archetypes are generated five times, each producing slight variations in the archetypal representations. Similarly, semantic similarity scores between folktales and these archetypes are calculated separately for each iteration. The final measure of thematic alignment is then obtained by averaging across these five versions, ensuring greater stability and reducing the impact of any single run's randomness.

[2]The Berezkin Catalogue can attribute motifs to regions within China but does so at a very coarse level, covering only around 10 traditions for the Han Chinese. This limitation makes it unsuitable for granular, within-country analyses.

A high degree of overlap emerges between the synthetic archetypes and actual folktales, partly reflecting textual similarities. To separate deeper thematic resemblance from purely stylistic overlap, random text is generated to mimic the narrative structure and linguistic patterns of the original tales. Measuring and controlling for the similarity between these randomized texts and the actual folktales isolates the extent to which similarity with the Synthetic Archetypes reflects genuine thematic alignment rather than stylistic coincidence.

Aggregating these similarity scores at the county level offers a regionally detailed view of cultural variation, showing how themes such as trust, charity, or deference to authority manifest geographically. To strengthen the robustness of the analysis, an additional classification approach based on large language models (LLMs) (Brown et al., 2020) is used in parallel. The LLMs assign each folktale to predefined thematic clusters, generating an independent set of classifications at the county level. Comparing the LLM-based classifications to the Synthetic Archetype scores provides a second perspective on the data, increasing confidence in the results by reducing reliance on a single modeling framework.

Validation To ensure the reliability and interpretability of the cultural measures derived from folktales, I conduct two principal checks: (1) examining the consistency of similarity scores across multiple synthetic story versions for the same theme, and (2) comparing cultural trait measurements obtained via the Synthetic Archetype–BERT method with those derived from an independent LLM-based classification approach. These steps confirm that the observed patterns reflect meaningful cultural constructs rather than methodological artifacts.

1. **Consistency Across Synthetic Story Versions.**
   At the tale level, Pearson correlation coefficients are calculated to compare similarity scores generated from different synthetic story versions for each cultural theme. High coefficients indicate that the scoring process yields robust and consistent results across versions. Table I illustrates this for the *Honesty* trait using five synthetic story versions; similar patterns hold for other themes. The high, statistically significant correlations underscore the stability and reliability of the similarity measures.

2. **Alignment with an Independent LLM-Based Classification.**
   Next, I compare two independent measures of cultural traits: (a) the Synthetic Archetype–BERT similarity scores and (b) an LLM-based classification of folktales

Table I: Tale-Level Correlation for the Honesty Theme Across Five Versions

	Version 1	Version 2	Version 3	Version 4	Version 5
Version 1	1.0	0.7331***	0.8366***	0.9319***	0.8199***
Version 2	0.7331***	1.0	0.7463***	0.7164***	0.7814***
Version 3	0.8366***	0.7463***	1.0	0.8443***	0.8600***
Version 4	0.9319***	0.7164***	0.8443***	1.0	0.8191***
Version 5	0.8199***	0.7814***	0.8600***	0.8191***	1.0

into predefined cultural themes. Table II presents Pearson correlation coefficients for each cultural trait at the tale level, showing statistically significant associations (all at $p < 0.01$) between the two sets of measures. This consistency indicates that the Synthetic Archetype–BERT approach and the LLM-based classification detect broadly similar patterns in cultural content.

Table II: Correlations between Synthetic and LLM-Based Cultural Trait Measures

Cultural Trait	**Correlation**
Honesty	0.1663 ***
Trust	0.1555 ***
Effort	0.2818 ***
Equality	0.1014 ***
Buddhism	0.1495 ***
Ancestor Worship	0.1079 ***
Clan	0.2553 ***
Gift Exchange	0.1584 ***
Benevolent Emperor	0.2694 ***

Note: Correlations are computed using `pwcorr` for the tale-level measures. Significance levels: *** for $p < 0.01$, ** for $p < 0.05$, and * for $p < 0.10$.

II.C Folktales of Ethnic Minorities

The original dataset includes folktales from China's 56 officially recognized ethnicities, but this study focuses exclusively on narratives from Han Chinese regions. Although material from ethnic minority groups was digitized as part of the broader project, it is not analyzed here. This decision reflects the study's objective of constructing a cultural map based on Han Chinese traditions, which align with the scope and methodological framework employed.

The exclusion of ethnic minority narratives stems from both practical and methodological considerations. The diversity and complexity of these groups would require tai-

lored approaches to effectively analyze their distinct traditions and narratives. Furthermore, the scarcity of supplementary data, such as survey responses, constrains the ability to rigorously connect these narratives to other variables. For instance, the Chinese General Social Surveys (CGSS) provide information for only six ethnic groups, limiting the feasibility of robust analyses involving ethnic minority folktales.

III Cultural Narratives in Folktales and Their Contemporary Relevance

The folktales collected during the 1980s, many of which originated centuries ago, provide a unique lens through which to examine the cultural values and beliefs of historical Chinese society. These narratives reflect a broad spectrum of themes, ranging from moral values that are widely recognized across societies, such as honesty, prosociality, effort, and equality, to those that have been particularly prominent in China's historical and philosophical traditions, including Buddhism, ancestor worship, and gift exchange.

This section investigates how historical cultural narratives relate to contemporary values and attitudes, using responses from the Chinese General Social Survey (CGSS) to measure modern beliefs and behaviors. By linking historical narratives to modern survey responses, the analysis sheds light on the persistence and evolution of cultural traits across time and space. To examine these connections, I estimate the following equation:

$$\mathit{Belief}_{i,c} = \alpha + \beta\, \mathit{Trait}_c + \gamma\, \mathit{StyleControl}_c + \mathbf{X}_c^H\Omega + \mathbf{X}_c^G\Lambda + \mathbf{X}_c^C\Pi + \mathbf{X}_i\Gamma + \epsilon_{i,c}, \quad (1)$$

where c denotes a county, and $\mathit{Belief}_{i,c}$ represents individual beliefs or values from the CGSS. The term Trait_c is the average similarity of folktales in county c to a particular cultural theme (e.g., trust, honesty, or charity), computed as the average across five iterations of similarity calculations. To account for purely stylistic overlap between folktales and the synthetic archetypes, $\mathit{StyleControl}_c$ represents the average similarity of the county's folktales to randomly generated texts (see Appendix A.2 for details).

$\mathbf{X}_c^H$ and $\mathbf{X}_c^G$ are vectors of historical and geographic controls, while $\mathbf{X}_c^C$ represents contemporary controls, such as economic and institutional variables. Individual characteristics like age and gender are included in $\mathbf{X}_i$.

Historical and geographical controls include factors such as agricultural suitability, trade networks, population density, terrain ruggedness, proximity to Confucius's birthplace, and treaty port status, along with fixed effects for historical macroregions. Con-

temporary controls include GDP per capita, urbanization, sectoral composition, education levels, ethnic composition, provincial capital status, and fixed effects for modern macroregions.

This analysis focuses on China proper and excludes ethnic minority regions. Figure A.2 displays the geographic distribution of the sample after omitting areas outside China proper and heavily ethnic regions.

III.A Moral Values and Worldviews

This subsection examines the relationship between folktale narratives and modern attitudes toward honesty, prosociality, equality, and effort, exploring how cultural themes in historical folktales relate to contemporary beliefs and behaviors in these areas.

Honesty Honesty and trust are foundational cultural values that underpin reliable interpersonal interactions and promote social cohesion. Algan and Cahuc (2010) underscore the importance of these traits in fostering social capital, while Cohn, Fehr, and Maréchal (2014) highlight their role in reducing transaction costs. I examine the alignment between folktale narratives that emphasize honesty and trust and modern attitudes toward vigilance against exploitation, as measured by the survey question, "Do you agree that people will take advantage of you if you're not careful?"

As shown in Columns 1 and 2 of Table III, counties where folktales more closely align with honesty and trust exhibit stronger agreement with the need for vigilance against exploitation. These findings reflect the alignment between traditional narratives and contemporary attitudes toward interpersonal caution.

Charity and Prosocial Behavior Prosocial behaviors, such as altruism and community engagement, are vital for maintaining social cohesion and resilience. Andreoni (1990) discuss the motivations behind altruistic behaviors, including the "warm-glow" effect of giving, while Hamilton (1964) frame prosociality within evolutionary contexts, emphasizing its role in fostering cooperation and reciprocal support. I explore the connection between prosocial themes in folktales and survey questions such as, "Would you intervene if children were damaging public property?" and "Can you borrow tools easily from your neighbors?".

As shown in Columns 3 and 4 of Table III, counties where folktales align more closely with prosociality exhibit a higher likelihood of intervening in public misbehavior and greater ease in borrowing tools from neighbors. These findings reflect the alignment

Table III: Honesty, Prosociality, Equality and Effort

	(1)	(2)	(3)	(4)	(5)	(6)
Honesty Narrative	-0.88**					
	(0.37)					
Trust Narrative		-0.97***				
		(0.30)				
Charity Narrative			1.35***	0.71*		
			(0.39)	(0.39)		
Effort Narrative					1.08**	
					(0.41)	
Equality Narrative						2.01**
						(0.91)
Age and Gender	Y	Y	Y	Y	Y	Y
Contemporary Controls	Y	Y	Y	Y	Y	Y
Historical Controls	Y	Y	Y	Y	Y	Y
Geographical Controls	Y	Y	Y	Y	Y	Y
R^2	0.017	0.017	0.053	0.058	0.049	0.069
Observations	21929	21929	4859	4859	4876	2426
Mean of Dep. Var.	3.03	3.03	2.76	2.79	4.40	3.86

Notes: The table shows the associations between folklore traits and measures of contemporary attitudes from the Chinese General Social Survey (CGSS). The dependent variables are based on responses to specific questions reflecting themes of trust and honesty, charity, effort, and equality. Columns 1 and 2 assess trust and honesty through responses to the statement, "In this society, if you're not careful, others will take advantage of you," with honesty- and trust-related folklore as key predictors. Columns 3 and 4 examine charity, focusing on responses to the question, "If children in your community or village were damaging flowers or public property, would you stop them?" and the ease of borrowing tools like a wrench or screwdriver from neighbors. Columns 5 explore beliefs about personal success, asking whether it is largely due to effort. Column 6 evaluates agreement with the statement, "Social wealth distribution is unfair, with excessive disparity," reflecting attitudes toward equality. The independent variables are the average similarity of folktales in a county to traits of honesty, trust, charity, effort, and equality. Controls include contemporary, historical, and geographic factors. Standard errors are clustered at the county level and are shown in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

between prosocial narratives and modern communal behaviors.

Effort Effort is widely recognized as critical for overcoming adversity and achieving success. Becker (1964) discuss effort as a crucial component of human capital development, while Kaplan, Lancaster, and Robson (2000) emphasize its adaptive significance in resource acquisition and planning. The relationship between effort-focused narratives

and modern beliefs about hard work is assessed using CGSS responses to the question, "Do you agree that personal achievements are largely due to hard work?" Column 5 of Table III reveals that individuals in counties where folktales align more closely with effort themes report stronger agreement with the importance of hard work. These results reflect the enduring emphasis on perseverance found in traditional narratives.

Equality Equality is a key value for promoting fairness and societal stability. Alesina and Rodrik (1994) connect perceptions of equality to sustainable economic growth and political stability, while Boehm (1999) discuss its role in reducing dominance hierarchies and fostering group cohesion. The relationship between equality-focused narratives and modern beliefs about fairness is assessed using CGSS responses to the question, "Social wealth distribution is unfair, with excessive disparity." Column 6 of Table III shows that individuals in counties where folktales align more closely with equality themes report heightened awareness of wealth disparities and stronger support for redistributive policies. These findings suggest a relationship between equality-related narratives and modern attitudes toward fairness.

The analysis shows significant relationships between folktale themes related to honesty, prosociality, equality, and effort and corresponding attitudes in modern survey responses. Counties with honesty-focused narratives are associated with heightened caution against exploitation, while prosocial narratives correspond to stronger engagement in communal behaviors, such as intervening in public misbehavior or borrowing tools from neighbors. Equality-focused narratives are linked to heightened awareness of wealth disparities and stronger preferences for redistributive policies. Finally, effort-focused narratives correspond to stronger beliefs in hard work and perseverance as drivers of success. These findings highlight how historical narratives correspond with regional variations in contemporary beliefs and behaviors.

III.B Traditional Beliefs, Faiths and Customs

This subsection examines the correspondence between cultural narratives tied to Confucianism, Buddhism, ancestor worship, and gift exchange and contemporary attitudes and practices.

Buddhism Buddhist beliefs, particularly ideals such as nirvana, have shaped moral and spiritual frameworks across many regions. Keown (1992) emphasize the role of Buddhist teachings in ethical reasoning, while Gombrich (2006) explore their influence on spiritual

practices. I identify Buddhist narratives in folktales and assess their association with contemporary Buddhist attitudes as measured in the CGSS. Specifically, Column 1 of Table IV captures respondents' religious affiliation with Buddhism ("Your religious belief – Buddhism"), and Column 2 measures their agreement with the concept of Nirvana ("Do you agree with the concept of Nirvana?" with responses: 1 = Agree, 2 = Neutral, 3 = Disagree). As shown in Table IV, stronger Buddhist narratives are significantly associated with higher identification with Buddhism (Column 1) and a lower likelihood of disagreement with Nirvana (Column 2), indicating greater acceptance of this spiritual ideal.

Ancestor Worship Ancestor worship emphasizes the supernatural influence of ancestors and their role in ensuring familial continuity. Watson (1988) highlight its significance in Chinese cultural practices, while Freedman (1979) discuss its centrality to family structures and rituals. I assess narratives related to ancestor worship by comparing them with respondents' beliefs in ancestral supernatural powers, as measured in the CGSS by the question "Do you agree that ancestors have supernatural powers?" (Responses: 1 = Agree, 2 = Neutral, 3 = Disagree). As shown in Column 3 of Table IV, there is a significant negative association between ancestor worship narratives and disagreement with the belief in ancestral supernatural powers. This result suggests that individuals in counties with stronger ancestor worship narratives are more inclined to accept the supernatural influence of their ancestors.

Clan Clan narratives highlight the significance of kinship networks and the informal institutions that emerge from them. Research by Greif and Tabellini (2017) suggests that clan structures historically provided alternative mechanisms to formal institutions by fostering localized cooperation and collective problem-solving. In his work, Zhang (2017) examines how kinship ties and property rights arrangements in preindustrial China were deeply intertwined with familial obligations and intergenerational support. Building on this framework, I identify clan-related themes in folktales and assess their association with contemporary familial attitudes as measured in the CGSS. Specifically, Column 4 of Table IV captures attitudes regarding filial responsibility in caring for aging parents, while Column 5 reflects respondents' willingness to promote beneficial actions among kin and friends. The results indicate that in counties where folktales more strongly reflect clan narratives, there is significantly higher endorsement of intergenerational care and family-based support, underscoring the enduring influence of traditional kinship norms

Table IV: Traditional Beliefs, Faiths and Customs

	(1)	(2)	(3)	(4)	(5)	(6)
Buddhist Narrative	0.359**	-1.228*				
	(0.142)	(0.657)				
Ancestor Narrative			-1.698*			
			(0.861)			
Clan Narrative				0.832**	2.418**	
				(0.371)	(0.926)	
Gifting Narrative						2.393**
						(0.985)
[2em] Age and Gender	Y	Y	Y	Y	Y	Y
Contemporary Controls	Y	Y	Y	Y	Y	Y
Historical Controls	Y	Y	Y	Y	Y	Y
Geographical Controls	Y	Y	Y	Y	Y	Y
R^2	0.0901	0.0564	0.0720	0.0687	0.0415	0.103
Observations	17095	1225	1599	19345	4855	4799
Mean of Dep. Var.	0.0533	3.833	3.683	0.535	3.203	

Note: This table examines the associations between traditional narratives and contemporary beliefs and practices as measured by survey responses. The dependent variables capture respondents' attitudes toward various cultural and traditional practices. Column 1 is based on the question "Your religious belief – Buddhism," Column 2 on "Do you agree with the concept of Nirvana?," column 3 on "Do you agree with the concept of ancestral supernatural powers?," column 4 on the belief that children should care for their aging parents, column 5 on the willingness to promote beneficial actions among kin and friends as captured by "I will try to ensure that my family and friends do what I believe is beneficial for them, even if they do not agree," and column 6 on the perceived financial burden of gift exchange, measured by "How much pressure does gift-giving expenditure place on your household?" The independent variables are the average similarity of county folktales to narratives associated with Buddhism, ancestor worship, clan affiliation, and gift exchange. Control variables include demographic factors (e.g., age and gender), contemporary factors, historical factors, and geographic factors. Standard errors are clustered at the county level and are shown in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

on modern familial obligations.

Gift Exchange Gift exchange practices, rooted in themes of reciprocity and social ties, are central to maintaining community bonds, particularly in traditional and rural societies. Mauss (1954) describe gifting as a mechanism for fostering mutual obligations, while Yang (1994) emphasize its role in reinforcing social networks. I measure the perceived financial burden of gift exchange based on survey responses assessing the level of financial pressure caused by gift-giving expenses. As shown in Column 6 of Table IV, there is a significant positive association between gifting narratives in folktales and the

reported financial burden of gift exchange, highlighting the alignment between historical narratives emphasizing reciprocity and contemporary experiences of social obligation in gifting practices.

The findings demonstrate associations between folktale narratives tied to Buddhism, ancestor worship, clan, and gift exchange and contemporary cultural attitudes and practices. Specifically, Buddhist narratives are associated with stronger religious identification and greater acceptance of Nirvana; ancestor worship narratives correspond to a higher likelihood of accepting ancestral supernatural powers; clan narratives align with enhanced familial trust and intergenerational obligations; and gifting narratives are positively linked to modern gifting practices. These results underscore the correspondence between historical narratives and contemporary cultural values.

III.C Attitudes Towards Centralized Authority

Trust in state institutions is critical for governance and political stability, shaping public compliance with policies, cooperation in state-led initiatives, and perceptions of governmental legitimacy. Scholars like Levi (1998) and Rothstein (2011) emphasize its importance in reducing transaction costs and fostering collective action.

We examine how narratives of a benevolent emperor—an embodiment of centralized authority in Imperial China—relate to regional differences in trust in state institutions. The cultural theme of benevolence centers on the emperor as a compassionate and just figure, symbolizing the ultimate source of order and stability. Folktales under this theme often depict the emperor as a protector of the people, reinforcing loyalty to the central government while downplaying the influence of local or intermediary authorities.

This analysis uses CGSS responses measuring trust in the central government, local governments, the military, the judicial system, and the police. As shown in Table V, individuals in counties featuring folktales that portray the emperor as a benevolent ruler exhibit greater trust in state institutions, particularly in the central government (column 1). There are also positive associations with trust in the military (column 3) and the police (column 4). However, no significant relationship is found for trust in local governments (column 2) or the judicial system (column 5).

The observed patterns suggest that folktales emphasizing the benevolence of the emperor correspond to greater trust in national institutions—such as the central government, military, and police—while showing no association with trust in local governance or the

Table V: Trust in State Institutions

	(1)	(2)	(3)	(4)	(5)
	Central Government	Local Government	Military	Police	Judicial System
Benevolent Emperor Narrative	1.350**	0.560	1.515**	1.946*	1.360
	(0.613)	(1.230)	(0.648)	(1.005)	(0.976)
[1em] Age and Gender	Y	Y	Y	Y	Y
Contemporary Controls	Y	Y	Y	Y	Y
Historical Controls	Y	Y	Y	Y	Y
Geographical Controls	Y	Y	Y	Y	Y
Adjusted R^2	0.129	0.0614	0.123	0.0981	0.0796
Observations	4873	4864	4861	4869	4859
Mean of Dep. Var.	4.393	3.600	4.373	3.919	3.907

Note: The table examines the associations between beliefs about emperors, as reflected in local folktales, and levels of trust in various institutions, using survey-based measures. The dependent variables reflect trust in the central government (Column 1), local government (Column 2), the military (Column 3), the police (Column 4), and the judicial system (Column 5). The key independent variable is the average similarity of local folktales to narratives about emperors. Control variables include demographic, contemporary, historical, and geographic factors, such as age, gender, GDP per capita, historical population density, and distance to Confucius's birthplace. Standard errors are clustered at the county level and are shown in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

judiciary. The positive associations align with Weber's theory of traditional authority, where charismatic rulers confer legitimacy on overarching state institutions, reinforcing centralized rule (Weber, 1978). The lack of association with local government trust is consistent with arguments by Tsai (2007), who suggest that centralized authority can overshadow local governance, reducing its visibility and accountability. Similarly, the absence of an effect on the judiciary reflects its historical role in Imperial China as an extension of the bureaucratic state rather than an independent institution.[3]

IV Cultural Traits and Economic Performance

This section examines the relationship between cultural traits, as derived from folktales, and regional economic outcomes, with a focus on GDP per capita in 2010. Cultural narratives provide an alternative lens for understanding regional disparities in economic performance, complementing traditional explanations centered on institutional quality,

[3]The lack of association with the judiciary is unsurprising, as historically, the judicial system in Imperial China was not a separate branch of government but rather a function of the bureaucratic state, administered by officials as part of their broader governing responsibilities. modern China, the judiciary is formally a separate branch of government, though it remains structurally subordinate to the Chinese Communist Party, limiting its independence in practice.

geographic endowments, and resource availability.

To formally investigate this relationship, I estimate the following equation:

$$\textit{GDP per capita}_c = \alpha + \sum_k \beta_k \, \textit{Trait}_{c,k} + \gamma \, \textit{StyleControl}_c + \mathbf{X}_c^C \Gamma + \mathbf{X}_c^G \Lambda + \mathbf{X}_c^H \Omega + \epsilon_c, \tag{2}$$

where $\textit{GDP per capita}_c$ is the log of GDP per capita in county c, and $\textit{Trait}_{c,k}$ represents the set of cultural traits derived from folktales (e.g., honesty, effort, ancestor worship), computed as the average semantic similarity between county c's folktales and the corresponding Synthetic Archetypes, across five iterations of similarity calculations. The term $\textit{StyleControl}_c$ is a county-level measure of average similarity between real folktales and randomly generated text, designed to capture stylistic overlap rather than substantive thematic content (see Section A.2).

$\mathbf{X}_c^C$, $\mathbf{X}_c^G$, and $\mathbf{X}_c^H$ are vectors of contemporary, geographic, and historical control variables, respectively. The set of controls remains consistent with the specification in Section III, ensuring comparability, but $\textit{GDP per capita}_c$ is now analyzed as the dependent variable rather than a control. The coefficients β_k capture the marginal effects of cultural traits on GDP per capita, and γ measures the contribution of stylistic similarity, while ϵ_c is the error term.

I implement Lasso regression to identify the subset of cultural traits and control variables most predictive of GDP per capita. This approach penalizes less relevant predictors, enhancing model interpretability and robustness. The analysis includes both a full sample of counties and a restricted sample of counties with at least five folktales, to ensure sufficient data coverage for the cultural trait measures.

Table VI presents the results. The estimates indicate that honesty narratives are consistently associated with higher GDP per capita across all three specifications. In the full sample (column 1), as well as in the restricted samples including only counties with at least three tales (column 2) and at least five tales (column 3), honesty maintains a positive relationship with economic performance. These findings align with research emphasizing the roles of trust and honesty in reducing transaction costs and fostering productivity (Putnam, 1993; Tabellini, 2008; Guiso, Sapienza, and Zingales, 2011; Becker, 1964).

By contrast, equality narratives exhibit a negative association with GDP per capita across all three specifications, supporting arguments that an emphasis on redistribution and economic leveling may be inversely related to economic performance (Okun, 1975;

Table VI: Lasso-Selected Cultural Traits Predicting GDP Per Capita (2010)

Cultural Trait	**Any Tale**	**At Least 3 Tales**	**At Least 5 Tales**
Honesty	0.0242	0.0732	0.0779
Effort	-0.0202	–	–
Equality	-0.0093	-0.0804	-0.0804
Buddhism	-0.0019	-0.0031	-0.0031
Ancestor Worship	–	–	–
Clan	-0.0367	–	-0.0002
Gift Exchange	0.0129	0.0072	0.0121
Benevolent Emperor	0.0026	-0.0114	-0.0193

Note: The table reports standardized coefficients from Lasso regressions predicting GDP per capita in 2010. The first column includes all counties with at least one folktale. The second column restricts the sample to counties with at least three folktales, and the third column further restricts the sample to counties with at least five folktales. Positive coefficients indicate a positive association with GDP per capita, while negative coefficients indicate an inverse association. Traits not selected by the Lasso procedure are indicated by "– ". Traits that were not selected in any of the three models are excluded from the table. All models include controls for style similarity, contemporary conditions, geographic variables, and historical factors.

Alesina and Rodrik, 1994). This suggests that cultural narratives prioritizing equality may reflect norms that dampen incentives for economic competition and investment, potentially hindering long-run development.

Clan narratives appear in the full sample (column 1) and in the restricted sample with at least five folktales (column 3), with negative coefficients in both cases. This points to a possible inverse relationship between strong kinship ties and economic performance. These findings align with research by Enke (2019) and Schulz et al. (2019), who show that intensive kinship structures can strengthen in-group cooperation but may reduce impersonal trust and broader economic exchange, potentially limiting regional economic dynamism.

Gift exchange narratives show a small but positive association with GDP per capita in all three specifications, suggesting that norms around reciprocity and social obligations in economic transactions may have some relevance for economic outcomes, though the magnitude of the effect is modest.

Benevolent emperor narratives consistently exhibit negative associations with GDP per capita across all three specifications. This suggests that historical narratives reinforcing centralized authority and moralized rulership may be linked to weaker economic outcomes, potentially through their effects on attitudes toward hierarchy, state control,

and economic autonomy. This result aligns with perspectives on the economic drawbacks of strong centralized authority (Hayek, 1945; Besley and Persson, 2010).

Buddhism and ancestor worship narratives are largely not selected in the Lasso models, suggesting no robust association between these cultural themes and contemporary regional economic performance.

Overall, the results suggest that certain cultural traits—honesty and, to a lesser extent, gift exchange—are positively linked to economic performance, while equality and clan narratives tend to be negatively related to growth. Benevolent emperor narratives consistently exhibit negative associations with GDP per capita, reinforcing the notion that historical cultural beliefs about governance and authority have lingering effects on economic outcomes. By incorporating cultural values into the study of regional economic performance, this analysis provides new insights into the cultural determinants of China's uneven development.

V Enduring Narratives and Historical Chronicles

This study uses folktales as a key data source to construct a cultural map of China, identifying regional variations in values and beliefs that are linked to economic outcomes. While this approach offers a powerful framework for examining culture through narrative traditions, other sources—such as local gazetteers and historical poetry—could also be employed to extract cultural indicators and broaden our understanding of historical cultural traits.

Local gazetteers, compiled over centuries, contain detailed records of regional practices, customs, and social organization. These documents could complement the folktales analyzed here by providing additional pathways to quantify cultural variation across both space and time. Historical poetry, with its rich linguistic and thematic content, likewise illuminates cultural values, social norms, and philosophical ideals that shaped different eras and locales. By examining poetry, researchers may uncover how cultural values were articulated, transmitted, and, potentially, transformed within literary traditions.

Furthermore, collecting and analyzing these sources at multiple points in time would allow for investigation into the origins and evolution of cultural narratives. Such an approach can reveal how deeply rooted beliefs emerged, were diffused, and persisted or changed across historical periods.

Xue (2024) uses both poetry and local gazetteers to derive cultural indicators, focusing

on customs, norms, and attitudes relating to women. Although these textual sources have so far been applied to a relatively narrow subset of cultural traits, their broader potential remains largely untapped. Expanding their use to include a wider range of themes could yield additional insights into historical cultural variation throughout China.

Because cultural traits often persist across generations—shaped by environmental stability and embedded social structures—narrative sources such as folktales, gazetteers, and poetry provide a powerful lens for studying cultural continuity (Giuliano and Nunn, 2021). In a companion project, I incorporate these diverse sources to investigate how historical cultural traits manifest and relate to both historical and contemporary economic and political outcomes. This complementary research underscores how cultural narratives serve as a bridge between past and present societal dynamics, offering insights into where such beliefs originated and how they have endured.

VI Conclusion

This study represents the first systematic effort to construct a cultural map of China by analyzing nearly 20,000 folktales collected in the 1980s. Through this large-scale examination of historical narratives, it uncovers important regional differences in values, mindsets, and beliefs, and links cultural traits—such as trust, honesty, and reciprocity—to contemporary economic outcomes. The results provide robust evidence that culture plays a key role in shaping recent economic growth in China, with traits like honesty, effort, and gift exchange showing strong positive correlations with GDP per capita, while ancestor worship exhibits a consistently negative association.

In addition to contributing to the literature on culture and economics, the findings offer fresh insights into the drivers of China's economic rise by connecting historical cultural narratives to modern economic performance. This perspective underscores the significance of deep-rooted cultural determinants in understanding regional and national growth patterns.

Prior attempts to classify Chinese folklore have been constrained by limited geographic coverage and insufficient regional specificity. By introducing systematic methods for analyzing folktales and their cultural content, this study opens up new opportunities to examine the historical development of cultural traits and their impact on present-day economic and social phenomena. Furthermore, the methodology is readily transferable to other regions with rich cultural archives, providing a valuable framework for exploring

the role of cultural narratives in shaping economic outcomes.

References

Alesina, Alberto, and Dani Rodrik. 1994. "Distributive Politics and Economic Growth." *Quarterly Journal of Economics* 109 (2): 465–490.

Algan, Yann, and Pierre Cahuc. 2010. "Inherited Trust and Growth." *American Economic Review* 100 (5): 2060–2092.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving." *Economic Journal* 100 (401): 464–477.

Barro, Robert J., and Rachel McCleary. 2003. *Religion and Economic Growth.* NBER Working Papers 9682. National Bureau of Economic Research, Inc, May. http://ideas.repec.org/p/nbr/nberwo/9682.html.

Bascom, William. 1965. "Four Functions of Folklore." *The Journal of American Folklore* 67 (266): 333–349.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* Chicago: University of Chicago Press.

Becker, Sascha O, Jared Rubin, and Ludger Woessmann. 2024. "Religion and Growth." *Journal of Economic Literature* 62 (3): 1094–1142.

Bénabou, Roland, and Jean Tirole. 2006. "Incentives and Prosocial Behavior." *American Economic Review* 96 (5): 1652–1678.

Berezkin, Yu E. 2015. "Folklore and Mythology Catalogue: Its Lay-Out and Potential for Research." *The Retrospective Methods Network,* no. S10, 58–70.

Besley, Timothy. 2020. "State capacity, reciprocity, and the social contract." *Econometrica* 88 (4): 1307–1335.

Besley, Timothy, and Torsten Persson. 2010. "State Capacity, Conflict, and Development." *Econometrica* 78, no. 1 (January): 1–34.

Boehm, Christopher. 1999. *Hierarchy in the Forest: The Evolution of Egalitarian Behavior.* Cambridge, MA: Harvard University Press.

Boyd, Robert, and P. J. Richerson. 1985. *Culture and the Evolutionary Process.* Chicago: University of Chicago Press.

———. 2005. *The Origin and Evolution of Cultures.* Oxford University Press.

Brandt, Loren, and Thomas G Rawski. 2008. *China's Great Economic Transformation.* Cambridge university press.

Brown, Tom, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. "Language models are few-shot learners." *Advances in neural information processing systems* 33:1877–1901.

Cohn, Alain, Ernst Fehr, and Michel A. Maréchal. 2014. "Business Culture and Dishonesty in the Banking Industry." *Nature* 516:86–89.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." In *Proceedings of NAACL-HLT,* 4171–4186.

Dundes, Alan. 1980. *Interpreting Folklore.* Bloomington: Indiana University Press.

Eberhard, W. 1952. *Chinese Festivals.*

Enke, Ben. 2019. "Kinship, Cooperation, and the Evolution of Moral Systems." *The Quarterly Journal of Economics* 134 (2): 953–1019.

Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde. 2018. "Global Evidence on Economic Preferences." *The Quarterly Journal of Economics* 133 (4): 1645–1692.

Fernández, Raquel. 2011. "Does Culture Matter?" *Handbook of social economics* 1:481–510.

Finnegan, Ruth. 1992. *Oral Traditions and the Verbal Arts: A Guide to Research Practices.* London: Routledge.

Freedman, Maurice. 1979. *The Study of Chinese Society: Essays by Maurice Freedman.* Stanford, CA: Stanford University Press.

Giuliano, Paola, and Nathan Nunn. 2021. "Understanding Cultural Persistence and Change." *The Review of Economic Studies* 88 (4): 1541–1581.

Gombrich, Richard. 2006. *Theravāda Buddhism: A Social History from Ancient Benares to Modern Colombo.* London: Routledge.

Gorodnichenko, Yuriy, and Gerard Roland. 2017. "Culture, Institutions, and the Wealth of Nations." *Review of Economics and Statistics* 99 (3): 402–416.

Greif, Avner, and Guido Tabellini. 2017. "The Clan and the Corporation: Sustaining Cooperation in China and Europe." *Journal of Comparative Economics* 45 (1): 1–35.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales. 2011. "Civic Capital as the Missing Link." In *Handbook of Social Economics,* edited by Jess Benhabib, Alberto Bisin, and Matthew O. Jackson, 417–480. Amsterdam: Elsevier.

Hamilton, W. D. 1964. "The Genetical Evolution of Social Behaviour." *Journal of Theoretical Biology* 7:1–16.

Hayek, F.A. 1945. "The use of knowledge in society." *The American Economic Review* 35:519–530.

Iannaccone, Laurence R. 1998. "Introduction to the Economics of Religion." *Journal of Economic Literature* 36, no. 3 (September): 1465–1495.

Kaplan, H. S., J. Lancaster, and A. Robson. 2000. "The Evolution of Intelligence and the Human Life History." *Proceedings of the National Academy of Sciences* 97 (9): 6351–6356.

Keown, Damien. 1992. *The Nature of Buddhist Ethics.* London: Macmillan.

Levi, Margaret. 1998. *Consent, Dissent, and Patriotism.* Cambridge: Cambridge University Press.

Lévi-Strauss, Claude. 1962. *The Savage Mind.* Originally published in French as *La Pensée Sauvage.* Chicago: University of Chicago Press.

Lin, Justin Yifu. 1995. "The Needham Puzzle: Why the Industrial Revolution Did Not Originate in China." *Economic development and cultural change,* 269–292.

Malinowski, Bronislaw. 1926. *Myth in Primitive Psychology.* New York: W. W. Norton Company.

Mauss, Marcel. 1954. *The Gift: Forms and Functions of Exchange in Archaic Societies.* London: Cohen West.

Michalopoulos, Stelios, and Melanie Meng Xue. 2021. "Folklore." *The Quarterly Journal of Economics* (January).

Mokyr, Joel, and Guido Tabellini. 2024. "Social Organizations and Political Institutions: Why China and Europe Diverged." *Economica* 91 (362): 347–382.

Nunn, Nathan. 2012. "Culture and the Historical Process." *Economic History of Developing Regions* 27:S108–S126.

Nunn, Nathan, and Raul Sanchez de la Sierra. 2017. "Why Being Wrong Can Be Right: Magical Warfare Technologies and the Persistence of False Beliefs." *American economic review* 107 (5): 582–587.

Okun, Arthur M. 1975. *Equality and Efficiency: The Big Tradeoff.* Washington, D.C.: Brookings Institution Press. ISBN: 9780815764755.

Perkins, Dwight, et al. 1975. *China's Modern Economy in Historical Perspective.* Stanford University Press.

Putnam, Robert D. 1993. *Making Democracy Work. Civic Traditions in Modern Italy.* Princeton, NJ: Pinceton University Press.

Reimers, Nils, and Iryna Gurevych. 2019. "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks." In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP),* 3982–3992.

Richardson, Philip. 1999. *Economic Change in China, c. 1800-1950.* Cambridge University Press.

Rothstein, Bo. 2011. *The Quality of Government: Corruption, Social Trust, and Inequality in International Perspective.* Chicago, IL: University of Chicago Press.

Schulz, Jonathan F., Duman Bahrami-Rad, Jonathan Beauchamp, and Joseph Henrich. 2019. "The Church, Intensive Kinship, and Global Psychological Variation." *Science* 366 (6466): eaau5141.

Speer, Robyn, Joshua Chin, and Catherine Havasi. 2018. *ConceptNet 5.5: An Open Multilingual Graph of General Knowledge.* arXiv: 1612.03975 [cs.CL]. https://arxiv.org/abs/1612.03975.

Tabellini, Guido. 2008. "Presidental Address Institutions and Culture." *Journal of the European Economic Association* 6 (2-3): 255–294.

Tsai, Lily L. 2007. *Accountability Without Democracy: Solidary Groups and Public Goods Provision in Rural China.* Cambridge: Cambridge University Press.

Uther, Hans-Jörg. 2004. *The Types of International Folktales: A Classification and Bibliography. Based on the System of Antti Aarne and Stith Thompson.* Vol. 284-286. FF Communications. Helsinki: Academia Scientiarum Fennica.

Von Glahn, Richard. 1991. "The Enchantment of Wealth: the God Wutong in the Social History of Jiangnan." *Harvard Journal of Asiatic Studies* 51 (2): 651–714.

———. 2004. *The sinister way: The Divine and the Demonic in Chinese Religious Culture.* University of California Press.

Watson, James L. 1988. "The Structure of Chinese Funerary Rites: Elementary Forms, Ritual Sequence, and the Primacy of Performance." In *Death Ritual in Late Imperial and Modern China,* edited by James L. Watson and Evelyn S. Rawski, 3–19. Berkeley: University of California Press.

Weber, Max. 1978. *Economy and Society: An Outline of Interpretive Sociology.* Vol. 1. University of California press.

Xue, Melanie Meng. 2024. *High-Value Work and the Rise of Women: The Cotton Revolution and Gender Equality in China.* Working paper.

Yang, Mayfair Mei-Hui. 1994. *Gifts, Favors, and Banquets: The Art of Social Relationships in China.* Ithaca, NY: Cornell University Press.

Zhang, Taisu. 2017. *The Laws and Economics of Confucianism: Kinship and Property in Preindustrial China and England.* Cambridge University Press.

Zhu, Xiaodong. 2012. "Understanding China's Growth: Past, Present, and Future." *Journal of Economic Perspectives* 26 (4): 103–124.

Appendix

1 Data

A Geographic Distribution of Folktales

Figure A.1 displays the geographic distribution of all folktales collected by the Folklore Collection Project. Each blue dot represents a county with recorded folktales, and the size of the dot corresponds to the number of tales documented in that county. The map provides a comprehensive visualization of the spatial coverage of folktales across China.

Note that this map includes *all* folktales in the dataset, not just those in the final analytical sample. While the sample used in this study focuses on China proper and regions historically inhabited by the Han Chinese, the figure illustrates the broader geographic scope of the folktale collection effort.

Additionally, some ethnic minority folktales could not be reliably assigned to specific locations, and thus are not shown. Folktales from ethnic minorities that lack clear county-level identifiers are excluded from both the figure and the empirical analysis to ensure consistency in geographic attribution.

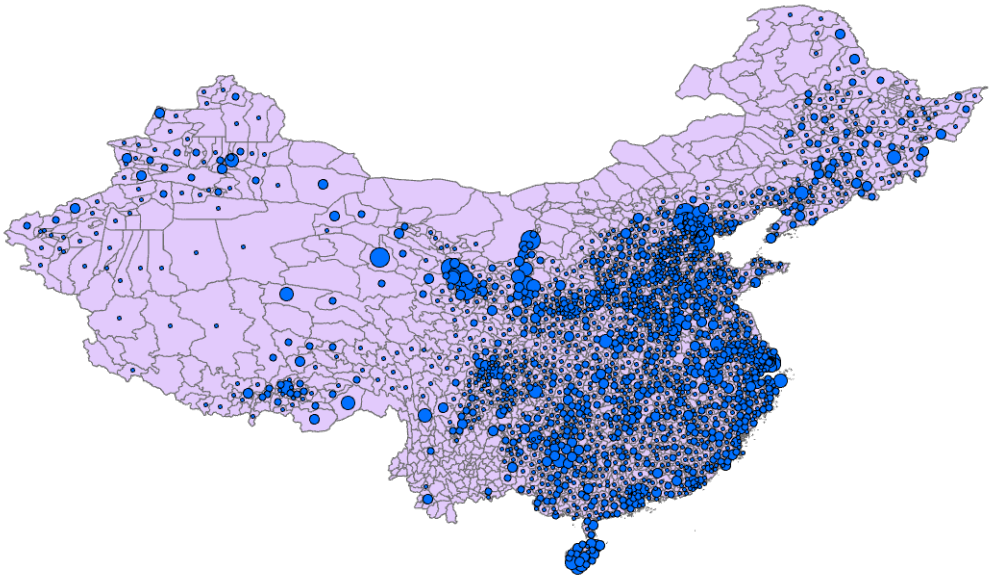


Figure A.1: Geographic Distribution of Folktales

B Sample Construction

The original dataset includes folktales from all of China's officially recognized ethnicities. However, this study focuses on regions traditionally occupied by the Han Chinese population. Although the survey data used to measure modern cultural values include some non-Han areas, the vast majority of respondents are Han Chinese, making it difficult to systematically analyze cultural values among ethnic minorities. Furthermore,

many folktale collections from minority regions lack clear geographic identifiers, preventing reliable linkage to contemporary outcomes.

The sample is constructed using the following criteria:

1. **Exclusion of Ethnic and Frontier Regions.**
   We exclude provinces and autonomous regions historically or currently recognized as ethnic or frontier areas, such that the sample approximates the Eighteen Provinces of the Qing Dynasty ("China proper"). Territories beyond the Ming Dynasty boundaries (1368–1644) are also excluded.

2. **Exclusion of Ethnic Autonomous Areas.**
   We exclude autonomous prefectures and counties where ethnic populations predominate, as well as regions with legal, cultural, or social norms that have historically diverged from Han Chinese practices.

3. **Population Composition Criterion.**
   We exclude counties where more than 20% of the population identified as non-Han in the 2000 census.



Figure A.2: Geographic Distribution of Folktales in the Sample

2 Methodology

A Synthetic-Archetype Based Measure

A.1 Overview

1. Digitizing Folktales

Scanned Chinese folktales were processed using Optical Character Recognition (OCR) software to convert them into machine-readable text. After the initial conversion, the text was reviewed to correct OCR errors. A cleaning process was applied to standardize terms and ensure consistency in linguistic structures across the dataset.

2. Preparing Synthetic Archetypes

A selection of 40 Chinese folktales was used to create a style summary capturing the linguistic and structural features of traditional narratives. These features were then used to generate synthetic stories representing specific cultural themes, such as honesty and charity.

3. Computing Semantic Similarity with BERT

The SentenceTransformer BERT model (`DMetaSoul/sbert-chinese-general-v2`) was used to calculate semantic similarity scores. Each folktale was compared against the synthetic stories to measure alignment with 10 cultural themes.

4. Aggregating Scores by Region

Each folktale was assigned to its corresponding county based on metadata. Semantic similarity scores were then aggregated at the county level, producing regional averages for each cultural theme.

5. Running Robustness Checks with GPT Models

As an alternative method, GPT models were used to classify folktales directly into cultural categories without relying on synthetic stories. The GPT-based classifications were aggregated at the county level for comparison with the BERT-based results.

6. Comparing Methods for Validation

County-level outputs from the BERT similarity analysis and GPT classification were compared. Agreement between the two methods was used to validate the robustness and reliability of the findings.

A.2 Random Text Generation and County-Level Text Similarity Control

To isolate stylistic similarities from the substantive thematic content captured by the Synthetic Archetype approach, I construct a county-level "text similarity" control variable by comparing each folktale to randomly generated text that mimics the linguistic style of genuine folktales but does not convey a coherent theme. By controlling for this measure in regressions, I ensure that the estimated relationships between cultural themes and modern outcomes are less likely to be driven by superficial textual overlap.

Random Text Generation. I employ a Markov chain text generator (adapted from the `markovify` library) trained on the entire folktale corpus, using a state size of five tokens. This parameter balances syntactic coherence and computational feasibility, generating sentences that appear stylistically consistent with Chinese folktales. Because the original code base is designed for English text, I adapt it to segment Chinese strings (e.g., via `jieba` tokenization). After training, the model outputs "null-text" that reproduces the vocabulary, phrasing, and sentence structures typical of the real tales, while avoiding any clear narrative theme or plot.

Computation of Text Similarity. Once the Markov model generates this null-text, I compute Sentence-BERT similarity scores between each real folktale and the null-text. By aggregating these similarity scores at the county level, I obtain a "style-only" measure of textual resemblance, denoted in regressions as $StyleControl_c$.

Interpretation and Usage. Substantial overlap between the synthetic archetypes and actual folktales may arise if BERT embeddings capture writing style rather than deeper thematic content. Including $StyleControl_c$ in regressions helps partial out such style-driven effects. If $Trait_c$ (i.e., the measured thematic alignment of a county's folktales with a cultural archetype) remains strongly associated with outcomes after controlling for $StyleControl_c$, then the residual similarity is more plausibly indicative of a genuine thematic match rather than mere linguistic or structural resemblance.

Empirical Validation. I gauge the usefulness of the null-text approach by comparing its similarity scores against those generated by the synthetic archetypes. Although random-text similarities can be moderately high (suggesting that BERT partly captures shared style), correlations between folktales and null-text are noticeably weaker than the alignment of folktales with their matching archetypes. This pattern indicates that, while stylistic resemblance explains some portion of the similarity, there remains meaningful scope for genuine thematic alignment in the data.

A.3 Selection Criteria for the Sample of 30 Tales

Purpose The random sample of 30 tales provides the basis for generating synthetic story styles, capturing the linguistic and structural characteristics of traditional Chinese narratives. These synthetic stories serve as stylistic benchmarks for analyzing the broader dataset without influencing substantive themes.

Criteria

1. The sample includes only Han Chinese tales, regardless of their regional origin, and excludes all tales from ethnic minority towns.

2. The 30 tales were randomly selected from the pool of Han Chinese narratives to avoid bias while adhering to the inclusion criteria.

Rationale

1. Including only Han Chinese tales minimizes linguistic and structural heterogeneity, ensuring that the synthetic stories reflect the narrative norms of Han Chinese traditions.

2. This choice aligns with the study's focus on Han Chinese narratives, which are prioritized due to the limited availability of survey data on ethnic minorities.

This selection process ensures that the synthetic story styles are representative of Han Chinese folktale traditions, improving the reliability of the methodology applied to the broader dataset.

A.4 Prompts for Generating Synthetic Archetypes

1. **Honesty**
   The story should mention a character's honest or deceitful behavior, whether such actions drive the plot or influence relationships. Specifically, the narrative should either show how a character wins trust, respect, or cooperation through steadfast honesty—highlighting its positive role in personal growth, interpersonal relationships, and social order.

2. **Trust**
   The story should mention interpersonal trust or suspicion, whether trust is being built, maintained, or lost in complex situations. Specifically, the narrative should either illustrate how characters establish or preserve trust to foster cooperation and social cohesion.

3. **Charity**
   The story should include acts of charity—such as donating, assisting, or giving alms—regardless of whether the outcome is straightforward or complex. Specifically, the narrative should demonstrate how charitable actions improve the recipient's life and promote social harmony and warmth.

4. **Effort**
   The story should mention a character's efforts or struggles, irrespective of whether these lead to success or failure. Specifically, the narrative should depict how perseverance and hard work enable the character to overcome obstacles and achieve goals, emphasizing effort as essential for personal development and social achievement.

5. **Equality**
   The story should address how equality is practiced, maintained, or contested in personal or social relationships. Specifically, the narrative should either illustrate how a character's struggle or reform promotes equality and drives social justice and cooperation.

6. **Buddhism**
   The story should depict how characters practice Buddhist teachings (such as karmic cause and effect, compassion, or meditation) or rely on sacred Buddhist powers (like a revered monk's guidance or a Buddha statue's protection) to overcome calamities or attain sudden enlightenment. It should illustrate how these doctrines shape characters' actions and beliefs, guiding them through karmic trials, dilemmas, or other challenges to find inner peace, resolve crises, or complete their spiritual cultivation.

7. **Ancestor Worship**
   The story should portray characters' reverence and remembrance of their forebears, showing how ancestors—seen as symbols of wisdom, virtue, and strength—shape descendants' beliefs, morals, and life choices. It should highlight how ancestor worship serves as a core force in the family's continuity, prompting characters to uphold familial honor, follow ancestral teachings, pass on traditional skills, or fulfill ancestral missions. In times of adversity or trial, the ancestors' protection, guidance, or legacy should provide spiritual support, strengthening the characters' resolve and intergenerational obligations.

8. **Clan**
   The story should describe the strict rules and hierarchical structures within a lineage, emphasizing members' absolute obedience to clan authority. It should highlight how characters follow clan elders' directives, protect the family's honor, or fulfill lineage responsibilities—even when confronted with personal sacrifice or moral conflict.

9. **Gift Exchange**
   The story should center on the giving and reciprocation of gifts, showing how this practice fosters important bonds, expresses respect, or fulfills obligations within the community. It should demonstrate how reciprocal gifting reinforces social ties, maintains harmony, or signals social status.

10. **Benevolent Emperor**
    The story should depict how a wise, enlightened emperor governs with benevolence, justice, and sound judgment, ensuring social stability, economic prosperity, and the well-being of the people. It should showcase how the emperor's decisions gain public support and maintain lasting order, presenting him as an ideal sovereign whose governance enhances the welfare of the entire empire.

B LLM-Based Measure

B.1 Prompt Description

You are a researcher proficient in the study of Chinese folktales. Your task is to analyze the text of each incoming Chinese folktale and determine its relevance to multiple predefined themes, as well as the folktale's stance on those themes. Output your structured analysis in JSON format using the following two-step process:

1. **Step One: Theme Relevance Judgment**
   Based on each theme's definition, decide whether the folktale text involves that theme:

   - If the folktale does not mention the theme at all, mark it as "D" (not mentioned).
   - If the folktale does mention the theme, proceed to Step Two.

2. **Step Two: Stance Judgment**
   If the folktale text is related to the theme, determine which stance it takes:

   - "A" (mentioned and supported): The folktale clearly expresses support for the corresponding viewpoint.
   - "B" (mentioned and opposed): The folktale clearly expresses opposition to the corresponding viewpoint.
   - "C" (mentioned without a clear stance): The folktale mentions the theme but does not explicitly indicate support or opposition.

Notes and Restrictions

- Base your judgments only on the folktale text plus the definitions and viewpoints provided for each theme; do not use external sources.
- Avoid far-fetched interpretations or overreading.
- The output must strictly follow a JSON format for further processing.

Folktale Text: [Place the folktale text here].

Themes and Definitions (with "A"-Level Supported Viewpoints):

Trust Definition: Trust such that one is willing to entrust responsibilities to another. If marked "A": Emphasize the importance of trust as a foundation of interpersonal relationships. Show how trust is established or broken, revealing harmony, cooperation, or consequences of betrayal.

Charity Definition: Caring for others, marked by compassion. If marked "A": Depict how benevolence improves lives and accumulates virtue and blessings, reflecting social responsibility and empathy for the vulnerable.

Honesty Definition: Aligning words and deeds with one's inner thoughts; not deceitful. If marked "A": Center the narrative on sincerity and consistency between words and deeds, highlighting trust, respect, and moral responsibility.

Effort Definition: Diligent application of strength and perseverance. If marked "A": Show how hard work and tenacity lead to overcoming challenges, achieving goals, or personal growth.

Equality Definition: Equitable treatment in social, political, economic, or legal realms. If marked "A": Emphasize consistent rights and opportunities across different backgrounds, suggesting respect for equal status.

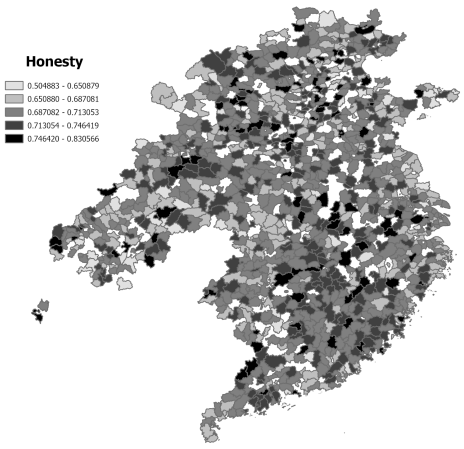
Buddhism Definition: A religion in China that originated in India. If marked "A": Highlight Buddhist teachings such as the Four Noble Truths or karmic retribution, focusing on spiritual liberation and purification.

Ancestors Definition: Forebears of a family or ethnic group. If marked "A": Depict reverence for ancestral heritage, teachings, or rituals, illustrating a sense of cultural or familial identity.

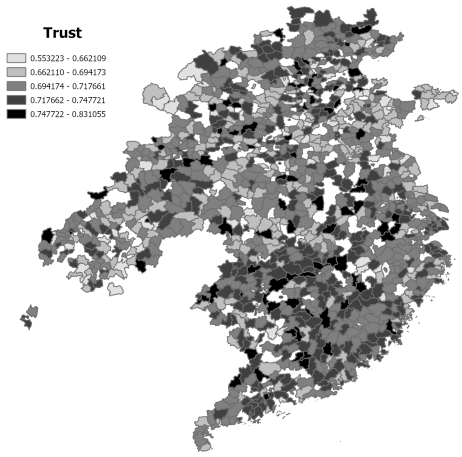
Clan Definition: A patrilineal family or its members (excluding married-out women). If marked "A": Focus on ancestral bloodline, family responsibility, and preserving tradition and honor.

Gift Exchange Definition: Bestowing a gift. If marked "A": Highlight reciprocal giving and goodwill, showcasing how courtesy fosters social harmony.

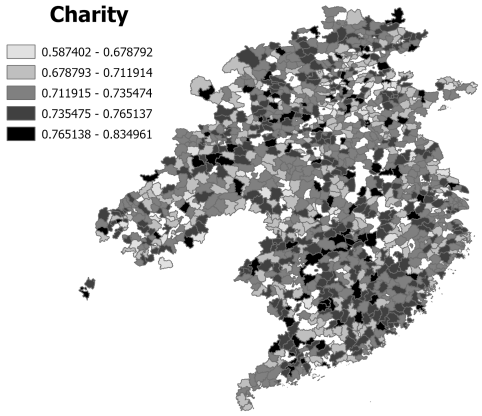
Emperor Definition: The highest ruler in a monarchical system. If marked "A": Center on virtues and governance, condemning tyranny and praising enlightened rulership, showing the ruler's profound impact on societal order.



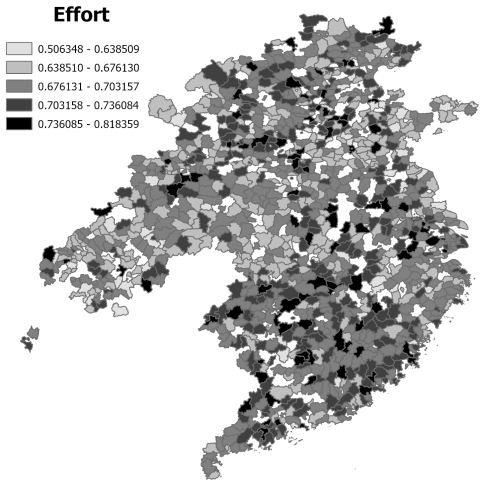
(a) Honesty



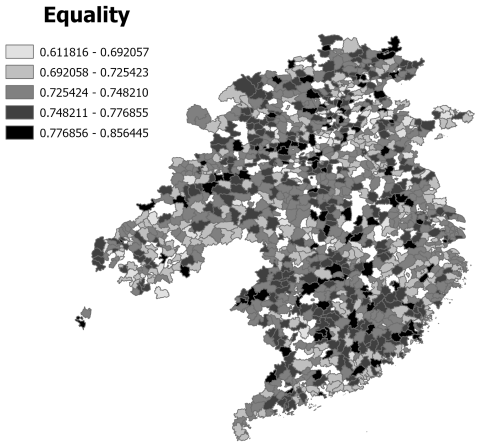
(b) Trust



(c) Charity

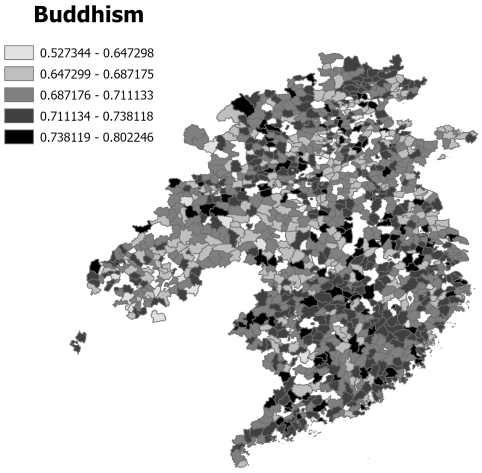


(d) Effort

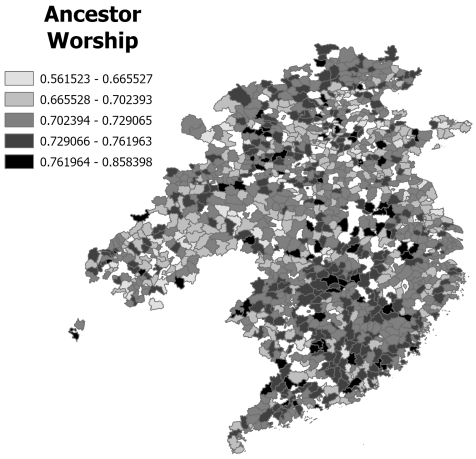


(e) Equality

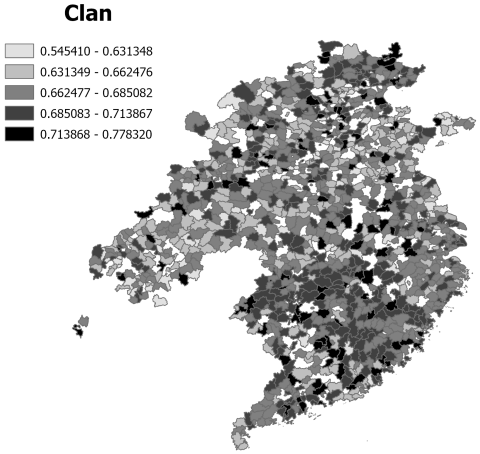
Figure A.3: Geographical Distribution of Selected Cultural Traits (1/2)



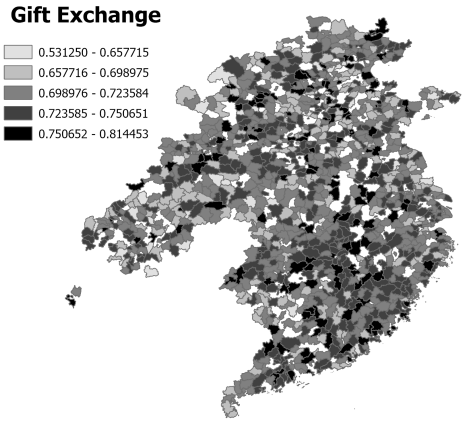
(a) Buddhism



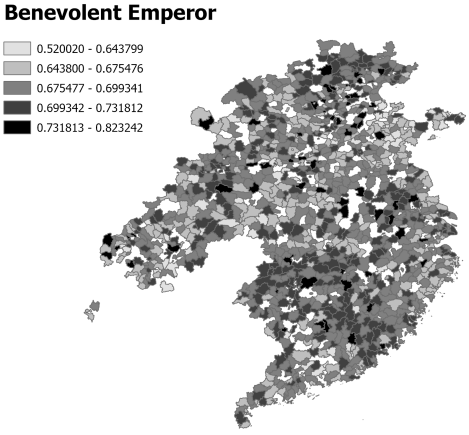
(b) Ancestor Worship



(c) Clan



(d) Gift Exchange



(e) Benevolent Emperor

Figure A.4: Geographical Distribution of Selected Cultural Traits (2/2)