University of St Andrews | Business School

Economics Discussion Papers

Cultural Capital and the Productivity of Ideas: Evidence from Historical Texts

Radoslaw Stefanski

Economics Discussion Paper No. 2602
Published: 23 July 2026
JEL classification: O41; O31; N13; Z10
Keywords: culture and growth; ideas production; growth accounting; innovation barriers; text as data

Department of Economics, Business School, University of St Andrews
Discussion Papers Series
ISSN 2978-4026
https://ideas.repec.org/s/san/econdp.html
bschoolecondor@st-andrews.ac.uk

Cultural Capital and the Productivity of Ideas

Evidence from Historical Texts

Radoslaw (Radek) Stefanski[†]

July 23, 2026

Total word count: 9,215

Abstract

Long-run growth is driven by new ideas, yet the cultural environment shaping their production is difficult to measure over time. We use large language models to read 23,000 books from the Western canon and score whether each endorses, rejects, or merely depicts six dimensions of culture. We accumulate the scores into inherited stocks and summarize them with an Innovation Wedge measuring cultural resistance to new ideas. Between 1000 and 1920 the wedge falls by 51 percent. Blinded expert readings and modern surveys validate the measure. An independent 5,000-book archive reproduces the decline. In a calibrated semi-endogenous growth model, the falling wedge raises 1920 productivity to 1.78 times its counterfactual level, explains two-thirds of the first sustained acceleration in productivity growth between 1500 and 1700, and accounts for 38.7 percent of productivity growth in 1920.

JEL classification: O41, O31, N13, Z10.

[†]University of St Andrews, Castlecliffe, The Scores, Fife, KY16 9AR, United Kingdom; University of Stavanger; and OxCarre, University of Oxford. Email address: rls7@st-andrews.ac.uk. I gratefully acknowledge the support of the Google Cloud Research Credits program for providing the computing resources necessary to conduct this research. I thank Alex Trew and Eugenio Proto for early conversations about reading Latin texts with language models, which helped motivate this project. For discussions of the literary corpus, the classics, and the approach taken here, I thank Jane Lightfoot, Georgy Kantor, and Marine Ganofsky. Finally, I am especially grateful to the scholars in literature, classics, and modern languages at the University of St Andrews whose generous participation, careful judgment, and specialist knowledge of the texts made the expert validation exercise possible.

1 Introduction

Between 1000 and 1920, the Western canon shifts away from deference and the fear of the untried toward autonomy and patience. We measure this transition with large language models (LLMs) that read more than 23,000 books in context and score whether each text endorses, rejects, or merely depicts positions along six cultural dimensions. We accumulate the scores into inherited stocks and summarize those stocks with two wedges. The Innovation Wedge is our main measure of cultural resistance to new ideas; Total Friction is a broader measure incorporating all six dimensions. Over this period, the Innovation Wedge falls by 51 percent and Total Friction by 28 percent (Figure 1).

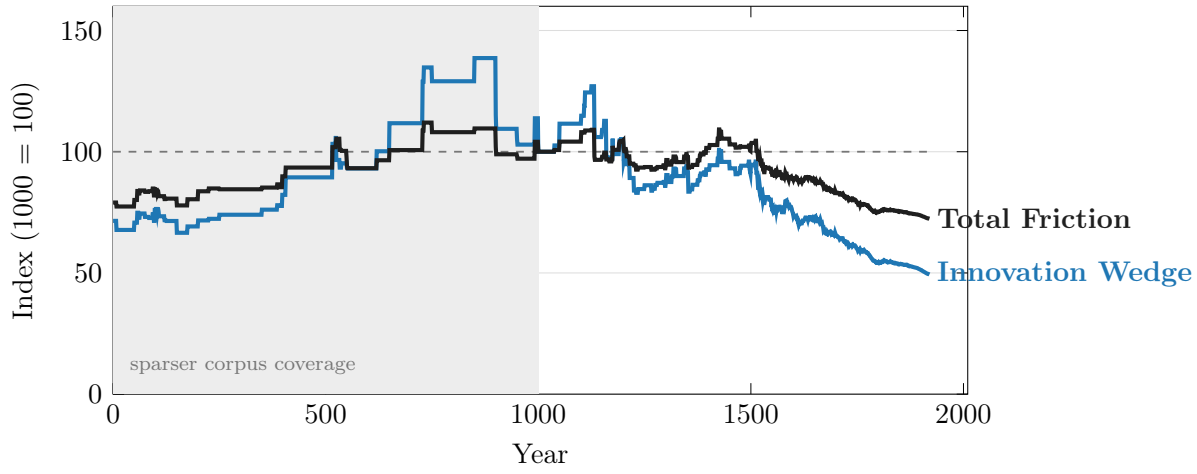


Figure 1: Cultural friction read from books.

Notes: The shaded pre-1000 segment is shown for directional context. The Innovation Wedge is $(PD + UA)/(INDIV + LTO)$; Total Friction adds the remaining two dimensions, $(PD + UA + INDUL)/(INDIV + LTO + MAS)$. The components are inherited stocks—power distance (PD), uncertainty avoidance (UA), individualism (INDIV), long-term orientation (LTO), indulgence (INDUL), and competitive striving (MAS)—read from more than 23,000 dated books. Section 3 details the construction; Figure 3 reports the six stocks in levels.

This transition matters in idea-based growth models because a slow-moving cultural environment can alter research productivity—the efficiency with which research effort produces useful knowledge (Romer, 1990; Jones, 1995). A large literature links culture to growth and development through historical accounts (Mokyr, 2016; McCloskey, 2016; de la Croix et al., 2018; Squicciarini, 2020), formal theory (Bénabou et al., 2022), and cross-country evidence (Tabellini, 2010; Gorodnichenko and Roland, 2011; Alesina and Giuliano, 2015; Nunn, 2012). A related literature traces Western individualism and impersonal cooperation to the erosion of intensive kinship under the Western Church (Schulz et al., 2019; Henrich, 2020).[1] Existing

[1]Our series measures some of the same margins but does not test that mechanism.

work, however, offers no cultural series that is both long enough to span the emergence of modern growth and sufficiently structured to enter a quantitative growth model.

We build such a series from books drawn from Project Gutenberg (2026), a public-domain library. The books in our corpus do not provide a direct measure of population beliefs. They preserve cultural stances expressed in texts written, read, and taught by literate elites—toward hierarchy, autonomy, uncertainty, achievement, indulgence, and patience. Recovering those stances, however, is a matter of reading rather than counting.

The closest antecedent is Michalopoulos and Xue (2021), who measure culture in the cross-section from a folklorist's catalogue of oral traditions mapped to ethnic groups. There the texts reach the measure only through the cataloguer's summaries; here the raw texts are read in context, dated, and their scores accumulated into a 1000–1920 time series. Recent work traces the rise of a progress vocabulary in the digitized English-language print record (Almelhem et al., 2026). Using a broader Western corpus, we instead ask whether each text endorses a value, rejects it, or merely depicts it. A progress-vocabulary proxy built on our own corpus shows that the two approaches measure related but distinct objects (Appendix A.6). They differ in two respects. First, the vocabulary measure counts what texts discuss, while our scores recover the stance texts take; the worked pair in Section 3 makes that distinction concrete. Second, the two measures move at different times. The progress vocabulary moves first, behaving like a flow of new ideas; the inherited stock turns later and more slowly, like the environment ideas enter.

We validate the measure in four ways. First, expert readers blinded to our scores agree with the direction of those scores in 93.0 percent of directional anchor comparisons among books they know. Second, although our historical series end in 1920,[2] the component stocks align with the cross-country orderings reported by modern cultural measures, especially on the dimensions used in the growth model of Section 6. Third, the component stocks usually move in the direction historians would expect around major episodes of European history. Finally, an independent archive of 5,000 historical books, drawn from four collections outside the main corpus, reproduces the decline when we apply the original scoring pipeline without modification.[3]

We use our measured Innovation Wedge to quantify culture's contribution to modern growth in a semi-endogenous model of idea production. A lower wedge raises research productivity, so a given amount of research effort produces more ideas, which accumulate into a higher level of total factor productivity (TFP). We estimate the model's key wedge elasticity and

[2]The 1920 endpoint is imposed by the archive. Later works largely remain under copyright, so a public-domain collection covers a rapidly shrinking share of what was written thereafter.

[3]The four collections are EEBO-TCP, Evans, ECCO, and the Internet Archive. Appendix B.4 provides their full names and construction details.

calibrate the remaining parameters. In the model, holding the wedge at its year-1000 value leaves productivity in 1920 at 56.3 percent of its realized value. In the long run, culture affects the productivity level, whereas population growth pins down the growth rate. Because the transition unfolds over centuries, however, the falling wedge also raises the growth rate along the transition path. By 1920, it accounts for 38.7 percent of productivity growth in that year. Between 1500 and 1700, the period to which Mokyr (2016) traces the cultural origins of modern economic growth, the falling wedge explains 64.8 percent of the first sustained acceleration in productivity growth. Population scale becomes the larger contributor to productivity growth only in the nineteenth century.

Two additional exercises clarify what the measure captures and how it varies across space. Across the six margins, passages take stances about as often in the nineteenth century as in the eleventh. The change, however, is in what texts endorse. Among passages that take a stance, the mean stance score moves by 0.511, about half the distance from the scale's midpoint to full endorsement. We then map the stocks across space. Within any year, the surface maps show where friction sat lowest; over time, the low-wedge frontier drifts northwest, and on a three-degree European grid, nearby power distance and uncertainty avoidance predict less urbanization over the following century while nearby individualism predicts more. We interpret these regional patterns as predictive rather than causal.

Our empirical object is the durable literary culture produced and preserved by Western elites and institutions. Literacy remained limited for most of the period, and the literate minority disproportionately wrote, taught, governed, and innovated. The cultural stances of that minority persist through the channels the literature documents for cultural and human-capital transmission—schooling (Squicciarini and Voigtländer, 2015; Becker and Woessmann, 2009), imitation (Bisin and Verdier, 2001), canon formation (Mokyr, 2016), institutions (Tabellini, 2010), and selection (Clark, 2007; Galor and Özak, 2016; Stefanski and Trew, 2024a). If norms relevant to idea production travel through these channels, they should appear in that minority's written record. The corpus's Western scope matches the episode we seek to explain, the takeoff of sustained growth in the West. The corpus does not directly represent non-Western literatures that influenced Western culture. Those influences enter the measure only insofar as they changed what the West wrote and kept.

The sharper sampling concern is selection into the digital archive. Gutenberg's coverage is selected twice. History selected which works were preserved, copied, and taught, and that selection is part of the object we study. Present-day digitizers select which surviving works are scanned, and those choices, shaped by language, popularity, and availability, are the

concern. The independent archive provides a check on this second selection. It draws on four independently assembled collections, and rebuilding the series there reproduces the decline.

The rest of the paper proceeds as follows. Section 2 presents the growth interpretation and explains where culture enters. Section 3 describes the corpus, the scoring protocol, and the construction of the stocks and wedges. Section 4 validates the measure against blinded expert readers, independent modern cultural benchmarks, and the second, non-overlapping archive. Section 5 documents the main time-series facts. Section 6 translates the measured transition into a calibrated model of idea production. Section 7 maps the measure across space and tests whether component stocks predict later urbanization on a coarser European grid. Section 8 concludes.

2 A Growth Interpretation

We use the simplest framework in which culture can affect the long-run path of ideas, and through ideas, TFP. We start with a standard semi-endogenous idea-production function:

$$\dot{A}_t = \delta_t H_{A,t}^{\lambda} A_t^{\phi}, \qquad 0 < \lambda \leq 1, \quad 0 < \phi < 1. \tag{1}$$

Here A_t is the stock of ideas, $H_{A,t}$ is the effort devoted to producing new ones, and δ_t is the productivity of that effort. New ideas raise aggregate productivity, so a higher path for A_t is a higher path for TFP.

We let culture in through one door, the productivity of research:

$$\log \delta_t = \log \bar{\delta} - \psi \, \tau_t^I, \qquad \psi > 0, \tag{2}$$

where τ_t^I is the Innovation Wedge, defined in Section 3 as a ratio of four of the six scored dimensions—hierarchy and uncertainty avoidance over individualism and long-term orientation. When this wedge is high—when deference and a fear of the untried weigh heavily against autonomy and a long horizon—a given amount of research effort yields fewer ideas.

In our model, culture does not shift preferences, rebuild institutions, and reorganize production all at once. The model also leaves aside task complementarities in O-ring economies and the trust, rules, and enforcement channels in models of institutions (Guiso et al., 2006; Tabellini, 2010; Algan and Cahuc, 2010). Those margins are real, and separating them from the idea channel is an open problem in its own right. We exclude these historically relevant margins

by construction to keep the accounting to a single, interpretable channel. We assume culture moves one term, the productivity of research, so a measured cultural series can be read directly into δ_t and the question becomes how much that channel moves TFP.

We feed the wedge only into δ_t because related work ties cultural attitudes—individualism, tolerance of uncertainty, openness to new knowledge—to innovation and scientific progress (Gorodnichenko and Roland, 2011; Bénabou et al., 2022). Above all the choice is tied to Mokyr (2016), for whom the decisive change was in the culture of Europe's educated elite, a market for ideas in which intellectual authority could be contested and useful knowledge accumulated. Our series can be read as a quantitative shadow of that account. The margins his mechanism needs—autonomy, patience, and tolerance of the untried—are those that move first and furthest in the measured stock.

The assumption $\phi < 1$ makes the model semi-endogenous in the sense of Jones (1995). The long-run growth rate of ideas is set by the growth rate of research effort—ultimately population—and not by δ_t, so a permanent fall in the wedge lifts the *path* of ideas without changing the asymptotic growth rate. This requires the decline in the wedge to die out in the long run, so that δ_t settles at a constant; along the transition, a falling wedge raises measured growth.

The long transition makes this level effect quantitatively important. A semi-endogenous model with our calibrated parameters moves between paths slowly,[4] and the decline in the wedge we document runs from 1000 to 1920. On a horizon that long, a level effect on the stock of ideas is large enough to matter for growth.

The mechanism also relates to recent work on long-run traits such as patience. Stefanski and Trew (2024a) argue that a secular rise in effective societal patience helps account for falling interest rates and the accumulation behind modern growth. Our object is the cultural environment in which traits such as patience are taught and praised, distinct from the household discount factor. The two channels are complements. The same inherited drift toward patience that lowers the price of time in their account raises the productivity of research in ours, working through accumulation there and through ideas here.

[4]The local half-life of convergence is $\ln 2/(\lambda n)$, which is roughly two and a half centuries at the calibrated values $\lambda = 1.00$ and $n = 0.29$ percent (Section 6).

3 Constructing the Measure

The model's input is a measured cultural series. We build it from historical books in three steps. We first identify the source work that each digital text represents and assign that work a date and origin. We then read the text for the cultural stances it endorses, rather than the subjects it merely depicts. Finally, we accumulate book-level scores into persistent stocks and combine the model-relevant stocks into an Innovation Wedge.

3.1 Corpus Construction and Scope

Our starting point is Project Gutenberg, a public-domain library that combines machine-readable texts with bibliographic records. We use its English-language files, which include works composed in English and English translations of works composed in other languages. A translation enters at the date and origin of the underlying source work, not at the date or place of a later English edition. The resulting object is therefore the Western literary tradition as preserved and transmitted in English, not a history of editions or circulation.

Catalog records are standardized and deduplicated at the work level. We then assign each work a composition date and origin using LLM-assisted search anchored in documentary sources. Before scoring, every text passes a screening review that removes broken files, duplicate editions, generic collections, and works that are not authored literary or intellectual texts. Automated diagnostics direct manual attention to uncertain cases. Appendix A.1 summarizes the path from raw catalog to the scored sample of 23,064 books; the aggregate audit table is in Online Appendix Section OA.B.6 and the work-level ledger is in the replication package.

The measurement object is durable literary culture, centered on the books that Western elites and institutions wrote, preserved, copied, and taught. It is not a survey of contemporaneous mass beliefs and it does not represent world literary production. Survival and canon formation select texts unevenly (Buringh and van Zanden, 2009). If the margin that matters for modern growth is upper-tail human capital, durable elite texts sit close to it (Squicciarini and Voigtländer, 2015). Selection into Gutenberg is a separate concern; Section 4.4 addresses it with an archive built from collections that Gutenberg does not cover.

Figure 2 reports the realized spatial support. The corpus has a Mediterranean-European core, which later expands across the North Atlantic, and contains scattered Western works written elsewhere. Those extra-core observations are largely works by Western colonists, missionaries,

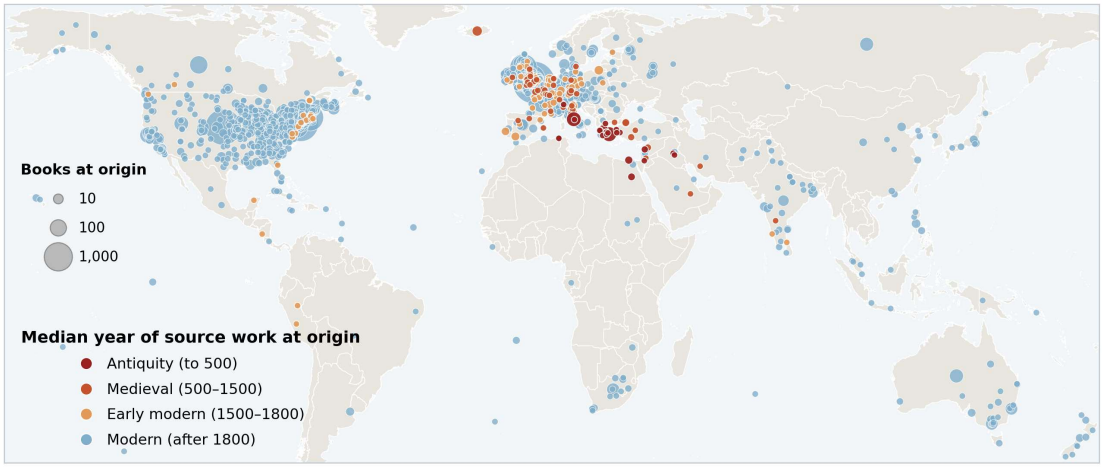


Figure 2: The sample is centered on the Western literary world, with a Mediterranean-European core and later North American expansion.

Notes: Each point is a unique mapped literary origin. Point size records the number of matched books and color records the median source-work year. The production run contains 23,064 matched geocoded book rows collapsed to 1,661 mapped origin-coordinate points. The figure describes the support of the corpus used in the paper; it is not a census of historical literary production.

and travelers abroad rather than evidence of locally representative literary traditions. The map should therefore be read as a support diagnostic for a Western literary stock.

3.2 Reading Cultural Stance

Dimensions. We borrow five cultural-margin labels that economists repeatedly study from Hofstede (2001). Each contrasts endorsement of one pole with endorsement of its opposite. Power Distance (PD) sets hierarchy, authority, and order against equality, decentralization, and meritocracy. Uncertainty Avoidance (UA) sets rules, clarity, safety, and dogma against ambiguity, experimentation, tolerance of disorder, and innovation. Individualism (INDIV) sets independence, personal rights, and individual identity against collective identity, loyalty, collectivism, and harmony. Long-Term Orientation (LTO) sets saving, pragmatic adaptation, and worldly planning for practical material returns against static tradition, honor or reputation, absolute truth, and social obligations. Masculinity (MAS) sets achievement, strength, competition, and material success against care, modesty, cooperation, and quality of life. Indulgence (INDUL), added later, sets fun, gratification, and freedom against restraint, asceticism, and strict social norms (Hofstede et al., 2010).[5] The labels name the margins we

[5]Appendix A.2 shows how the labels operate in scoring, including the narrowed Long-Term Orientation rule—religious or moral future orientation alone does not count as high LTO unless tied to worldly planning,

measure in text; the construction is not calibrated to Hofstede's survey scores and does not assume that historical authors used modern survey concepts.

Endorsement rather than vocabulary. The key feature of our scoring is that it separates what a text depicts from what it endorses. Machiavelli's *The Prince* (1513) treats hierarchy as an instrument of political action, whereas Venture Smith's *A Narrative of the Life and Adventures of Venture, a Native of Africa* (1798), the autobiography of a man enslaved in New England who purchased his own freedom, depicts extreme hierarchy while condemning domination and asserting self-ownership. A Power Distance dictionary ranks the books in the opposite order—66.7 for *The Prince* and 100.0 for Smith's *Narrative*—because the latter repeatedly uses the language of command and enslavement. Contextual reading instead assigns scores of 100.0 and 0.0, respectively. Appendix A.2 reports the complete trace for *The Prince* and the parallel contrast for Venture Smith; Online Appendix Section OA.C prints abbreviated records for both.[6]

Scoring. Scoring proceeds in four stages—a context pass, literary analysis, behavioral extraction, and synthesis—each inheriting the output of the one before it. The context pass classifies the work and its framing; literary analysis reads every selected passage through that context, deciding whether words on the page are endorsed, satirized, or merely reported; behavioral extraction, given that reading, records what each passage shows about the six dimensions; and synthesis combines framing and evidence into a directional score. Each passage is judged with the whole work's framing in view rather than in isolation.

The production system assigns Gemini 2.5 Pro to the context pass and Gemini 2.5 Flash to literary analysis, behavioral extraction, and synthesis.[7] Each book is divided into fixed-length passages of 30,000 characters with 500-character overlap; the context pass reads each book's opening and closing 10,000 characters. An initial tranche of 2,753 books was scored on every passage; the rest are scored on three to five evenly spaced passages per book. Online Appendix Section OA.C gives the maintained prompts and stage sequence, Appendix A.3

material investment, or practical future payoff. Online Appendix Section OA.C prints the production instructions and records the task-to-model map.

[6]The pair is not special; Appendix A.4 runs the same dictionary comparison for every book in the corpus, where agreement is modest book by book (0.16) for the same reason it is modest here.

[7]Gemini 2.5 Pro and Gemini 2.5 Flash are commercial LLMs from Google; Pro is the more capable and costlier model, Flash the faster and cheaper one. The context pass runs once per book and gets the stronger model; the passage-level stages run hundreds of thousands of times and get the cheaper one. Online Appendix Section OA.C.4 reports the cost accounting.

reports the sampling checks, and Appendix A.4 reports the prompt, model, and full-corpus method checks. At the passage level, dimension d receives

$$x_d \in \{-1, 0, +1\},$$

where $+1$ denotes endorsement of the high-trait pole, -1 denotes endorsement of the low-trait pole or rejection of the high pole, and 0 denotes no net directional evidence. The zero category includes both balance and silence, and zeros enter the passage average. A single directional passage in an otherwise silent book is therefore diluted toward neutrality rather than defining the book. Treating silence as missing leaves the accounting essentially unchanged (Appendix A.5).

Book-level scores are averages of the scored passages, mapped linearly to the 0–100 scale,

$$S_d = 50(\bar{x}_d + 1). \tag{3}$$

The scale has a direct cardinal reading. Because passages are scored -1, 0, or $+1$, a book's score is an affine transform of the net share of its passages endorsing the high pole, so a one-point difference corresponds to two percentage points of net endorsement.[8] What the scale does not claim is intensity—a vehement endorsement and a mild one both count $+1$—so the external exercises in Section 4 test rankings, and Appendix D.4 shows that rank-transforming the book scores leaves the accounting essentially unchanged.

The scores are noisy book by book but stable in the stocks built from them.[9] Sampling matters least. For the books read in full, scores built from only three to five passages match the full scores at mean rank correlation 0.911 and the Innovation-Wedge path at 0.983. The remaining noise sources wash out in the stocks. Independent end-to-end reruns reproduce the wedge path at correlations no lower than 0.985 (Online Appendix Table OA.5). Rescoring with alternative prompts and the independently trained DeepSeek v4 family[10] leaves median book-level rank agreement between 0.629 and 0.786. Prior-free lexical and embedding scorers agree with the staged scores at only 0.16 book by book yet still track the wedge path at 0.79 and 0.88. That modest book-level agreement is expected—dictionaries count vocabulary while the staged reading judges stance (Appendix A.4).

[8] A book at 80 on PD endorses hierarchy on net in 60 percent of its scored passages, a book at 20 rejects it on net in 60 percent, and a book at 50 is balanced or silent.

[9] Noise here means variation across runs, prompts, and model families in the reading of a passage and, for sampled books, the particular passages the sampling rule selects.

[10] DeepSeek v4 is an LLM from a different developer, so agreement with the Gemini-based scores cannot come from shared training.

Finally, the decline is not driven by books falling silent on the six dimensions. Passages take a clear position about as often in the nineteenth century as in the eleventh; what changes is the position they take (Section 5; Appendix A.5).

3.3 From Book Scores to Cultural Capital Stocks

A work can shape teaching, imitation, and institutions long after it is written, and models of cultural transmission imply gradual rather than annual adjustment (Bisin and Verdier, 2001; Fernández, 2013). We therefore accumulate the dated book scores into a slowly depreciating stock of inherited culture, using the perpetual-inventory construction familiar from physical capital.

Let N_t be the number of scored books dated to year t and let $\bar{S}_{d,t}$ be their average score on dimension d. The baseline allows years with more books to carry more weight while dampening the steep rise in surviving publication volume.

$$v_t = \log(1 + N_t), \qquad m_{d,t} = \begin{cases} \bar{S}_{d,t} \log(1 + N_t), & N_t > 0, \\ 0, & N_t = 0. \end{cases} \tag{4}$$

Here v_t is the observed cultural-volume inflow and $m_{d,t}$ is its score-weighted counterpart. The inherited volume and momentum stocks obey

$$V_t = (1 - \delta_K)V_{t-1} + v_t, \tag{5}$$

$$M_{d,t} = (1 - \delta_K)M_{d,t-1} + m_{d,t}, \tag{6}$$

and the cultural-capital stock is

$$K_{d,t} = \frac{M_{d,t}}{V_t}. \tag{7}$$

Thus $K_{d,t}$ is a surviving weighted average of inherited cultural content, not a measure of the absolute quantity of texts. Volume and momentum depreciate at the same rate, so an empty year leaves the composition $K_{d,t}$ unchanged. We set the stock-depreciation parameter δ_K to 0.005 in the preferred specification, a 99.5 percent annual retention rate with a half-life of roughly 138 years. The value is chosen to represent persistence across generations and institutions (Voigtländer and Voth, 2012; Nunn and Wantchekon, 2011); it is not a structural estimate. Appendix C varies it over a wide range.

The baseline log damping prevents dense modern publication years from mechanically overwhelming the medieval stock while retaining the idea that a year with more textual production should matter more. Appendix C records the construction details, support, timing, a worked England example, and the accumulation and composition checks, and Appendix A.5 the separate silence treatment; the long-run pattern survives all of them.[11]

3.4 The Innovation Wedge

The main economic object combines the four margins most directly tied to the production of ideas. The Innovation Wedge is

$$\tau_t^{\text{Innovation}} = \frac{K_{PD,t} + K_{UA,t}}{K_{INDIV,t} + K_{LTO,t}}, \tag{8}$$

written τ_t^I in the model sections. The numerator combines hierarchy (PD) and the fear of the untried (UA); the denominator combines autonomy (INDIV) and patience (LTO). A higher wedge therefore represents greater cultural resistance to dissent, experimentation, and patient accumulation. The component orientations draw on the culture-and-growth literature (Gorodnichenko and Roland, 2011; Tabellini, 2010) and Hofstede's dimension definitions (Hofstede, 2001); the ratio itself is our construction.

For description, we also report the broader Total Friction measure,

$$\tau_t^{\text{Total}} = \frac{K_{PD,t} + K_{UA,t} + K_{INDUL,t}}{K_{INDIV,t} + K_{LTO,t} + K_{MAS,t}}. \tag{9}$$

MAS and INDUL are less clearly signed for discovery, so they do not enter the model's Innovation Wedge. In particular, Indulgence can also be read as freedom from rigid social control rather than as weak discipline. The quantitative analysis therefore centers on τ_t^I, while Total Friction remains a descriptive companion. Nothing depends on the exact formula—taking differences instead of ratios, dropping single margins, signed principal components, and alternative stock constructions all preserve the long-run decline reported in Section 5 (Appendix C).

[11] Online Appendix Section OA.A states the exact conventions behind every headline computation—dates, samples, country names, numerical routines—so each number can be reproduced without guesswork.

4 Validating the Measure

We check the measure at four levels, from individual books to the entire nine-century path. Expert readers, blinded to the scores, compare pairs of books they know. Modern cross-country surveys supply the orderings the inherited stocks should reproduce. Dated historical episodes provide directional checks on the path itself. A second archive of roughly 5,000 books, scored with the pipeline unchanged, rebuilds the path. None of these checks alone validates the whole path or its cardinal scale but together they brace the measure at each level at which the paper uses it.

4.1 Expert Recognition of Book-Level Scores

The first check asks whether expert readers see the same contrasts the scorer finds. It is a recognition test of selected high-contrast book-level scores, not an audit of every book. Faculty members in literature, classics, and modern languages at the University of St Andrews—14 of 30 invited—completed blinded comparison packets containing anchor pairs spanning the six dimensions and additional rotating pairs.

Expert readers saw the titles, authors, dates, a neutral orientation aid, and a plain-English description of the margin, but not the model score, predicted direction, dimension label, or internal scoring output. They could abstain when they did not know both works.

Among anchor comparisons that experts knew and answered directionally, agreement with the score-implied ranking is 93.0 percent, and every anchor pair has majority agreement. Counting ties and cannot-judge responses in the denominator gives a conservative confirmation rate of 65.5 percent. Familiar-both coverage is 72.6 percent, the cannot-judge rate is 22.6 percent, and the rotating supplements agree in 95.8 percent of familiar-both directional responses. Appendix B.1 reports packet design, response counts, coding, and dimension-level results.

4.2 Modern Cross-Country Benchmarks

The second check asks whether a cultural stock read from pre-1920 texts recovers persistent cross-country orderings measured by modern surveys rather than noise. Nothing in our construction is calibrated to Hofstede's survey values, so agreement is a finding rather than a construction. Nor should the stocks reproduce any single battery exactly; the test is whether they recover country orderings on any cultural survey. We use a fixed dimension-and-sign map

for every external benchmark, so a positive Spearman correlation always denotes agreement in country rankings. The benchmarks include official Hofstede scores (Hofstede, 2001; Hofstede et al., 2010), the major value surveys, the Quality of Government battery (Teorell et al., 2026), GLOBE practices (House et al., 2004), and Schwartz values (Schwartz, 1992, 2008); Appendix B.2.1 reports the complete map.

The primary sample is Europe and North America, which contains 97.5 percent of complete scored book rows (22,493 of 23,064). Texts written elsewhere are often Western works produced abroad, whereas the external surveys measure local populations. We therefore use the non-Western comparison as an out-of-scope diagnostic. Agreement should weaken rather than extend mechanically beyond the corpus's support.

Table 1 reports the results. Across 138 comparisons in 11 benchmark families, the inherited text stock has mean signed rank correlation 0.061, the predicted sign in 65.2 percent of tests, and 9 of 11 families positive (family-level $p = 0.033$). Restricting attention to the four Innovation-Wedge components gives 76 tests across 11 families, mean $\rho = 0.095$, predicted signs in 68.4 percent, and 10 of 11 families positive ($p = 0.006$). The ceiling comparison below evaluates this same four-component set.

The out-of-scope non-Western comparison has mean correlation -0.076. Applying the same signed map to log income gives -0.092, and residualizing both the text stock and the benchmarks on log 1920 income leaves mean partial rank correlation 0.091 (9 of 11 families positive, $p = 0.033$). Agreement weakens where corpus support weakens and survives income residualization. One remaining concern is that the model may recognize a book's origin and import modern beliefs about that country into its scores. Dictionary and embedding scorers carry no such priors and reproduce the aggregate paths, so the paths do not rest on them (Appendix A.4).

Comparing the stock directly with the six official Hofstede dimensions gives weak correlations, with mean $\rho = 0.014$, though with only 6 tests the comparison is highly underpowered (Appendix B.2.4). A tougher and better-powered comparison puts our culture stock measure head to head with official Hofstede scores at predicting the *other* modern surveys. On direction, the stock nearly ties, matching the predicted sign in 74.4 percent of 86 tests against 75.6 percent for Hofstede. On strength, it trails; its correlations run 70.7 percent as large as Hofstede's at the mean and 35.0 percent at the median, improving to 86.9 percent on the four components the model uses (0.126 versus 0.145).

The four Innovation-Wedge components do not perform equally. Individualism and Long-Term Orientation have mean correlations 0.095 and 0.209, and Uncertainty Avoidance has 0.107.

Table 1: External validation of the inherited cultural stock

Object and sample	Tests	Mean N	Mean ρ	Median ρ	Test sign	Fam. +	Sign p
Primary all-available modern benchmarks, Europe and North America	138	19.0	0.061	0.058	65.2%	9/11	0.033
Comparison: log 1920 income, same signed map	138	16.6	-0.092	-0.162	42.0%	–	0.975
Out-of-scope non-Western check, all available modern benchmarks	138	23.4	-0.076	-0.082	34.1%	2/11	0.994
Innovation components, Europe and North America	76	19.0	0.095	0.096	68.4%	10/11	0.006
Innovation Wedge, Europe and North America	5	20.0	0.111	0.075	80.0%	–	0.188
Total Friction, Europe and North America	39	17.8	-0.021	0.008	51.3%	4/9	0.746
Text stock vs. non-Hofstede modern benchmarks	86	18.0	0.135	0.122	74.4%	9/10	0.011
Official Hofstede vs. same non-Hofstede benchmarks	86	18.0	0.191	0.349	75.6%	8/10	0.055
Text stock, innovation components vs. non-Hofstede benchmarks	72	17.5	0.126	0.104	73.6%	9/10	0.011
Official Hofstede, innovation components vs. same benchmarks	72	17.5	0.145	0.308	70.8%	6/10	0.377
Direct comparison to official Hofstede dimensions, Europe and North America	6	23.3	0.014	0.003	50.0%	–	0.656
1850–1920 bridge to official Hofstede dimensions, Europe and North America	6	23.3	0.092	0.115	83.3%	–	0.109

All rows use the same fixed sign map. Mean ρ and Median ρ summarize signed rank correlations, Test sign is the share of comparisons with the predicted sign, Fam. + is the number of benchmark families with positive mean correlation, and Sign p is the family-level sign-test probability. The 1850–1920 bridge extrapolates each country's text-stock trend to 1970 and compares the six implied dimensions with the official Hofstede dimensions. The out-of-scope non-Western row is a check on the signed map outside the primary geographic sample, not a randomization placebo.

The clearest external mirror for Power Distance is reversed Schwartz egalitarianism, which correlates 0.304 with the text stock in Europe and North America and 0.558 in the Western core. Its pooled same-name comparisons remain weaker, averaging -0.059 with 31.6 percent correctly signed. A reduced wedge built from UA and LTO alone, the two best-validated components, follows a similar decline to our baseline wedge, so the headline pattern does not lean on the margins where external validation is weakest.

4.3 Dated Historical Episodes

The above cross-sectional exercises support persistent country rankings, not the historical path or its cardinal scale. A separate check uses dated historical episodes. We assemble a fixed set—the Reformation, the Enlightenment, the French Revolution, and others—each with dates, affected countries, and the component movement standard histories would lead one to expect. In 11 of 15 primary comparisons, the component moves in the expected direction, although control regions often move the same way, consistent with broad diffusion rather than event-localized breaks. The one primary miss is the Napoleonic episode, where legal reforms imposed territory by territory are hard to map onto the panel's country units (primary panel and sign summary in Appendix B.3; full registry and results in Online Appendix Sections OA.E.2.1–OA.E.2.2). The direct test of the path is the independent archive below.

4.4 An Independent Archive

We assemble this second archive from collections outside Gutenberg and apply the original prompts, passage-selection rule, and stock recursion without re-estimating or retuning any part of the pipeline. The archive stages 5,037 books from EEBO-TCP, Evans, ECCO, and the Internet Archive after exact and close title–author matches with the main corpus are excluded, and retains 4,996 books after screening.[12] Unlike the source-work dating used in the main corpus, these books enter at publication or edition dates, the dates that the external collections supply consistently. The exercise therefore changes the archive and stresses the dating convention.

We evaluate whether the Innovation Wedge declines from the earliest well-covered period to the late nineteenth century, whether its decade path co-moves with the Gutenberg path, and whether the size of its decline lies within the range produced by resampling Gutenberg itself. The cleanest size comparison starts in 1550, after the shared cold start. From there the archive wedge changes by -0.743, against -0.356 in a matched Gutenberg sample. The steeper archive decline is what publication dating should produce because books enter at edition dates rather than their often earlier composition dates.

Over the longer 1500–1890 span, the archive wedge declines by 35.8 percent (from 2.241 to 1.438), compared with 60.7 percent for matched Gutenberg. Across 40 decades, the archive and Gutenberg paths have Spearman rank correlation 0.902 in levels and Spearman rank

[12]Works matching Gutenberg by exact or close title–author comparison are excluded when the archive is assembled. The analysis keeps complete six-dimension records dated 1500–1900 and drops residual newspapers, serials, and periodicals. Online Appendix Section OA.E.3.2 reports the screen, count by count.

correlation 0.349 in decade first differences. The decline lies at the 15th percentile of the Gutenberg resampling distribution. Seeding the archive from Gutenberg's inherited stock produces a decline of 0.555, while the unaccumulated decade flows fall by 1.076. The result does not depend on the empty-stock initialization or on accumulation alone.

The four collections enter in three successive chronological blocks, with EEBO-TCP and Evans both contributing to the earliest, so level changes at collection joins cannot be separated fully from time changes. Within the early collections, however, the wedge declines through 1800; the Internet Archive segment is flatter in the nineteenth century. The joins leave the exact full-period magnitude less secure than the direction and broad shape of the historical transition. The exercise tests Gutenberg's selection among surviving works, not survival itself, which is part of the measured object; and because the collections share no common subject classification, composition differs across archives. Appendix B.4 reports the decade path and criteria verdicts; collection-specific results and resampling tests are in Online Appendix Sections OA.E.3.1–OA.E.3.4.

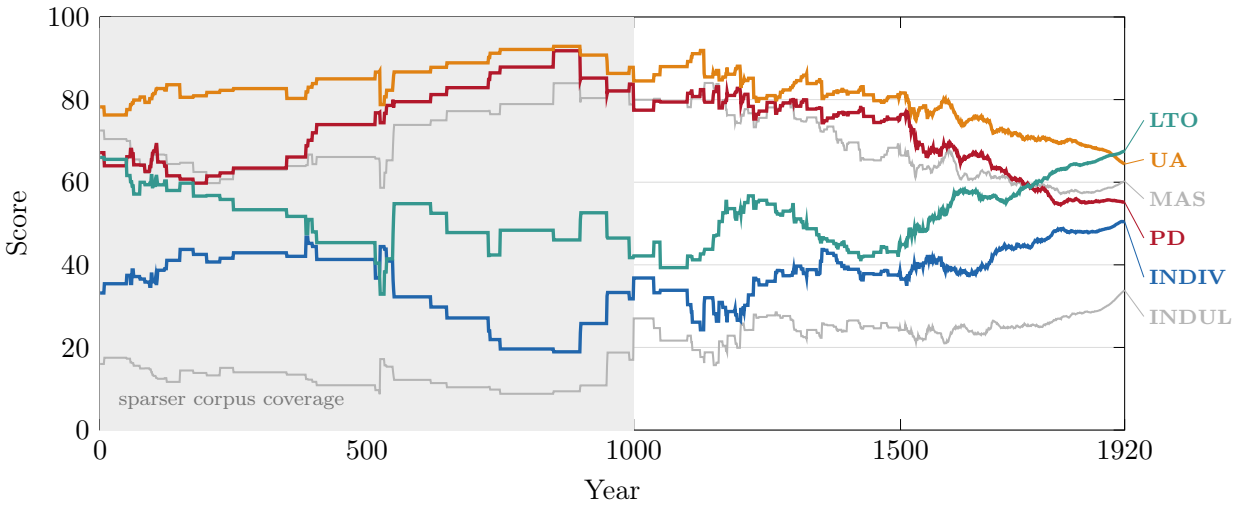
5 The Historical Decline in the Innovation Wedge

We now characterize the measured transition itself—how the six inherited stocks move between 1000 and 1920 and how those movements combine into the Innovation Wedge and the broader Total Friction measure. Three facts organize the evidence. The wedge falls by roughly half. Total Friction falls by less. And substantial movement predates 1500, the early-modern window Mokyr (2016) emphasizes.

On the four margins entering the Innovation Wedge, the inherited literary stock in 1920 is less hierarchical and rule-bound, and more individualist and future-oriented, than in 1000 (Figure 3). PD falls from 77.4 to 55.2 and UA from 84.5 to 64.4, while INDIV rises from 36.8 to 50.5 and LTO from 42.2 to 67.5. The two margins outside the Innovation Wedge also change. Competitive striving (MAS) falls from 80 to 60.3, while indulgence (INDUL) rises from 27 to 33.9.

These movements reinforce one another in the Innovation Wedge, which places the two blocking stocks over the two constructive ones; the wedge falls from 2.05 to 1.01, a 51 percent decline (Figure 1). Every component contributes; the decline is not driven by a single margin.[13] Total Friction, the broader six-dimensional composite, falls from 1.19 to 0.86, a decline of 28 percent.

[13]The decline does not depend on the exact formula. Additive-index and signed-principal-component versions of both measures, Total Friction variants that omit MAS or INDUL, reduced two-component



Notes: The display begins before the paper's year-1000 accounting window to show the direction of the earlier movement under the same stock construction. The shaded pre-1000 segment has substantially thinner textual support and provides directional historical context; it is not used for break dating or quantitative accounting. PD = hierarchy, INDIV = autonomy, MAS = competitive striving, UA = insistence on rule and certainty, LTO = future orientation, INDUL = gratification. Figure 1 plots the two wedges, $(PD + UA)/(INDIV + LTO)$ and $(PD + UA + INDUL)/(INDIV + LTO + MAS)$, as indices.

Figure 3: The post-1000 transition reverses an earlier movement toward greater cultural friction.

Because MAS enters its denominator and INDUL its numerator, lower MAS and higher INDUL each raise the ratio, partly offsetting the decline generated by the four core margins.

The early centuries rule out one mechanical worry. If the stock construction itself pushed the series in one direction, the pre-1000 record should drift the same way as the later one. It does the opposite. Before 1000, five of the six stocks move against their net 1000–1920 direction. PD, UA, and MAS rise through late antiquity and the early Middle Ages, while INDIV and LTO fall; INDUL remains low. The same construction that delivers the later decline first delivers centuries of rising friction, so the fall after 1000 is a turn in what the books say, not a drift built into accumulation. Because textual support before 1000 is sparse,[14] we use that segment only for directional context, not to date the reversal or to enter the quantitative accounting.

The fall in the wedge is not explained by later texts ceasing to take positions.[15] Conditional on taking a position, the mean stance of a passage moves by 0.511 on the −1 to +1 scale,

Innovation Wedges, and an Innovation Wedge built from rank-transformed book scores all preserve the 1000–1920 decline. Appendix Table 18 reports the external-validation signs.

[14]The corpus contains few dated works per century between roughly 200 and 1000; Appendix A.1 reports support by period.

[15]Across the six dimensions, the share of passages taking a directional position moves from 99.4 to 97.8 percent between the eleventh and nineteenth centuries—small declines on every dimension, averaging 1.6 percentage points.

roughly half the distance from neutrality to full endorsement (Appendix A.5). The comparison cannot separate changed judgment from changed subject—whether later books judge kings differently or simply write about parliaments instead. In either case, the written record endorses different things at the end of the period than at the beginning.

The decline is staggered rather than a single break. The largest hundred-year fall in the Innovation Wedge occurs over 1130–1230. At the component level, LTO records its largest rise over 1074–1174, UA its largest fall over 1130–1230, and INDIV its largest rise over 1200–1300; PD moves later, with its largest decline over 1514–1614 (Appendix Table 23). These windows are descriptive, not formal break dates, and their bootstrap ranges overlap. They therefore do not identify distinct breaks. The point estimates nonetheless place the largest century movements in the wedge, LTO, UA, and INDIV before the largest measured decline in PD.

The early movement is not monotone—several series partly retrace it in the later Middle Ages. Even so, the Innovation Wedge remains below its medieval peak in 1500. Within the 1500–1700 interval emphasized by Mokyr (2016), the dominant component movements are the decline in PD and a further rise in LTO. Taken together, the component paths suggest a sequence—future orientation, tolerance of uncertainty, and autonomy move early, while hierarchy moves later. The early-modern transformation therefore extends and deepens a transition already under way rather than initiating it.

6 Culture and Productivity Growth

We use the measured Innovation-Wedge and population paths in the model of Section 2 to quantify how the cultural transition affects productivity growth. The counterfactual is simple—hold the Innovation Wedge at its year-1000 value, let population follow its historical path, and compare the resulting productivity path to the realized one.

Quantitative framework. Using $\delta_t = \bar{\delta}\exp(-\psi\tau_t^I)$, the idea-production equation from Section 2 becomes

$$\dot{A}_t = \bar{\delta}\exp(-\psi\tau_t^I)H_{A,t}^{\lambda}A_t^{\phi}, \qquad 0 < \phi < 1, \tag{10}$$

with innovative effort assumed to be proportional to population through the constant research share $0 < s_A < 1$,

$$H_{A,t} = s_A L_t. \tag{11}$$

Population therefore determines the scale of innovative effort, while the Innovation Wedge determines how effectively that effort produces ideas.

We normalize $A_{1000} = 1$. Then, solving forward from the normalization date gives

$$A_t = \left[A_{1000}^{1-\phi} + (1-\phi) \int_{1000}^{t} \bar{\delta} \exp(-\psi \tau_s^I)(s_A L_s)^{\lambda} \, ds \right]^{1/(1-\phi)}. \tag{12}$$

Appendix D.1 derives this solution in full.

Calibration. The key parameter is ψ. It maps a change in the Innovation Wedge into a proportional change in the productivity of research and therefore determines the economic size of the measured cultural transition. We estimate it from differences in the timing of cultural change across countries and examine its sensitivity to small-cluster inference and alternative timing. The remaining parameters govern returns to research effort, spillovers from the existing stock of ideas, and the scale of idea production. We discipline them using established growth evidence and one normalization. The accounting combines these parameters with the corpus-wide path of the Innovation Wedge and the Western population path.

The ψ estimation panel needs two things for each country, a wedge and a measure of idea production. Each country's wedge is constructed in the same way as the corpus-wide measure, using only books written in that country.[16] Births of prominent scientists and inventors proxy for idea production roughly a generation later. The birth records come from Pantheon, a database of historically notable individuals identified from Wikipedia pages that appear in many languages (Yu et al., 2016). The panel begins in 1500, when country-level book support and the birth record become dense enough to compare. We retain only country-periods with at least 15 cumulative matched books, so no wedge rests on a handful of texts. Thirteen countries meet this threshold.[17]

We obtain ψ by comparing how fast the Innovation Wedge fell across countries with how many notable innovators each produced a generation later, net of common period effects and country size. The estimating equation is

$$\mathbb{E}\big[B_{c,t}^{25} \mid \tau_{c,t:t+24}^I, L_{c,t}, \alpha_c, \eta_t\big] = \exp\Big(\alpha_c + \eta_t + \lambda_L \log L_{c,t} - \psi \, \bar{\tau}_{c,t}^{I,25}\Big). \tag{13}$$

[16] Books map to countries by modern borders. See Online Appendix Section OA.A.

[17] Canada, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, the Netherlands, Spain, Switzerland, the United Kingdom, and the United States. Austria, Belgium, Norway, Sweden, Poland, and Turkey sit inside the scope but never accumulate fifteen matched books.

Here $B_{c,t}^{25}$ is the number of core Pantheon innovation births in country c over the 25-year window beginning in t, and $\bar{\tau}_{c,t}^{I,25}$ is the forward 25-year mean of its Innovation Wedge. Population $L_{c,t}$, from HYDE, the History Database of the Global Environment (Klein Goldewijk et al., 2017), enters with an estimated coefficient λ_L.[18] Country and period fixed effects are denoted by α_c and η_t.

Births are counts, and many country-periods record no innovation births at all, so we estimate the conditional mean by Poisson pseudo-maximum likelihood and cluster standard errors by country (Appendix D.2). The coefficient on $\bar{\tau}_{c,t}^{I,25}$ is $-\psi$. The estimate is $\psi = 0.631$, with country-clustered $p = 0.023$. In units, a fall of one-tenth in the wedge, roughly a century of the measured transition, raises idea production at given inputs by about 6.5 percent ($\approx e^{0.1\psi}$).

Thirteen clusters make conventional inference optimistic. Across the small-cluster procedures, the p-value ranges from 0.018 to 0.096 (Online Appendix Table OA.36). Significance at five percent therefore depends on the procedure, while the estimated magnitude remains similar. The most demanding timing specification measures the wedge over the preceding 25 years. The books then precede the innovation-birth window, so subsequent births cannot influence the regressor. This specification gives $\psi = 0.567$ with $p = 0.004$. The fixed effects absorb permanent country differences and common changes in population, recorded science, and culture. They do not remove country-specific confounds or establish that cultural change is exogenous. Online Appendix Section OA.G.4 reports a compact set of additional confound and outcome-scope checks. The estimate remains positive with university-founding and GDP controls and under country trends, while Pantheon comparisons show a similar association for state-power figures and the opposite sign for the arts.

The elasticity is estimated from within-country variation alone, so the aggregate series of innovator births across the West never enters and can serve as a check. At the panel estimate, the model's implied path fits that series with $R^2 = 0.846$. Because the aggregate objective is flat in ψ, the fit is a timing-and-shape diagnostic rather than independent identification of the elasticity. The accounting is also invariant to whether culture raises each researcher's productivity or changes the allocation of talent into research. Both channels enter idea production through the same composite term, so the counterfactual depends only on their sum (Appendix D.5).[19] Appendix D.2 gives the benchmark identification design, while Online Appendix Section OA.G reports alternative scopes and proxy assumptions.

[18] Estimating λ_L lets the data determine how strongly notable births scale with population rather than imposing proportionality.

[19] Culture could instead redirect a fixed European talent pool. Pantheon assigns individuals to their birthplaces, so adult migration does not relocate their recorded births. Restricting the panel to individuals who died in their birth country leaves the elasticity at 0.591, within a standard error of the benchmark

A second external benchmark compares the model paths with measured English TFP built from the Broadberry–de Pleijt accounts (Appendix D.3). Those data never enter the model, and the comparison is an England/Great Britain shape check on a Western-aggregate path. Under the standard produced-capital convention, measured TFP tracks the realized path across five centuries.[20] By 1870 measured TFP lies within 3.1 percent of the realized path and 29.8 percent from the counterfactual.[21]

For the accounting, we combine this estimate with the corpus-wide wedge and the HYDE population path aggregated over the Western accounting region.[22] We sum population across these countries, interpolate annually in levels, and normalize $L_{1920} = 1$.

Two remaining parameters govern the curvature of idea production. The elasticity λ determines how innovative effort enters, while ϕ determines how the existing stock of ideas affects subsequent research productivity. Bloom et al. (2020) show that maintaining a given rate of TFP growth has required increasing research effort, which motivates $\phi < 1$. We set $\lambda = 1.00$ and long-run frontier TFP growth to $g_A^\infty = 1.0$ percent per year, a calibration consistent with the same historical TFP-growth evidence. We set long-run population growth to $n = 0.29$ percent per year, approximately the rate implied by the UN medium-variant projection to 2100 (United Nations, Department of Economic and Social Affairs, Population Division, 2024). The balanced-growth condition implies

$$g_A^\infty = \frac{\lambda n}{1-\phi} \quad \Rightarrow \quad \phi = 0.710. \tag{14}$$

Appendix D.4 varies λ over $\{0.75, 1.00\}$ and ϕ over $\{0.50, 0.71, 0.85\}$. Finally, $\bar{\delta}$ and the research share s_A enter through the composite scale $\bar{\delta} s_A^\lambda$. We choose this scale so that annual productivity growth on the realized path equals 1 percent in 1920. This normalization fixes the 1920 growth rate. The long-run growth rate disciplines ϕ. Table 2 collects the calibration.

(Online Appendix Table OA.37). Reallocation within a fixed pool also cannot account for the aggregate innovation path used in the nine-century timing check.

[20] Under Broadberry and de Pleijt's own convention, which excludes dwellings and assigns capital a larger share, measured TFP rises less steeply and lands between the two paths, so that convention does not distinguish them.

[21] Measured as log points; a gap of x log points is approximately x percent.

[22] The Western accounting population includes Europe excluding Russia, with Turkey, plus the United States, Canada, Australia, and New Zealand.

Table 2: Benchmark calibration for the semi-endogenous idea-production model

Object	Benchmark value
Innovation Wedge τ_t^I	Corpus-wide annual series from Section 5, 1000–1920
Population L_t	Western accounting population (HYDE), $L_{1920} = 1$
Cultural elasticity ψ	0.631, estimated on the 13-country Pantheon/HYDE panel
Idea-production elasticity λ	1.00
Long-run population growth n	0.29 percent per year
Long-run growth g_A^∞	1.0 percent per year
Idea-stock curvature ϕ	0.710, implied by λ, n, and g_A^∞
Composite scale $\bar{\delta} s_A^\lambda$	0.0355, re-anchored so that $g_{A,1920} = 1.0\%$

Notes: The matched source pool contains the countries for which books, HYDE population, and Pantheon births can be aligned. The realized estimation panel contains 13 countries that pass the cumulative-book threshold. The headline accounting applies that panel elasticity to the corpus-wide Innovation-Wedge path and the Western accounting population. The long-run values of n and g_A^∞, together with $\lambda = 1$, imply ϕ, and the composite scale matches the 1920 productivity-growth target.

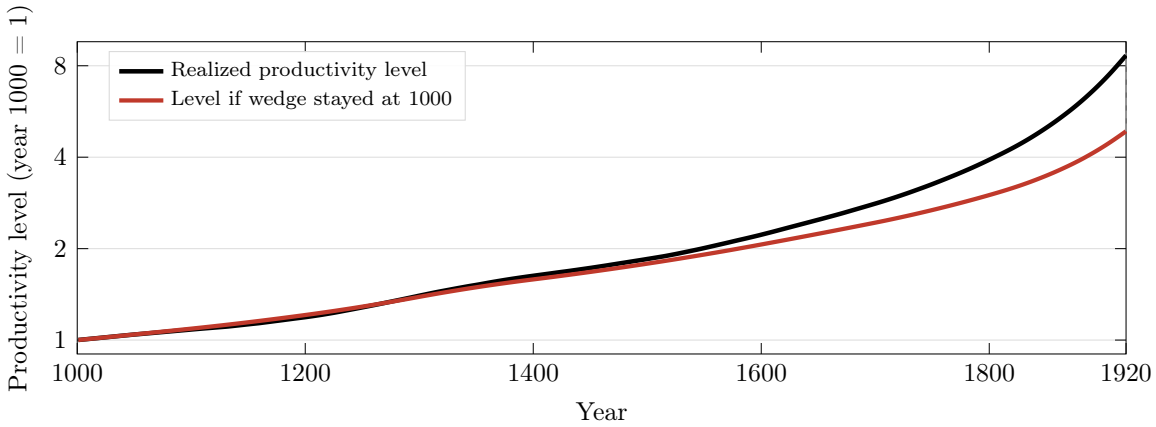
The historical transition. The benchmark counterfactual freezes the Innovation Wedge at its year-1000 value, leaving population and every calibrated parameter unchanged.[23] The difference between the realized and counterfactual paths is the wedge's contribution to productivity. The wedge falls by roughly one unit, from 2.05 to 1.01, and at $\psi = 0.631$ realized 1920 productivity is 1.78 times what it would have been had the wedge stayed at its year-1000 value.[24] Figure 4 plots the two level paths. The 1920 gap also understates the wedge's eventual effect. With the wedge held at its 1920 value forever, productivity would keep rising as the idea stock adjusts; by 1920 only about 25.4 percent of that long-run gain had arrived in log terms.[25]

The gap does not depend on where the counterfactual is anchored or on how the scores are scaled. Fixing the wedge at its 1300 value, or rebuilding the stocks from rank-transformed book scores, leaves the counterfactual in the same range (Appendix D.4). Nor does the gap depend on the wedge formula. Appendix B.5.2 repeats the accounting with the reduced wedge built from the two best-validated components, UA and LTO, both at the benchmark elasticity and with its own re-estimated ψ; holding that wedge fixed leaves 1920 productivity at 54 and 55.4 percent of its realized value, near the benchmark's 56.3 percent. It remains a robustness check rather than a replacement benchmark, since its significance weakens under exact small-cluster inference.

[23] Both paths are model solutions—"realized" means solved under the measured wedge, not observed TFP, which enters only the England comparison below.

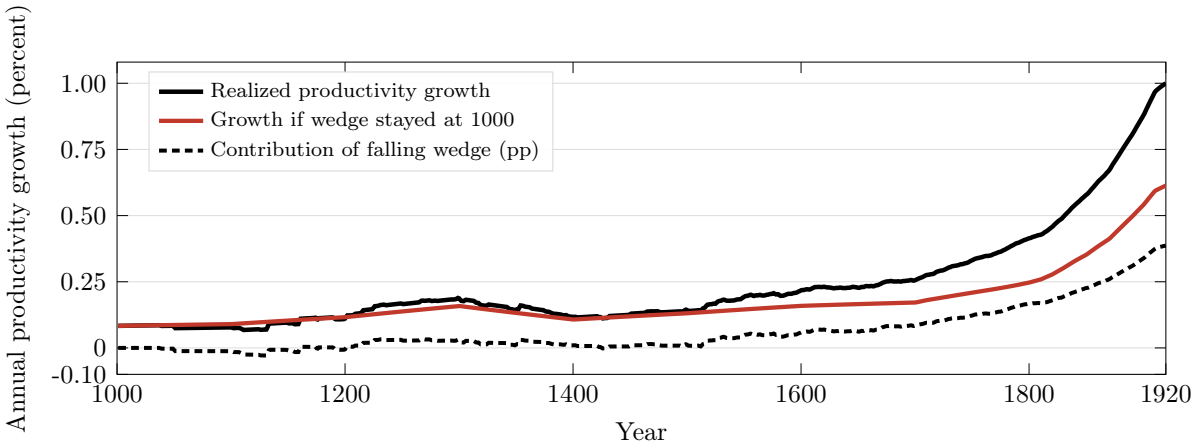
[24] At the 1920 endpoint, the lower wedge makes a given research effort $\exp\{\psi[2.05 - 1.01]\} = 1.93$ times as productive in generating new ideas. The productivity level rises by less, 1.78, because the wedge fell only gradually and knowledge spillovers are less than one for one ($\phi < 1$).

[25] The eventual rise is $\exp\{\psi \Delta\tau^I/(1-\phi)\} = 9.58$—the familiar $1/(1-\phi)$ amplification of semi-endogenous models.



Notes: The realized path is the solved productivity path implied by the benchmark semi-endogenous mapping, and the counterfactual path holds the Innovation Wedge at its year-1000 value while population follows its historical path. By 1920, the counterfactual level is 0.563 of the realized level. The vertical axis is logarithmic, so equal vertical distances are equal proportional gaps. Figure 5 shows the same counterfactual in annual growth-rate terms.

Figure 4: In the benchmark semi-endogenous mapping, holding the Innovation Wedge at its year-1000 value leaves the 1920 productivity level at little more than half the level on the realized-input path.



Notes: The figure plots annual productivity growth on the realized path, $100\log(A_t/A_{t-1})$, annual productivity growth in a counterfactual that holds the Innovation Wedge fixed at its year-1000 level, $100\log(A_t^{cf}/A_{t-1}^{cf})$, and the growth gain from the falling Innovation Wedge, $100\log(A_t/A_{t-1}) - 100\log(A_t^{cf}/A_{t-1}^{cf})$. The series are transitional, not asymptotic. The composite scale is chosen so that annual productivity growth on the realized path equals 1 percent in 1920.

Figure 5: In the benchmark semi-endogenous mapping, realized productivity growth rises above the fixed-wedge counterfactual; the growth gain from the falling Innovation Wedge reaches 0.387 percentage points by 1920.

Figure 5 plots realized growth, fixed-wedge growth, and their difference. In the calibrated model, the falling Innovation Wedge contributes 0.012 percentage points to annual productivity

growth by 1500, 0.084 percentage points by 1700, and 0.39 percentage points by 1920.[26] The contribution has two phases. Before 1500, the accounting assigns the wedge only 5.7 percent of average growth; what growth there is comes almost entirely from population scale. Between 1500 and 1700, model-implied annual productivity growth roughly doubles, rising by 0.120 percentage points, and the falling wedge accounts for 64.8 percent of that rise. Only after 1800 does the scale channel reassert itself, delivering the majority of a much faster acceleration. The wedge's share of the 1700–1920 rise is 40.7 percent. Cultural change therefore does much of the early work of the model-implied acceleration, and culture and population reinforce each other later. The window in which the wedge does most of its work, 1500–1700, is precisely the one Mokyr (2016) identifies as the transformation of Europe's elite culture of ideas. The gradual acceleration is also consistent with Crafts (1985). British productivity growth rose over a long transition rather than in a single sharp takeoff, and the model offers one possible mechanism for that pattern.

Accounting robustness. Rebuilding the accounting from the alternative-prompt and DeepSeek rescorings of Section 3, holding the elasticity fixed, gives 1920 contributions between 0.29 and 0.39 percentage points, at or below the benchmark (Appendix A.4). Rebuilding the stocks from 1500, so that nothing written before the print era enters, gives 0.356 against the benchmark's 0.387; the early canon is not doing the work (Online Appendix Table OA.47). Resampling the corpus itself keeps the wedge's share of 1920 growth between 27.7 and 53.5 percent under the same fixed elasticity (Appendix D.4).

Which cultural changes matter? The Innovation Wedge combines a blocking numerator, $PD + UA$, and a constructive denominator, $INDIV + LTO$. Table 3 holds each block at its year-1000 value while the other dimensions follow their realized paths. These are separate, overlapping counterfactuals rather than additive shares. Holding the blocking pair alone leaves 1920 productivity at 80 percent of its realized value, and holding the constructive pair alone leaves 72.2 percent. Both sides of the wedge matter, and the decomposition matches the timing evidence of Section 5—Long-Term Orientation, Individualism, and Uncertainty Avoidance move earliest, and the later fall in Power Distance reinforces what those earlier changes began. Appendix D.4 reports the individual-dimension counterfactuals.

[26] Using the former full-26 population scope at the same panel elasticity gives a 1920 contribution of 0.36 percentage points, or 36 percent of 1920 productivity growth; excluding Australia and New Zealand instead leaves the contribution essentially unchanged (0.386 percentage points).

Table 3: Innovation-Wedge block counterfactuals

Counterfactual	Block	Fixed share	Δg_{1700}	Δg_{1920}
PD fixed	PD	0.880	0.022	0.079
UA fixed	UA	0.902	0.017	0.075
INDIV fixed	INDIV	0.927	0.016	0.060
LTO fixed	LTO	0.803	0.026	0.106
PD+UA fixed	Numerator block	0.800	0.038	0.150
INDIV+LTO fixed	Denominator block	0.722	0.045	0.199
All four fixed	Full Innovation Wedge	0.563	0.084	0.387

Notes: Psi is fixed at the 13-country panel estimate. Every row uses the corpus-wide all-books wedge, the Western accounting population, and the normalization that realized productivity growth equals 1 percent in 1920. Each counterfactual fixes the listed dimension block at its year-1000 value while leaving the remaining blocks and historical population path unchanged.

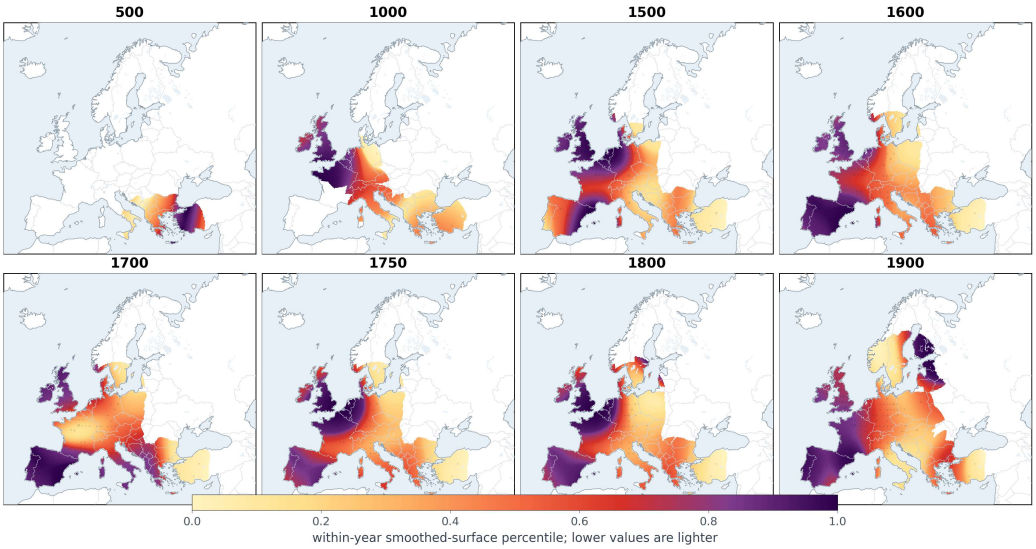
7 The Geography of the Measure

The aggregate series treats the West as one place. This section puts the measure on the map instead, following Europe's low-wedge innovation frontier over time and asking whether the culture surrounding a place predicts how fast it later urbanizes. The frontier drifts northwest over the period, while nearby hierarchy and uncertainty avoidance predict less urbanization and nearby individualism predicts more.

The country stocks used above pool literary origins within countries. The spatial construction instead places books on fixed one-degree cells, applies the recursion from Section 3.3 within each cell, and compares locations within the same year. Locally geocoded source works enter their own cells, while country-level source origins enter the cell containing the modern capital of the standardized country. The urbanization regressions rebuild the stocks on three-degree cells from point-resolved books only. Appendix E.1 summarizes the descriptive and regression samples; the complete origin rules, stock recursion, smoothing, support, and weighting are in Online Appendix Section OA.H.

A caution applies throughout this section. The aggregate series pools thousands of books, so book-level noise largely averages out. A single grid cell rests on far fewer books, and its stock is correspondingly noisier, most of all where support is thinnest. The spatial evidence carries a lighter burden than the aggregate series, and we read it as descriptive and predictive rather than precise.

Figure 6 shows the movement in two ways. Panel A reports the within-year relative surface, while Panel B weights stocked cells toward the lower end of the Innovation-Wedge distribution so that the path follows the low-friction frontier rather than the continental average.



A Within-year relative smoothed Innovation-Wedge stock surface.

B Soft lower-tail path of the Innovation Wedge.

Figure 6: The within-year surface and soft lower-tail center describe the changing geography of Europe's Innovation Wedge.

Notes. Panel A shows fixed one-degree cultural-capital cells for Europe excluding Russia, with Turkey retained. Each displayed year's smoothed values are converted to within-year percentiles, so lighter areas rank lower and darker areas rank higher within that year. The surface is descriptive of spatial ordering rather than absolute magnitudes. Locally geocoded works enter their own cells, while country-level origins enter the cell containing the modern capital of the standardized country. Panel B reports the connected lower-tail path from the same stocked cell panel. Within each 100-year bin beginning in 500, cells are ranked by their volume-stock-weighted period wedge, weighted toward the low tail by normalized rank, and averaged by cell centroid. The plotted labels are node years, and the path is displayed through its 1900 node. Both panels are descriptive summaries and are separate from the three-degree point-only urbanization design.

The panels answer different questions because the relative surface removes the common continental decline while the lower-tail path follows the frontier. The path begins in the Mediterranean and Adriatic support, passes through the Italian-Adriatic and central European corridor, and settles inside the denser northwestern and central European stock; Appendix E.3 applies the same construction to North America and the pooled transatlantic stock. The relative surface reveals spatial contrasts that the frontier path cannot. Descriptively, the Iberian Peninsula darkens after 1500, in the centuries of the Inquisition, while interior France lightens across the 1700 and 1750 snapshots, in the decades of the Enlightenment. The figure does not attribute those contrasts to those or any other events; Appendix B.3 reports the disciplined event comparisons.

We then ask whether nearby cultural stocks predict urbanization over the following century on the point-only three-degree panel.[27] City growth in this era responded to more than culture—the spread of printing mattered too (Dittmar, 2011)—so culture enters here as one predictor among several. We merge decade-average stocks with HYDE urbanization shares. Baseline decades run from 1000 through 1820, with outcomes through 1920 because book inflows end in 1920. We estimate

$$\Delta u_{i,t\to t+100} = \beta^m \left(\bar{c}_{i,t}^{m,\text{local}} - c_{i,t}^m \right) + \gamma^m c_{i,t}^m + \rho u_{i,t} + \alpha_i + \tau_t + \varepsilon_{it}^m, \tag{15}$$

where $u_{i,t}$ is baseline urbanization, $c_{i,t}^m$ is the cell's own stock, and $\bar{c}_{i,t}^{m,\text{local}}$ averages observed neighbors within the stated spatial radius. The coefficient β^m sits on the neighbor-minus-own gap, while ρ measures the association with baseline urbanization. The regressions include own culture, baseline urbanization, cell effects, and decade effects, with Bartlett–Conley spatial and temporal standard errors. The most demanding specification absorbs country-by-period shocks.

The baseline uses 712 cell-decades in 43 cells. A neighboring PD stock 10 points higher predicts 3.7 percentage points less urbanization over the next century; the same gap in UA predicts 5.4 less, and in INDIV 3.9 more. The baseline LTO gradient is imprecise, although it remains positive. With country-by-period effects added (Panel B of Table 4), the three margins keep their signs but the estimates are smaller and less precise.

Appendix E.2 reports the full eight-measure and four-specification table with MAS, INDUL, Total Friction, the Innovation Wedge ratio, joint components, and volume-weighted neighbors, while Appendix E.4 reports alternative grids, support thresholds, non-overlapping windows,

[27] The sample covers Europe excluding Russia, with Turkey retained; Appendix E.4 reports the global perimeter and other grids.

Table 4: Nearby hierarchy, uncertainty avoidance, and individualism predict later urbanization on a three-degree European grid.

	PD	UA	INDIV	LTO
Panel A. Baseline separate regressions				
Local gap	-0.373^{***}	-0.535^{***}	0.388^{***}	0.054
	(0.057)	(0.077)	(0.075)	(0.082)
Cell FE	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes
Country-period FE	No	No	No	No
Inference	Conley	Conley	Conley	Conley
Observations	712	712	712	712
Cells	43	43	43	43
Panel B. Country-by-period fixed effects				
Local gap	−0.066	−0.139	0.031	−0.040
	(0.056)	(0.100)	(0.061)	(0.078)
Cell FE	Yes	Yes	Yes	Yes
Decade FE	Yes	Yes	Yes	Yes
Country-period FE	Yes	Yes	Yes	Yes
Inference	Conley	Conley	Conley	Conley
Observations	712	712	712	712
Cells	43	43	43	43

Notes. The table reports the four Innovation-Wedge components from the three-degree urbanization design of Section 7. PD denotes Power Distance, UA Uncertainty Avoidance, INDIV Individualism, and LTO Long-Term Orientation. Panel A reports separate regressions. Panel B adds country-by-period fixed effects and is the most demanding specification in the main analysis. Every regression includes cell and decade fixed effects, the cell's own cultural stock, and baseline urbanization. Parentheses contain Bartlett–Conley standard errors with the same spatial and temporal cutoffs as Table 31 in Appendix E.2. The sample uses point-resolved books for Europe excluding Russia, with Turkey retained. ***, **, and * denote significance at 1, 5, and 10 percent.

regional stability, reverse timing, and spatial placebos. The grid, perimeter, assignment, and functional-form checks preserve the pattern, but the high-support restriction does not (Online Appendix Section OA.I.3). The associations are regional, attenuate when components enter jointly or country-period shocks are absorbed, and are predictive correlations, not evidence of cultural diffusion or causal effects. They complement the aggregate accounting of Section 6.

8 Conclusion

Reading more than 23,000 books from the Western canon—judging what each text endorses rather than which words it contains—delivers an inherited stock of cultural capital that moves slowly and a long way. Between 1000 and 1920, the cultural barriers most tied to innovation fall by half. The fall begins before the early-modern window where existing cultural accounts start and runs through industrialization.

The available checks point the same way. Blinded experts recover the book-level contrasts, and modern surveys recover the country orderings, most strongly on the model's margins, not through income, and not outside the corpus's scope. Prior-free scorers recover the same aggregate paths despite weak book-level agreement with the staged reading, and an independent archive of roughly 5,000 books, scored with the pipeline unchanged, reproduces the decline.

In a semi-endogenous model of idea production, with the cultural elasticity estimated from within-country variation, holding the Innovation Wedge at its year-1000 value leaves 1920 productivity at 56.3 percent of its realized value. Because knowledge spillovers are less than one for one, this is a level effect. Population pins down long-run growth, while culture shifts the level path. But the transition between levels runs for centuries, and over the nine centuries we measure, the falling wedge raises growth all along the path. On this horizon, a level effect of this size is what matters. The timing also matters for the cultural narrative of growth. Much of the decline is already underway before 1500, so the cultural preconditions that Mokyr (2016) locates in the early-modern centuries appear to have been accumulating well before the Republic of Letters. The geography points the same way—on a three-degree European grid, places surrounded by deference and the fear of the untried urbanize less over the following century, and places surrounded by autonomy urbanize more.

Three limitations remain. First, the expert-recognition exercise is a high-contrast book-level check, not an audit of every score. Second, the spatial evidence is predictive, not causal. Religion, institutions, literacy, printing, universities, and elite knowledge networks may all carry the cultural stock we measure. Controlling them away would not isolate culture; it would remove the channels through which culture travels. Whether shocks to the circulation of ideas actually move this stock—printing, the Reformation, censorship, and the changing geography of publication—is a question the fixed-grid machinery built here is designed to answer, and it is the subject of companion work in progress. Third, the corpus is centered on Western elite literary culture. That may be the right place to begin for a paper about modern growth, but it is not the end of the project. Extending the corpus across languages and traditions will show whether the measure applies beyond the canon it was built on.

Culture is often treated as the residual left over when measurement runs out—important, persistent, and unmeasured. Nine centuries of books suggest it does not have to be treated that way. Culture can be read from what societies wrote and kept, checked against independent evidence, and entered into a growth model in measured units. On the evidence here, the accounting assigns it a large role in the growth of ideas.

APPENDIX CONTENTS

A Measurement Pipeline and Robustness

This appendix expands the measurement construction of Section 3.2. It proceeds from corpus assembly and its support over time to a worked book-level example and within-book sampling, then turns to prompt, model, and method robustness, silence and stance, and the progress-vocabulary comparison.

A.1 Corpus Assembly

This subsection records corpus assembly from raw catalog to scored panel. Screening keeps 83.4 percent of the deduplicated catalog, 88.9 percent of screened works are successfully scored, and 74.1 percent of cataloged works therefore enter the analysis. Scoring proceeds in waves—operational batches, not historical periods. An alpha wave covers canonical, high-priority works first; the rest follows in roughly 1,000-book blocks (waves 000 through 025), each stratified by broad year tier, sorted within tier by priority, downloads, and classification confidence, and arranged so repeated authors fall in different waves. A wave fixes only the order of scoring and whether a book is read on every passage or on a within-book sample (Appendix A.3); it has no effect on where the book enters the series.

Cleaning uses LLM-assisted boundary detection and a trained machine-learning model, with manual review for ambiguous cases. Boilerplate headers and footers, translator notes, editorial commentary, tables of contents, footnotes, appendices, and other paratext are removed when line spans can be identified and retained only when clearly part of the authored text; ambiguous cases go to review, and a book is scored only after its boundaries are accepted. Each cleaned text retains a complete record with line ranges, retained paratext, snippets, confidence, search attempts, review reasons, and fallback markers. Without this step, the scorer would often mistake archival scaffolding for cultural content.

The production run—the alpha wave plus waves 000 through 025—yields 23,064 scored books, every one with complete six-dimension scores, a valid analysis year, and a geocoded origin. Almost nothing is lost between raw scoring runs and the assembled panel. A handful of books fail before assembly, some because the provider's content filter refused their passages outright, including Wilde's *Salomé*, and a small set of works leaves through the non-Western screen. The omission ledger in the replication package names each book. The geography spans 113 standardized country labels, 27 of which meet the 15-book country-stock threshold; dropping country-level locations leaves 18,403 point-resolved books across 703 one-degree cells. Each

stage writes its own records—cleaned catalog rows, dated works, origin assignments, selection ledgers, review statuses, geocode outputs—so a reader can trace why any book entered, left, or remained; Online Appendix Section OA.B.6 reports the full audit table.

Coverage is uneven over time. The corpus averages 2.6 books per decade in 1000–1499, 24.4 in 1500–1699, 57.1 in the eighteenth century, and 1523.7 in 1800–1920. Figure 7 shows the pattern by decade, and Online Appendix Section OA.F reports how that support carries into the inherited stocks, for the full corpus and the grid-matched sample. The early sample is informative but thin, so early-period movements should be read together with the counts behind them; Appendix C.1 reports what this support looks like once books accumulate into the inherited stocks.

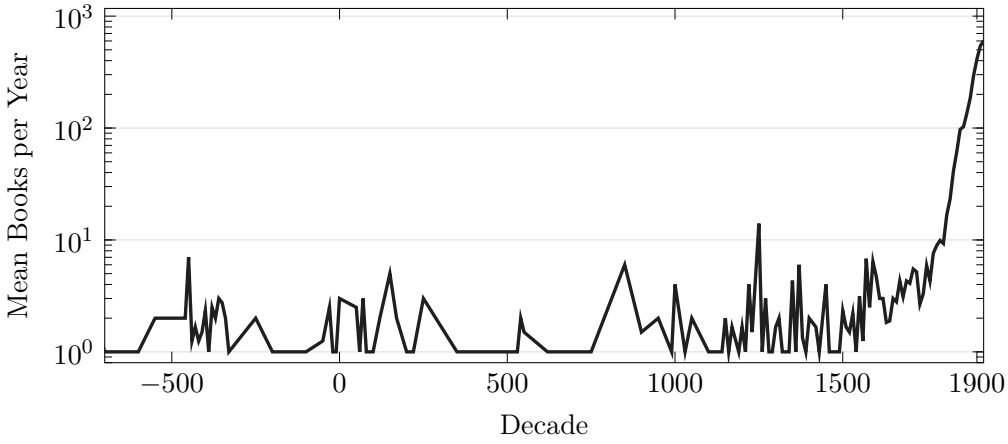


Figure 7: Book support becomes much denser in the early modern and industrial parts of the sample.

Notes: The figure plots mean dated scored books per year by decade. The vertical scale is logarithmic. The display is restricted to the paper's main historical support and ends in 1920.

A.2 Worked Example: From Text to Book Score

The implementation calls the fixed-length passages of Section 3.2 chunks; this appendix and the replication package use that term. The worked records show how the scoring procedure operates at book level. Table 5 summarizes the two records used here as an example. For each book, the record preserves the dataset's bibliographic information, the context pass, the per-chunk task reports, the final classification text, and the extracted score vector.

The two examples are useful because they separate depiction from valuation. *The Prince* is an elite manual of statecraft. It treats hierarchy, command, and princely authority as the facts of political life and often as instruments of survival. It therefore scores high on

Table 5: Two worked examples from the scored book records.

Book	Year	Canonical origin	Chunks	PD	INDIV	MAS	UA	LTO	INDUL
The Prince	1513	Florence, Italy	6	100.0	100.0	100.0	66.7	100.0	0.0
Venture Smith Narrative	1798	New London, Connecticut, United States	2	0.0	100.0	100.0	50.0	100.0	0.0

Notes. Both rows are drawn from the stored scoring records. Analysis year and source-work origin are the values used by the paper; Venture Smith's raw final-year field is 1795, while the resolved analysis year is 1798. Scores are saved book-level averages across successful selected chunks. PD denotes Power Distance, INDIV Individualism, MAS Masculinity, UA Uncertainty Avoidance, LTO Long-Term Orientation, and INDUL Indulgence. These columns report book-level score points.

Power Distance even though it also scores high on Individualism, Masculinity, and Long-Term Orientation. The book praises agency, strength, and practical planning inside a hierarchical political world. Its Uncertainty Avoidance score is mixed rather than low. Some chunks praise adaptation and new measures, but other chunks value order, control, and political discipline, leaving the book-level average at 66.7. Venture Smith's slave narrative is almost the opposite case for Power Distance. It is a first-person account of enslavement, sale, forced labor, self-purchase, and family reconstruction. The text depicts extreme hierarchy, but the normative force of the narrative is against domination and ownership of persons. That is why it can score low on Power Distance while scoring high on Individualism and Long-Term Orientation. The valued movement is toward self-ownership, work, property, and durable independence.

The word-count audit is useful precisely because it is transparent, but this pair shows why it cannot be read book by book without context. In the word-count audit, the Power Distance dictionary assigns *The Prince* a score of 66.7 and Venture Smith's narrative a score of 100.0. Our staged scorer assigns the opposite ordering—100.0 for *The Prince* and 0.0 for Venture Smith. Word-count evidence remains informative in aggregate because the dictionary-based stocks produce similar broad time-series patterns. At book level, a dictionary rule can mistake represented hierarchy for endorsed hierarchy. Venture Smith's narrative contains many words associated with command, ownership, and enslavement because it is describing slavery, but the text's normative force is against that hierarchy. Table 9 in Appendix A.4 works through both books and Swift's *A Modest Proposal* in detail.

A.2.1 How A Book Is Scored

The scoring procedure is deliberately staged. Turning a long book into six numbers requires separate judgments about genre conventions, irony, characters who speak against their author, and behavior that is depicted but not admired. A model asked to make every judgment at

once will misread some. Staged, no single step is hard. The first step establishes the book's context and tone. The passage reader then uses that context to judge what the author praises or condemns and whether the praise is sincere, while a separate report records what people do and how the text treats those actions. The final step weighs the reports into scores.

The stages are those of Section 3.2. The context stage does not receive the full book. It receives a concatenated excerpt built from the opening 10,000 characters and the closing 10,000 characters of the cleaned text, prefixed by the book's title and author. Its job is to produce a short interpretive lens. That lens is then passed forward to the chunk-level scorers. The literary scorer receives the lens together with one chunk. The behavioral scorer receives the full chunk and the literary-stage output as prior context. The synthesis scorer receives both prior reports and returns the final net normative score. The order matters. Context first, then literary framing and behavioral evidence, and only then synthesis. A deterministic parser extracts the score vector from the synthesis text, and book-level scores average those vectors across successfully scored selected chunks.

The scoring design uses a stronger model for the book-level context pass, a faster model for the repeated literary-analysis, behavioral-extraction, and synthesis passes, and the parser to extract the six-score vector from the final synthesis text. Although the production settings retain Gemini 2.5 Flash-Lite as a dormant data-clerk route, the production runner never invokes it. All 99,919 successful selected chunks report deterministic extraction, while a failed parse is closed for that attempt and can be recovered only by a fresh chunk retry. No stage substitutes for another. The context pass frames the book, the analyst passes produce evidence, the synthesis pass adjudicates the evidence, and the parser reads the score vector.

Online Appendix Section OA.C prints the verbatim effective instructions for the context, literary-analysis, behavioral-analysis, and final-adjudication stages, from maintained, hash-provenanced sources. This appendix retains the staged sequence and the scoring boundary needed to interpret the measure.

A.2.2 Why Long-Term Orientation Receives A Special Rule

Most dimensions are scored directly from the definitions of Section 3.2—the scorer asks whether the passage endorses one pole or its opposite, and the definitions carry the judgment. Long-Term Orientation is the hardest dimension to score because ordinary language has many kinds of future orientation. A saint's life can praise fasting, obedience, martyrdom, heaven, salvation, resurrection, and divine judgment. A dynastic history can praise lineage, ancient

law, continuity, and legacy. Those are future-facing in a broad semantic sense, but they are not automatically high LTO in the Hofstede sense used here. The concept is narrower. It covers pragmatic adaptation, thrift, education, investment, institutional learning, strategic patience, and this-worldly planning for practical future payoff.

The scorer therefore restricts LTO endorsement to this material and pragmatic sense, so that sacred temporality alone—heaven, salvation, eternity, preservation of ancient truth—does not count toward it. The restriction matters in the data. Against an independent dictionary-based LTO rule, on the 275 alpha-wave books both current instruments cover, the rank correlation moves from 0.029 to 0.122. Prompt variants that chase the dictionary agreement harder can reach higher LTO-only correlations, but they move the other dimensions and the implied wedges more aggressively. We retain the stricter boundary as the baseline and report the variants as robustness checks. Appendix A.4 reports the cross-model and cross-prompt comparison. The effective literary- and behavioral-analyst instructions implementing the restriction are reproduced in Online Appendix Section OA.C.

The rule is not a claim that religious books must be low-LTO. A religious text can score high LTO when it praises practical education, durable institutions, worldly planning, material investment, adaptive reform, or patient work for this-worldly returns. The rule says only that sacred temporality—God's plan, heaven, salvation, eternity, prophecy, or preservation of ancient truth—does not by itself establish the economic patience-and-adaptation margin that this paper needs LTO to measure.

A.2.3 Provenance Setup: Source Place/Year Inference, Place Canonicalization, and Geocoding

The provenance procedure is separate from scoring and runs before the book record is finalized. Three maintained instructions govern it. The source-inference instruction asks for the underlying source work rather than any later edition, requires an explicit relationship classification (original work, faithful translation, adaptation, commentary, and five further classes), and demands the narrowest defensible origin—a city when one is established, with an explicit rule against broadening a known city to a country because the broader answer feels safer, and equally explicit permission to return a broad polity or a century-level date when that is the honest best answer. Publication, printing, and performance places are excluded. Each record returns with a confidence tier and up to two independent supporting web citations, with Project Gutenberg and pages derived from its metadata excluded as evidence. The canonicalization instruction then treats the inferred place as a default to

preserve, normalizing or honestly broadening it into a single canonical form, never inventing precision and never returning multiple places. Geocoding is deterministic where possible; only places that fail local lookup go to the geocoding-normalization instruction, which maps a historical name to its modern geocodable equivalent—Weimar, Germany rather than a historical wrapper—without changing the underlying origin. Coordinates and a granularity class are recorded with each resolved place. Title matching and year parsing are deterministic rules, with no model in the loop.

The unit is a source-work observation, not a sequence of later editions. Later reprints, English editions, rediscoveries, and Gutenberg edition locations do not create new dated-origin observations. This is both practical and conceptual—Gutenberg coverage of later editions is incomplete, and using observed reprints would mix historical circulation with modern archive selection. If an older work is rediscovered and shapes later culture, that influence should appear indirectly through the later works written under that influence.

Unresolved cases remain broad or enter review rather than receiving false precision. Online Appendix Section OA.C prints all three instructions verbatim.

A.2.4 A Full Scoring Trace: *The Prince*

The example begins with a simple historical assignment. The English Gutenberg text is treated as a faithful translation of Machiavelli's *Il Principe*, so the scored object is dated to 1513 and assigned to Florence, Italy, where Machiavelli wrote the work while in exile. The canonical record keeps that city-level origin and year with high confidence, and the geocoder places it at latitude 43.769600 and longitude 11.255800. The relevant object is a Renaissance statecraft text tied to the historical source work rather than to a later English edition.

On the scoring side, the stored record includes the context pass's output verbatim. For *The Prince* it begins.

> "The global tone of The Prince is a complex fusion, not a simple classification. On its surface, it operates as a stark Historical Chronicle, a pragmatic and coldly analytical political science manual. Machiavelli's authorial voice is that of a clinical observer, using historical case studies to deduce timeless principles of acquiring and maintaining power. However, this clinical analysis is framed by a profound sense of Tragedy, most evident in the final chapter's passionate "Exhortation to Liberate Italy." Here, the voice shifts from detached analyst to desperate patriot, lamenting Italy's subjugation and pleading for a redeemer. . . ."

The complete context record appears in Online Appendix Section OA.C.

All 6 chunks are selected and successfully scored, and each contributes one score vector with entries in $\{-1, 0, +1\}$. The chunks agree on every dimension except UA—every chunk scores $+1$ on PD, INDIV, MAS, and LTO and -1 on INDUL. UA splits, 4 chunks at $+1$ against 2 at -1. The book-level rows of Table 5 are nothing more than these 6 vectors averaged and mapped linearly to the 0–100 scale of Section 3.2. Unanimous $+1$ maps to 100, unanimous -1 maps to 0, and the UA mean of $+1/3$ maps to 66.7. The book is therefore not classified as low-UA overall; the trace shows why a hierarchical statecraft text can still contain adaptive, low-UA passages.

Chunk 6 is a useful trace because it shows why the scoring prompt can assign high LTO without mechanically assigning high UA. The literary analyst reads the final chapters as praising individual agency, strong leadership, competitive achievement, adaptation, and future-oriented institutional reform. Its UA judgment is negative because the text calls for "new laws and new ordinances" and says that "the old order of things was not good" rather than adhering to old methods.

The behavioral stage is even more explicit. For the same chunk, the recorded behavioral output gives an evidence scan rather than a short verdict. Its LTO reasoning emphasizes Machiavelli's criticism of short-sighted princes and the metaphor of building "defences and barriers" against raging rivers and future floods. Its UA reasoning points in the other direction because the chunk praises adaptability and innovation rather than fixed rules.

The synthesis stage then resolves the reports into one final vector. For this chunk, the literal language is as follows.

> "UA: -1 ... The reports indicate that low uncertainty avoidance behavior (adaptability, innovation, breaking from old orders, strategic deception) is depicted and explicitly praised as a means to achieve practical results and honor, rather than adherence to rigid rules. ... LTO: +1 ... The reports show high long-term orientation behavior (strategic planning, patience, actions for future stability and material outcomes) being depicted and framed as pragmatic and essential for statecraft and restoration. ... INDUL: -1 ... The reports consistently depict low indulgence behavior (restraint, discipline, calculated measures, suppression of personal desires for political goals) and frame it as effective and necessary for state stability, with no praise for personal freedoms or immediate gratification."

The final saved vector for the chunk is $(1, 1, 1, -1, 1, -1)$. That is exactly the vector carried by the recorded score object and final synthesis text. The example also shows the role of the strict LTO boundary. The analysts assign the text high LTO because they connect its future

orientation to adaptive statecraft, institutional reform, and practical worldly security rather than to destiny, divine history, or inherited tradition.

The detailed walk-through uses *The Prince* because its provenance chain and chunk-level disagreement are compact enough to print cleanly. Venture Smith's narrative—the second record of Table 5, assigned analysis year 1798 at New London, Connecticut, with raw final-year field 1795—receives the same treatment in Online Appendix Section OA.C. The abbreviated worked records for both books, including passage-level reports and synthesis text, are printed there; the complete machine records remain in the replication package.

A.3 Passage Sampling

Scoring every chunk of every book with a frontier model is expensive at corpus scale. At Gemini Batch API rates, a complete-book run averages about $0.33 per book and a sampled run about $0.08, so all-chunk scoring of the full corpus would have cost roughly $7,600—about 2 times that at standard rates—against roughly $2,450 realized, or $0.11 per book (Online Appendix Section OA.C.4 reports the full accounting). The alpha wave and waves 000 through 002 were scored on every chunk; later waves score a within-book sample. The sample size is 10 percent of a book's chunk count, rounded up, bounded below by three and above by five; books with fewer than three chunks are read in full. Formally,

$$n_{\text{sel}} = \min\big(N_{\text{chunks}},\, 5,\, \max(\text{floor},\, \lceil 0.10 \cdot N_{\text{chunks}} \rceil)\big), \tag{16}$$

where the floor is three for books with at least three chunks and the book's full chunk count otherwise. In practice, books with up to thirty chunks contribute three, thirty-one to forty contribute four, and forty-one or more contribute five. The selected chunks are the midpoints of equal intervals through the text, so the sample spans beginning, middle, and end—an eighteen-chunk book selects chunks 4, 10, and 16. Table 6 collects these settings.

Three facts make the sampling close to immaterial. First, the five-chunk ceiling truncates the ten-percent rule for only 0.6 percent of sampled books (130 of 20335); the complete chunk-count distribution is in the replication package. Second, sampling does not move the scores. Rescoring the all-chunk waves under the sampling rule leaves book-level scores almost unchanged (Table 6), and a test-retest program plus a completion experiment that scored every remaining chunk of the sampled books find no detectable bias against complete readings (Online Appendix Section OA.D). Third, because the sample is spread evenly through the

text, what survives at book level is idiosyncratic noise rather than systematic bias, and it averages out once thousands of books are accumulated into the stocks.

Table 6: Chunk-sampling agreement and configuration

Panel A. Production chunk sampling

Rule	Books	Mean Pearson	Mean Spearman	Innovation-Wedge path correlation	Innovation-Wedge 1920 ratio
10 percent, min 3, max 5	2753	0.925	0.911	0.983	0.994

Panel B. Stock-level agreement under chunk sampling

Sample share	Books	Dimension-path correlation	Innovation-Wedge path correlation	Innovation-Wedge 1920 ratio
10 percent	2754	0.988	0.950	0.999
25 percent	2754	0.995	0.981	1.000
50 percent	2754	0.998	0.993	1.000
75 percent	2754	0.999	0.998	1.000

Notes. Panel A applies the production sampling rule to the all-chunk initial alpha wave and waves 000–002. Panel A reports 2,753 books because its comparison sample excludes one of the 2,754 complete-passage records under the non-Western screen, while Panel B reports all 2,754 records. Panel B compares inherited stock paths built from sampled chunks with all-chunk paths. Complete draw and coverage matrices remain in the replication package.

A.4 Prompt, Model, and Method Robustness

This appendix tests the scoring method itself, twice. The first exercise rescores fixed samples of already-scored books under alternative prompts and an independently trained model family. The second discards the staged reading entirely and rescores the corpus with prior-free instruments. Both leave the aggregate paths in place while agreeing with the staged scores only modestly book by book. A third strand—independent end-to-end reruns and repeated rescoring of fixed book samples—is summarized with the noise accounting in Section 3.2 and reported in full in Online Appendix Section OA.D.

Prompt and model rescoring. The first exercise holds the model fixed and varies the staged instructions. Most of the variants alter the strict LTO rule of Appendix A.2, and the results reported there explain the design—variants that chase higher dictionary agreement on LTO move the other dimensions and the implied wedges more aggressively, which is why the stricter boundary remains the baseline. The second holds the instructions fixed and rescores the same books with DeepSeek v4, a model family trained independently of Gemini, so shared training cannot manufacture agreement. The samples are the alpha wave—the canonical high-priority tranche that production scored on every chunk (Appendix A.1),

which frees the comparison from within-book sampling noise—and a sample of ordinary wave books rescored across model families. Table 7 summarizes the agreement. Book-level agreement is modest, while the analysis uses aggregates. Rerunning the growth accounting with the disagreement treated as measurement noise and the panel elasticity held fixed leaves the headline contribution within a narrow band around the baseline. Online Appendix Section OA.D reports every dimension and comparison and states the audit's scope; it is not corpus-wide.

Table 7: Prompt and model score agreement

Sample	Comparison	Books	Median Spearman	Range across dimensions
Alpha sample	Gemini vs. DeepSeek v4	275	0.786	0.273–0.893
Alpha sample	Prompt and LTO-rule variants	275	0.770	0.080–0.916
Sampled-wave sample	Gemini vs. DeepSeek v4	250	0.629	0.053–0.780

Notes. The table summarizes book-score Spearman correlations across the six dimensions for rescoring comparisons within the named samples. The alpha sample is the canonical tranche scored on every passage; the sampled-wave sample is drawn from ordinary waves. DeepSeek v4 is the independently trained second model family, and the variants are instruction and LTO-rule variants. The Online Appendix reports every dimension and comparison. LTO denotes Long-Term Orientation.

Prior-free full-corpus methods. We also test whether the aggregate cultural paths depend on the scoring method itself. Two prior-free instruments rescore the corpus with no staged reading anywhere in the loop. The exact-lexical count tallies fixed high- and low-pole seed vocabularies for each dimension—the same dictionary instrument as the word-count audit of Section 3.2, here applied to the whole corpus rather than the audit sample. The embedding-expanded version widens each pole to semantically similar terms before counting. Both instruments count pole-term matches across every fixed-length passage of each book, scaled by passage length, and both then feed the identical downstream machinery—the same book averaging, the same stock recursion, the same wedge formulas—so any divergence between the resulting paths and ours reflects the scoring method and nothing else. Online Appendix Section OA.D prints the seed lexicons, normalization and matching rules, embedding model, expansion thresholds, and aggregation details.

The instruments run on every text that survives three screens—source resolution, Gutenberg-boilerplate removal, and a requirement that a book keep at least 1,000 tokens and produce counts on all six dimensions. That leaves 22,722 usable texts, of which 22,696 belong to the 23,064-book panel. The comparisons below use those matched books. The panel books that drop out fail one of the same three screens, and a few usable texts fall outside the panel. These are method comparisons on the same books, not coverage checks. A dictionary

misranks a book when represented and endorsed values diverge, but such books are a minority, concentrated in devotional writing and juvenile fiction (Online Appendix Table OA.13), so the aggregate paths can survive even though book-level agreement is modest. One caveat remains. The staged scorer could in principle recognize a canonical work and lean on what it already knows about the country that produced it, and removing titles and authors would not prevent that, because famous texts are recognizable from their words alone. The staged reading is needed to interpret individual books; the long-run paths do not depend on it.

At the book level the two instruments agree modestly. A book's component scores imply a book-level wedge—the wedge formula applied to the book's own scores rather than to any stock—and the comparison keeps the 22,012 books for which both methods deliver finite components and a nonzero denominator. The per-book rank correlation between the exact-lexical proxy and our staged score is 0.13 for Total Friction and 0.16 for the Innovation Wedge; the embedding version is comparable at 0.21 and 0.16. These book-level comparisons are flows—the scores of the books themselves—as distinct from the inherited stocks accumulated from them. They are real but partial shared signals between methods that differ by construction, and far from the near-identity that would make agreement at coarser frequencies mechanical. Averaging many books amplifies that modest shared signal relative to book-level noise.

The methods converge as scores are averaged at the frequencies used in the analysis. Table 8 reports the climb. Annual flows already lift the correlation well above the book level, decade flows lift it further, and once periods are required to hold at least ten scored books—so that thinly covered early decades do not dominate—the embedding proxy tracks the Total Friction and Innovation Wedge paths at Spearman 0.83. The inherited stock paths, which add persistence on top of averaging, reach 0.87–0.88 for the embedding proxy and 0.74–0.79 for the exact-lexical one (Table 8). The decade result is the one that matters. It shows the strong stock correlation is not manufactured by the stock recursion, because the two methods already converge at 0.83 at decade frequency, before any smoothing, wherever the data are dense enough to measure culture at all—the post-1000 window in which the paper's claims sit.

Online Appendix Section OA.D reports the category, century, and direction breakdowns. Table 9 works through the disagreement cases. Machiavelli's *The Prince* endorses hierarchy about as openly as a book can; the staged reading scores its PD at 100, while the dictionary, counting words rather than weighing stance, puts it lower at 66.7. Venture Smith's narrative is the mirror image. Its pages are dense with the vocabulary of command and ownership because they describe slavery, so the dictionary assigns PD 100.0 to a text whose normative

Table 8: Proxy–LLM agreement by averaging frequency

			Exact lexical		Embedding-expanded	
Average	Restriction	Periods	Total ρ	Innovation ρ	Total ρ	Innovation ρ
Annual	All	600	0.278	0.291	0.397	0.458
Annual	≥ 10 books	163	0.455	0.490	0.476	0.489
Decade	All	149	0.350	0.314	0.554	0.510
Decade	≥ 10 books	50	0.465	0.642	0.828	0.835
Century	All	32	0.508	0.550	0.527	0.559
Century	≥ 10 books	16	0.526	0.524	0.794	0.688

Notes. Spearman rank correlations between the proxy and LLM wedge series, computed at each averaging frequency. "≥10 books" restricts to periods with at least ten scored books, so that thinly covered early periods do not dominate. Book-level correlations are 0.13 and 0.16 (exact lexical); inherited stock paths reach 0.74–0.79 (exact lexical) and 0.87–0.88 (embedding-expanded). LLM denotes large language model, and Innovation refers to the Innovation Wedge.

force runs entirely against hierarchy; the staged reading scores it 0. Swift's *A Modest Proposal* fails the count for a different reason—irony. Read word by word it brims with authority, calculation, and control, and the dictionary scores it accordingly; the staged reading, knowing what kind of document it is, does not take the proposal at its word. One instrument counts what a text mentions. The other judges what it defends.

Table 9: Worked cases: surface counts versus staged reading.

	The Prince		Venture Smith		Swift	
	Staged LLM	Exact lexical	Staged LLM	Exact lexical	Staged LLM	Exact lexical
PD	100.0	66.7	0.0	100.0	0.0	100.0
UA	66.7	100.0	50.0	50.0	0.0	100.0
IW	0.83	1.33	0.25	1.50	0.00	2.00

Notes. Selected book-level scores on the 0–100 dimension scale and the implied Innovation Wedge, for the staged LLM reading and the exact-lexical proxy. LLM denotes large language model, PD Power Distance, UA Uncertainty Avoidance, and IW the Innovation Wedge.

Three qualifications apply. First, the staged reading remains the paper's measure. Experts validate its book-level scores (Appendix B.1), the external battery checks the country stocks built from it (Section 4), and on the worked cases above it is the reading, not the proxies, that recovers the authorial stance. The cheap proxies are judged against it, not the other way around. Second, two smooth accumulating series can move together simply because corpus composition drives both, so the stock-level agreement corroborates the broad movement without proving it independently. The decade-level agreement is the evidence that matters; it is computed before any accumulation. Third, the conclusion is about the measure, not the method. The broad Total Friction and Innovation-Wedge declines are properties of the

text that even a scorer with no priors recovers. What the staged reading adds is trust in individual books.

A.5 Silence and Stance

Table 10 recomputes the benchmark accounting under the alternative treatments of silent chunks—silence as neutral, as in the baseline, and silence as missing—with the panel elasticity held fixed. The accounting is essentially unchanged across treatments. Between the eleventh and nineteenth centuries, the dimension-balanced directional-position share moves from 99.4 to 97.8 percent. All 6 dimension-level changes are declines, with an average absolute change of 1.6 percentage points, compared with 0.511 for conditional stance. Online Appendix Section OA.D records the estimand and explains why the zero category pools balance with silence.

Table 10: Fixed-share and productivity-growth accounting under alternative treatments of silent chunks

Convention	Fixed share	Contribution to 1920 productivity growth (percentage points)
Baseline (silence scored 50)	0.563	0.386
Silence as missing	0.543	0.402
Salience weighted	0.563	0.386

Notes. The table recomputes the benchmark accounting under the stated treatment of silent chunks with the panel elasticity fixed. Fixed share is counterfactual 1920 productivity divided by realized 1920 productivity.

A.6 Progress Vocabulary Bridge

We compare our measure with text evidence already used in economics. Almelhem et al. (2026) track the rise of progress-oriented language in early modern books. We build a simple version of that object on our own corpus—a count of 54 whole-word progress terms in the scored passages of each book, drawn from fourteen families—advance, betterment, civilisation/civilization, development, discovery, enlightenment, experiment, improvement, industry, invention, machinery and manufacture, progress, science, and technology—counted case-insensitively as whole words. For each book, we count how often these terms appear per 10,000 words of scored text. The rate puts long and short books on the same footing and accumulates into a stock exactly as the cultural scores do. This is a proxy built from our excerpts, not a replication of the published dictionary.

Table 11 reports how the progress vocabulary moves with the wedge. The vocabulary measures attention to progress rather than friction, so its correlation with the wedge is negative, most strongly at the stock level; the corresponding agreement for the dictionary wedges, where the expected sign is positive, is in Table 8. As the measured friction retreats, progress talk rises—an independent, mechanical reading of the same shelves pointing the same way. The timing behind the introduction's claim is in the lead-lag grid. Agreement is strongest when the vocabulary is shifted several decades ahead of the wedge (Online Appendix Table OA.12), so progress talk does move first, and the inherited stock turns later. The full construction and lead-lag grid are in Online Appendix Section OA.D; the per-century top-inflow lists and run records are in the replication package. The accounting rests on the inherited stock alone.

Table 11: Progress vocabulary and the Innovation Wedge

Progress-vocabulary and staged Innovation-Wedge correlations in supported decades

Series construction	Periods	Minimum books	Spearman ρ
decade flow average	31	10	-0.281
inherited decade stock	31	10	-0.698

Notes. Supported-decade Spearman correlations. The rows compare the progress-vocabulary attention measure with the staged wedge, so the expected sign is negative; the dictionary-wedge agreement, where the expected sign is positive, is in Table 8. The complete lead-lag grid is in Online Appendix Table OA.12; the per-century detail and run records remain in the replication package. Staged reading is the paper's book-scoring procedure with context, literary, behavioral, and synthesis stages. The Innovation Wedge is the inherited-stock ratio $(PD + UA)/(INDIV + LTO)$, where PD denotes Power Distance, UA Uncertainty Avoidance, INDIV Individualism, and LTO Long-Term Orientation.

B Validation Details

This appendix collects the validation record, moving from the book level to the country level to history. Appendix B.1 reports the expert-recognition test of individual book scores. Appendix B.2 records the survey validation behind Table 1, including the checks that the agreement survives changes to the accumulation rule, lands on the intended dimension rather than its siblings, and cannot be reproduced by income. Appendix B.3 tests the measure against dated historical events, and Appendix B.4 against the independent second archive of Section 4.4. Appendix B.5 closes by asking whether the historical decline and the survey agreement survive alternative wedge formulas.

B.1 Expert Recognition of Book-Level Scores

The expert-recognition test summarized in Section 4.1 checks whether the large book-level differences the pipeline assigns are visible to expert readers comparing the two works directly. The test is deliberately high-contrast and built from recognizable books.

Each expert received a short blinded packet of pairwise comparisons. Each pair listed title, author, approximate date, a short neutral orientation aid, and a plain-English description of the trait being compared. The packet showed no model score, no predicted direction, no internal dimension label, and no source labels, file paths, or score gaps. Experts answered on a five-point scale running from "A clearly more" through "about equal" to "B clearly more," and could always answer "not familiar enough / cannot judge"—nobody who did not know both books was forced to guess.

Every packet contained the same six anchor comparisons, one per dimension, plus four supplement comparisons that rotated across packets. The anchors carry the main evidence, since every respondent saw them, and A/B order was counterbalanced so the score-predicted higher book appeared on each side across the deployment. The supplements buy breadth across more book contrasts at the cost of fewer repeated judgments per pair, so we read them as supporting evidence rather than as an equally powered test.

Responses are coded as an ordered A-minus-B score, from $+2$ for "A clearly more" through 0 for "about equal" to -2 for "B clearly more," and multiplied by the hidden displayed-side direction, so a positive value always means agreement with the model-implied contrast. We report two rates. The preferred rate conditions on responses where the expert knew both works and gave a directional answer. The conservative rate keeps every response in the denominator, so cannot-judge and about-equal answers count against the model.

Of 30 packets sent, 14 came back, a response rate of 46.7 percent, carrying 140 item-level responses—84 on the common anchors and 56 on supplements. Table 12 reports the results. Among responses where the expert knew both books and gave a direction, 93.0 percent agree with the score-implied direction, and every anchor pair has majority agreement. Counting every response against the model, including cannot-judge and about-equal, still leaves the anchor confirmation rate at 65.5 percent. The supplements draw fewer usable answers but point the same way.

Table 13 breaks the same responses out by dimension. Agreement among familiar-both directional responses is at least 84.2 percent everywhere and unanimous for Individualism, Indulgence, and Uncertainty Avoidance. Power Distance is the hard one to field. Experts

Table 12: Expert recognition of book-level score contrasts

Role	Responses	Consistent	Opposite	Equal	Cannot judge
Anchor	84	55	4	6	19
Supplement	56	28	1	6	21

Notes. Responses are expert judgments of blinded book pairs. Anchor pairs appear in every packet, while supplement pairs rotate across packets. Consistent and Opposite denote agreement or disagreement with the score-implied ordering; Equal and Cannot judge retain nondirectional responses.

declined to judge 50.0 percent of its comparisons, which drags its conservative confirmation rate down to 35.0 percent; but when experts did know both books and gave a direction, even Power Distance points the model's way. That echoes the external validation, where Power Distance is also the weakest dimension—though here the weakness is that experts often could not judge it, not that they disagreed.

Table 13: Human validation by scored dimension

Dimension	Responses	Consistent	Opposite	Equal	Cannot judge
INDIV	25	14	0	5	6
INDUL	21	14	0	0	7
LTO	28	17	1	1	9
MAS	25	16	3	2	4
PD	20	7	1	2	10
UA	21	15	0	2	4

Notes. Responses are expert judgments of blinded book pairs, grouped by scored dimension. Consistent and Opposite denote agreement or disagreement with the score-implied ordering; Equal and Cannot judge retain nondirectional responses. INDIV denotes Individualism, INDUL Indulgence, LTO Long-Term Orientation, MAS Masculinity, PD Power Distance, and UA Uncertainty Avoidance.

Two cautions apply to these rates. Responses cluster by expert and by item, and the pairs were chosen to be recognizable and high contrast, so the sign counts are descriptive rather than a formal test. Still, experts were never told which way the model leaned, always had the cannot-judge escape, and chose the model-implied direction in nearly all usable anchor responses.

The test cannot speak to passage-level accuracy or to the precision of low-contrast scores. What it shows is that the large book-level contrasts entering the stock are visible to human experts. When the model says one recognizable work gives more weight to hierarchy, autonomy, achievement, uncertainty control, future orientation, or gratification than another, expert readers usually see the same contrast.

B.2 The Survey Validation

This subsection records the construction and results behind Table 1, in enough detail that a reader can see exactly what is being validated without turning the main text into a catalogue of survey items. The central object never changes. Country-year book flows from the production corpus are accumulated into inherited country stocks, and those stocks are compared to external country rankings under fixed directional mappings. The closing checks test the agreement's stability across choices and its specificity to the intended margins.

B.2.1 Benchmark Sources and Sign Map

The validation compares the text stock to external country measures that were not used to construct it. The sources are official Hofstede scores (Hofstede, 2001; Hofstede et al., 2010), country aggregates from the World Values Survey, European Values Study, Integrated Values Survey, International Social Survey Programme, European Social Survey, the Quality of Government Standard Dataset, GLOBE societal practices, and Schwartz cultural values. GLOBE values are not used in the headline validation because they measure aspirations; the practice scores are closer to realized social organization.

Each external variable is mapped to a text dimension and sign before computing any correlation. The map follows the economic interpretation of the six dimensions. Authority, deference, hierarchy, and power-distance measures map positively to PD. Autonomy, generalized trust, postmaterialism, individual self-direction, and reversed collectivism or embeddedness map positively to INDIV. Assertiveness, mastery, and hierarchical gender-role attitudes map to MAS. Expert-rule and certainty-oriented uncertainty-avoidance items map to UA. Work duty, future orientation, saving, thrift, and perseverance map to LTO. Happiness, life satisfaction, and affective autonomy map to INDUL. Composite validation rows use the same component logic. The Innovation Wedge is $(PD + UA)/(INDIV + LTO)$, while Total Friction adds MAS and INDUL to the broader six-dimensional ratio.

The sign convention is as follows. If an external proxy is predicted to move negatively with the text dimension, the proxy is multiplied by -1 before computing the correlation. All reported Spearman correlations are therefore signed so that positive means agreement with the fixed economic direction.

Online Appendix Section OA.E.1.2 reports the registry-level audit and the complete fixed sign map.

B.2.2 Text Input, Timing, and Samples

The text input is the scored output of the alpha wave and waves 000 through 025, with each wave entering through its final scored summary. The book-level fields used for the baseline are the production analysis year, analysis country or location, and the six scored cultural dimensions—Power Distance (PD), Individualism (INDIV), Masculinity (MAS), Uncertainty Avoidance (UA), Long-Term Orientation (LTO), and Indulgence (INDUL).

Country stocks are built exactly as in Section 3.3—country-year average scores, the same log-damped weight and depreciation—accumulated from the earliest valid year and read at 1920; K_{dct} is held at its 1920 value throughout the validation.

The headline validation sample is Europe and North America. This is not a convenience restriction. It matches the object being measured—Western elite literary culture. In the production book file, 22,493 of 23,064 complete scored rows, or 97.5 percent, are assigned to Europe or North America under the validation definition. Countries outside that region enter the all-overlap sample only as an out-of-scope comparison. In those countries, the country label often reflects where a Western author was located, not a representative national literary stock of the local population.

B.2.3 Family and Dimension Summaries

Panel A of Table 14 reports the validation by benchmark family. Each family is a set of external measures mapped to text dimensions under the fixed sign map of Appendix B.2.1. A test is one such measure, and the family's row reports how many tests it contributes, the mean and median signed Spearman correlation across them, the share carrying the predicted sign, and the share in which the intended text dimension ranks among the top three of the six candidates. Sources differ in how directly they measure national culture, and the results reflect that focus. The country-level batteries built for the purpose—the WVS/QoG extended battery, Schwartz values, GLOBE practices—agree with the stock on average, with WVS/QoG the strongest. Item-level microdata pooled across heterogeneous questions are noisier, and ISSP is close to zero in the pooled summary. Construction of each family and its fixed sign map are in Appendix B.2.1; Online Appendix Section OA.E.1.3 reports the family and dimension summaries. The signal also tracks how well the corpus represents the population being ranked—the mean correlation rises from 0.026 across all overlapping countries to 0.058 in Europe and North America and 0.145 in the Western core.

Table 14: External validation by benchmark family and text dimension

Panel A. Benchmark families

Benchmark family	Tests	Mean ρ	Median ρ	Positive signs (%)	Top-three specificity (%)
Official Hofstede	6	0.014	0.003	50.0	–
WVS/QoG compact	11	0.118	0.151	81.8	54.5
WVS/QoG extended	21	0.200	0.164	81.0	61.9
WVS online	17	0.004	-0.021	47.1	41.2
Raw WVS/EVS/IVS	18	-0.028	-0.035	38.9	27.8
Raw WVS/EVS factors	18	0.043	0.067	72.2	22.2
ISSP	14	-0.053	0.019	57.1	28.6
ISSP supplemental	12	0.063	0.096	58.3	33.3
ESS	4	0.094	0.138	100.0	25.0
GLOBE practices	7	0.136	0.211	71.4	57.1
GLOBE values	6	-0.006	0.048	50.0	33.3
Schwartz	10	0.110	0.151	90.0	50.0

Panel B. Text dimensions and wedges

Text variable	Tests	Mean ρ	Median ρ	Predicted sign (%)	Specificity rank percentile
PD	19	-0.059	-0.080	31.6%	0.333
INDIV	29	0.095	0.109	82.8%	0.683
MAS	12	0.056	0.050	75.0%	0.657
UA	10	0.107	0.180	60.0%	0.691
LTO	22	0.209	0.096	81.8%	0.730
INDUL	6	0.037	0.050	66.7%	0.645
IW	5	0.111	0.075	80.0%	0.845
TF	39	-0.021	0.008	51.3%	0.436

Notes. Rows use the 1920 text-stock ranking matched to later-dated benchmarks in Europe and North America. Specificity is the intended variable's rank percentile among candidate text variables after applying the predicted sign. The official Hofstede family has no specificity entry because its comparison is dimension-to-dimension. Wedge rows have fewer direct tests and are aggregate diagnostics. WVS denotes World Values Survey, QoG Quality of Government, EVS European Values Study, IVS Integrated Values Survey, ISSP International Social Survey Programme, and ESS European Social Survey. Panel A reports sign and top-three shares in percent, while Panel B reports the intended variable's mean rank percentile as a proportion.

Panel B of Table 14 gives the same evidence by text dimension. This panel explains why the main text emphasizes the Innovation Wedge and not Total Friction. INDIV, LTO, UA, MAS, and INDUL are positive on average in the pooled Europe/North America battery; PD is wrong-signed in this cross-section. The Innovation Wedge remains positive, while Total Friction is essentially zero. The margins Total Friction adds validate individually, so its failure is a property of the broader composite rather than of any single component; Appendix B.5 returns to this point.

B.2.4 Same-Name Hofstede Comparison

The direct same-name comparison—each text dimension against the corresponding official Hofstede dimension (Hofstede, 2001; Hofstede et al., 2010)—is the single-instrument test the main text sets aside, and it is weak. The comparison has 6 tests, mean $\rho = 0.014$, and predicted signs in 50.0 percent. With only 6 dimension-level tests it is badly underpowered, and it is dragged down by Power Distance and Long-Term Orientation.[28]

B.2.5 Additional Cross-Country Checks

Three further checks ask whether the validation means what Section 4 claims—that the stocks carry country-level cultural signal, specific to the intended dimensions, within the corpus's scope.

First, the agreement does not depend on the accumulation rule. Table 21 consolidates the stock-law sensitivity, and Online Appendix Section OA.E.1.5 reports the maintained grid summary. The pattern survives every depreciation and count-dampening choice.

Second, when a benchmark is built to measure a particular margin, the matching text variable should be among its best predictors, not merely positively correlated. Ranking all candidate text variables against each benchmark, the intended variable sits in the top half on average for the country-level families—WVS/QoG extended and compact, Schwartz, GLOBE practices, ESS—while raw respondent-level item batteries are weaker. At the wedge level the contrast is sharp. The Innovation Wedge ranks at the 0.845 percentile (permutation $p = 0.008$); Total Friction is indistinguishable from chance (0.436, $p = 0.900$).

Third, the income and all-overlap rows of Table 1 guard against a mechanical reading. Income correlates with almost everything, yet it does not reproduce the fixed signed cultural map. The worldwide comparison is flat, consistent with the corpus not representing the populations being ranked. It asks a Western corpus to rank populations that largely did not author it, and on non-Western countries alone the agreement disappears—mean $\rho = -0.076$, 34.1 percent of tests in the predicted direction, 2 of 11 families positive. The validation is positive within the corpus's geographic scope and disappears outside it.

[28]The 6 same-name correlations are PD −0.097, INDIV 0.130, MAS 0.050, UA −0.001, LTO 0.003, and INDUL −0.003. Individualism carries the row. The two weak dimensions differ. Long-Term Orientation validates across the broader benchmark battery but not on Hofstede's own LTO scale, so the same-name figure understates it; Power Distance is weak almost everywhere.

B.3 Historical Event Benchmarks

The test uses dated historical episodes, with dates drawn from standard historical sources and component signs assigned from the historical narratives; each episode names a specific component and a specific treated region. On the primary panel, treated regions move in the predicted direction in 11 of 15 component comparisons (Table 16); the strict difference-in-differences count is weaker because control regions often drift in the same direction, consistent with broad diffusion of cultural change rather than event-localized breaks.

The English Civil War settlement predicts lower hierarchy and uncertainty avoidance in the United Kingdom, and both treated-region movements have the predicted sign. The scientific-society wave predicts higher individualism and long-term orientation and lower uncertainty avoidance across the United Kingdom, France, the Netherlands, Italy, and Germany; individualism and uncertainty avoidance move as predicted, while long-term orientation does not. The Enlightenment public-sphere expansion predicts higher individualism and long-term orientation and lower hierarchy in the same five-country group, and all three treated-region movements have the predicted sign. In the Napoleonic-code regions, none of the three treated-region movements has the predicted sign, although the relative changes have the predicted signs for PD and INDIV. This is the only primary event with no treated-sign hit, and its treatment is the hardest to map onto the panel's country units because French legal reforms were imposed territory by territory within countries represented as single units (Acemoglu et al., 2011). The German reform and education wave moves long-term orientation and uncertainty avoidance upward as predicted, and Atlantic abolition moves hierarchy down and individualism up in the United Kingdom and United States. The primary event-level estimates are reported in Table 15; the feasibility screen and secondary events remain in Online Appendix Sections OA.E.2.1 and OA.E.2.2.

B.4 The Independent Archive

The independent archive draws from four collections outside Gutenberg. EEBO-TCP, the Early English Books Online–Text Creation Partnership, contains machine-readable transcriptions of English books printed from 1473 to 1700; Evans covers early American imprints through 1800; ECCO, Eighteenth Century Collections Online, covers eighteenth-century printed works; and the Internet Archive supplies the nineteenth-century block. EEBO-TCP and Evans form the 1500–1699 block, ECCO the 1700–1799 block, and the Internet Archive the

Table 15: Treated-minus-control component movements in the primary historical-event panel

Event-component estimates

Event	Component	DiD	Treated-direction hit
English Civil War to Glorious Revolution settlement	PD	15.67	Yes
English Civil War to Glorious Revolution settlement	UA	13.35	Yes
Scientific-society founding wave	LTO	-0.25	No
Scientific-society founding wave	INDIV	4.86	Yes
Scientific-society founding wave	UA	-4.63	Yes
Enlightenment public-sphere expansion	INDIV	-3.96	Yes
Enlightenment public-sphere expansion	LTO	-7.07	Yes
Enlightenment public-sphere expansion	PD	-0.67	Yes
Napoleonic-code regions	PD	-0.92	No
Napoleonic-code regions	UA	1.45	No
Napoleonic-code regions	INDIV	1.33	No
Prussian/German reform and education wave	LTO	0.57	Yes
Prussian/German reform and education wave	UA	11.46	Yes
Atlantic abolition and anti-slavery mobilization	PD	-0.61	Yes
Atlantic abolition and anti-slavery mobilization	INDIV	3.51	Yes

Notes. Primary events satisfy the fixed treated-side support rule. DiD is the treated-minus-control stock change; the final column records the companion treated-region sign check. PD denotes Power Distance, UA Uncertainty Avoidance, INDIV Individualism, and LTO Long-Term Orientation. Dating sources are English settlement (North and Weingast, 1989); scientific societies (Mokyr, 2016; David, 2008); Enlightenment (Mokyr, 2009; Squicciarini and Voigtländer, 2015); Napoleonic code (Acemoglu et al., 2011); Prussian reforms (Becker et al., 2011; Clark, 2006); Atlantic abolition (Drescher, 2009; Brown, 2006).

Notes. PD denotes Power Distance, UA Uncertainty Avoidance, INDIV Individualism, and LTO Long-Term Orientation.

Table 16: Historical event benchmark sign summary

Panel	Events	Tests	Difference-in-differences			Treated region		
			Hits	Share	p	Hits	Share	p
Primary events	6	15	9	0.600	0.304	11	0.733	0.059
Thin-side secondary events	5	13	10	0.769	0.046	9	0.692	0.133
All signed event-components	11	28	19	0.679	0.044	20	0.714	0.018

Notes. Difference-in-differences hits count event-component rows whose treated-minus-control change has the fixed sign; treated hits count whether the treated region alone moved in the fixed direction. The binomial probabilities are descriptive because event-component rows are not independent. In 4 of the 6 strict primary DiD misses, treated regions moved in the predicted direction but controls moved further in the same direction.

1800–1899 block. Sampling is balanced across half-centuries, and exact or close title–author matches to the main corpus are removed.

The archive uses publication or edition dates rather than inferred composition dates. It therefore changes both the archive and the dating convention, while retaining the scoring prompts, passage-selection rule, and stock recursion. The design does not weight books by readership or surviving copies. The four collections form three chronological blocks, with EEBO-TCP and Evans both contributing to the earliest block, so changes at collection joins

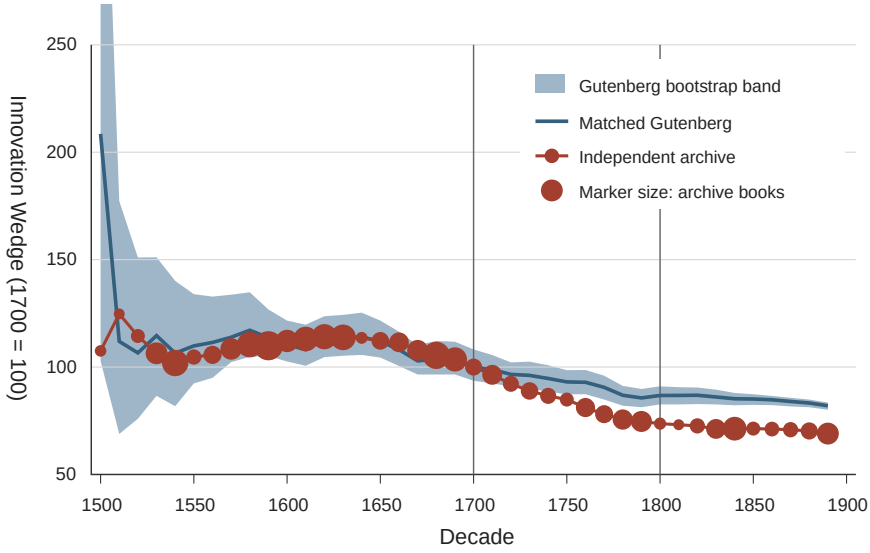


Figure 8: Independent-archive and matched Gutenberg Innovation-Wedge paths.

Notes: Both paths share the 1500 cold start and are indexed separately to 100 in 1700. From 1550 to 1890, the archive Innovation Wedge changes by -0.743 and the matched-Gutenberg wedge by -0.356; this comparison omits the first cold-start half-century. For reference, the archive change from 1500 to 1890 is -0.802. The shaded band is the matched-Gutenberg bootstrap band expressed using the same Gutenberg 1700 anchor. The vertical markers show collection joins at 1700, when Early English Books Online–Text Creation Partnership and Evans give way to Eighteenth Century Collections Online, and at 1800, when Eighteenth Century Collections Online gives way to the Internet Archive.

combine source and time differences. The within-collection paths and the initialization checks separate what can be learned without relying on those joins.

The archive path and frozen criteria verdicts provide the evidence summarized in Section 4.4. The reconstructed path must decline, must track the corpus path in levels, and must pass the frozen magnitude screens; Figure 8 and Table 17 report those outcomes. We also report the decade first-difference correlation, an addition beyond the frozen criteria; it comes to 0.349. The design summary, sampling screen, decade correlations, resampling comparison, and within-archive trends are reported in Online Appendix Sections OA.E.3.1–OA.E.3.5; the complete design document and the genre feasibility check remain in the replication package.

B.5 Wedge Robustness

This subsection tests the measured decline and the growth-accounting conclusions across wedge formulas. The variant family is common to both parts—additive indices, reduced two-component ratios, and signed principal components—so they read as one argument. Both the measure and its model-based interpretation are stable across the formulas. The

Table 17: Independent-archive criteria verdicts

Construction	Criterion	Verdict	Statistic
Primary, matched cold start	Endpoint decline (sign)	pass	-0.802
Primary, matched cold start	Decade Spearman (sign)	pass	0.902
Primary, matched cold start	Inside Gutenberg bootstrap band	pass	-0.802
Primary, matched cold start	Genre-reweighted (secondary)	indeterminate	–
Seeded continuation	Endpoint decline (sign)	pass	-0.555
Seeded continuation	Decade Spearman (sign)	pass	0.805
Seeded continuation	Inside Gutenberg bootstrap band	indeterminate	–
Seeded continuation	Genre-reweighted (secondary)	indeterminate	–
Decade flows	Endpoint decline (sign)	pass	-1.076
Decade flows	Decade Spearman (sign)	pass	0.660
Decade flows	Inside Gutenberg bootstrap band	indeterminate	–
Decade flows	Genre-reweighted (secondary)	indeterminate	–

Notes. Endpoint-decline rows report the early-to-late endpoint comparison; all three constructions compare 1500 to 1890. Decade-Spearman rows use the design-rule 10-book threshold, giving 40 eligible decades. The Gutenberg bootstrap-band criterion is evaluated for the primary matched endpoint only; its band is -4.310 to -0.265. For seeded continuation and decade flows the bootstrap-band criterion is indeterminate by design. Genre-reweighted is indeterminate because the independent archives provide archive-specific subject fields but the Gutenberg comparison panel lacks a comparable genre or Library of Congress subject field.

accounting part adds pure units checks—log, rank, and score-point transforms—which cannot alter the measure's ordering by construction and therefore appear only there. Total Friction appears only in the measure part, because the accounting rests on the Innovation Wedge alone.

B.5.1 The Measure Under Alternative Formulas

Table 18 reports whether the 1000–1920 decline survives alternative formulas and whether those constructions retain agreement with modern surveys.

Panel A of Table 18 rebuilds the historical paths under alternative formulas. For Total Friction the variants are an additive signed index, ratios dropping MAS or INDUL, the four-margin ratio, and a signed first principal component; for the Innovation Wedge they are the additive index, the two-component ratios PD/INDIV, UA/INDIV, and UA/LTO, and the signed first principal component. For each variant the panel reports the 1000 and 1920 values in the variant's own units, the decline in percent where the scale has a meaningful zero, the decline in annual-path standard deviations, and the correlation of the variant's path with the baseline path. Every variant declines, for both wedges, and every path tracks its baseline closely.

Panel B of Table 18 repeats the external validation of Section 4 for the alternative constructions; each row reports the number of external tests, the mean signed correlation, and the share of tests carrying the predicted sign. The Innovation Wedge carries the predicted sign under the baseline ratio, the additive index, the signed first principal component, and the two simple component ratios. Total Friction does not. The external data therefore discriminate among aggregates rather than blessing any combination of the components, and the narrower wedge is the one they support.

The weak component is Power Distance, as Section 4 reports. The growth accounting need not rely on it. The main text therefore treats reduced-wedge paths as robustness rather than replacement benchmarks, while the construction-specific panel estimates and matched accounting appear below. The complete model-link comparison is in Online Appendix Section OA.E.1.5. The cross-section is also a harsh test for PD. Among roughly twenty broadly similar Western countries in 1920 there is little PD variance left to detect, while the movement the paper actually uses—the long decline from medieval hierarchy toward modern egalitarianism over nine centuries—is within-Western and much larger.

B.5.2 The Accounting Under Alternative Formulas

Table 19 reports two accounting exercises that separate path robustness from numerical units. Panel A of Table 19 holds the benchmark panel elasticity fixed while replacing the corpus-wide wedge path, so its magnitudes are common-elasticity stress tests conditional on each construction's units. Panel B of Table 19 re-estimates the same country-panel specification for each construction and applies that elasticity to the matching corpus-wide path. Re-estimation makes the accounting invariant to a pure affine rescaling.

Across the full four-component alternatives fixed in advance of estimation, construction-specific accounting puts the model-implied 1920 contribution between 0.310 and 0.369 percentage points, compared with 0.387 under the benchmark ratio. The UA/LTO reduction gives 0.384 percentage points when the benchmark elasticity is held fixed and 0.364 after re-estimating its own elasticity. Its country-clustered p-value is 0.010, while the exact wild-score value is 0.093. The $PD/INDIV$ and $UA/INDIV$ reductions are likewise excluded from the full-component range.

Table 18: Wedge-form robustness and external validation

Panel A. Historical paths under alternative formulas

Variant	1000	1920	Decline (%)	Decline (SD)	Correlation with baseline path
Innovation Wedge					
Baseline ratio	2.051	1.013	50.6	2.6	1.000
Additive signed index	41.493	0.770	–	2.9	0.996
PD/INDIV reduction	2.103	1.093	48.0	1.9	0.929
UA/INDIV reduction	2.296	1.276	44.4	2.0	0.938
UA/LTO reduction	2.005	0.953	52.5	3.0	0.952
Signed first principal component	1.577	-3.837	–	2.9	0.992
Total Friction					
Baseline ratio	1.189	0.861	27.6	2.7	1.000
Additive signed index	10.018	-8.287	–	2.8	0.999
Ratio, drop MAS	2.393	1.300	45.7	2.7	0.959
Ratio, drop INDUL	1.019	0.671	34.2	2.7	0.991
PD+UA over INDIV+LTO	2.051	1.013	50.6	2.6	0.956
Signed first principal component	1.637	-4.690	–	2.9	0.891

Panel B. External validation in Europe and North America

Variant	Tests	Mean ρ	Median ρ	Share with predicted direction (%)
Innovation Wedge				
Baseline ratio	5	0.111	0.075	80.0%
Additive signed index	5	0.109	0.044	80.0%
PD/INDIV reduction	5	0.103	0.074	80.0%
UA/INDIV reduction	5	0.045	0.030	60.0%
UA/LTO reduction	5	0.110	0.086	100.0%
Signed first principal component	5	0.119	0.096	80.0%
Total Friction				
Baseline ratio	39	-0.021	0.008	51.3%
Additive signed index	39	-0.050	-0.012	43.6%
Ratio, drop MAS	39	0.021	0.046	61.5%
Ratio, drop INDUL	39	-0.036	0.007	51.3%
PD+UA over INDIV+LTO	39	-0.003	0.000	48.7%
Signed first principal component	39	0.023	0.056	59.0%

Notes. Higher values mean more friction. Panel A uses the aggregate annual log-damped stock series and orients signed principal components toward higher baseline friction. Decline (%) reports the percentage decline from 1000 to 1920 for ratio constructions and a structural dash for signed additive and principal-component constructions because their zero is arbitrary. Decline (SD) reports the endpoint fall divided by the population standard deviation of the annual path from 1000 through 1920 for every construction, using the path before display rounding. Panel B uses the 1920 text-stock ranking matched to later benchmarks. Each external test has a direction in the fixed sign map, and the displayed share is the fraction matching it. The table tests whether the historical decline and external signal survive alternative constructions. It does not select a new index. PD denotes Power Distance, UA Uncertainty Avoidance, INDIV Individualism, LTO Long-Term Orientation, MAS Masculinity, and INDUL Indulgence.

C Stock Construction, Support, and Historical Detail

Building on Section 3.3, this appendix first reports the support inside the inherited stock, then tests the stock law and corpus composition. It next dates the transition sequence and closes with a worked England stock.

Table 19: Alternative wedge constructions under fixed and re-estimated elasticities

Panel A. Common benchmark elasticity

Construction	Fixed share	1920 contribution (pp)
Baseline ratio	0.563	0.387
Additive signed index	0.646	0.321
PD/INDIV reduction	0.598	0.387
UA/INDIV reduction	0.623	0.398
UA/LTO reduction	0.540	0.384
Signed first principal component	0.474	0.958
Log ratio	0.673	0.281
Additive index (score points)	0.849	1.000
Rank-transformed scores	0.616	0.334

Panel B. Construction-specific panel re-estimation

Construction	$\hat{\psi}$	Cluster p	Exact wild p	1920 contribution (pp)
Baseline ratio	0.631	0.023	0.096	0.387
Additive signed index	0.645	0.011	0.026	0.327
PD/INDIV reduction	0.121	0.297	0.324	0.084
UA/INDIV reduction	0.189	0.156	0.161	0.133
UA/LTO reduction	0.593	0.010	0.093	0.364
Signed first principal component	0.091	0.013	0.021	0.310
Log ratio	0.782	0.008	0.035	0.341
Additive index (score points)	0.013	0.011	0.026	0.327
Rank-transformed scores	0.708	0.028	0.080	0.369

Notes. Higher values denote greater innovation friction. Panel A changes the corpus-wide path while holding the benchmark panel elasticity fixed. Panel B re-estimates the maintained country-panel specification for each construction and applies the resulting elasticity to its matching path. The panel sample and accounting inputs are otherwise unchanged. Clustered values use country-clustered inference, and exact wild values use the maintained wild-score procedure. The log ratio, additive signed index, signed principal component, and rank-transformed wedge form the full-component family fixed in advance of estimation. The two-component ratios are reduced-component diagnostics. The raw additive index and its displayed rescaling are affine equivalents and are both shown among the units checks. Here $\hat{\psi}$ is the construction-specific cultural elasticity, and pp means percentage points. PD denotes Power Distance, INDIV Individualism, UA Uncertainty Avoidance, and LTO Long-Term Orientation.

C.1 Inherited-Stock Support

The stock support is not the same object as the raw book count of Appendix A.1. Old inflows continue to enter the inherited stock after their publication year, while new inflows determine how much the stock can move in the current year. Table 20 reports this distinction at four benchmark dates, for the full corpus and, beside it, for the grid-matched sample used in the spatial analysis. The effective-support columns convert the surviving discounted weights into an equivalent number of equally weighted books; when a few high-weight books dominate the stock, effective support is low. Because old books persist in the stock, support at any date

exceeds the raw count of new books, and the paired columns show that the same pattern holds in the thinner grid sample.

Table 20: Support in the corpus-wide and grid-matched inherited stocks

Stock mass and current inflow

Year	Corpus stock mass	Grid stock mass	Corpus current inflow share	Grid current inflow share
1000	7.8	4.5	20.6%	15.3%
1500	27.8	21.0	2.5%	3.3%
1700	144.3	133.8	1.3%	1.5%
1920	549.8	478.2	1.2%	1.1%

Active-year mass and effective support

Year	Corpus active-year mass	Grid active-year mass	Corpus effective support	Grid effective support
1000	7.7	5.6	9.9	9.6
1500	31.2	25.4	51.0	40.2
1700	113.8	108.9	136.7	129.8
1920	171.0	169.4	162.6	170.4

Notes. The corpus columns use the full scored-book sample, and the grid columns use the grid-matched sample. Mass columns use the recursion's weighted-volume units; effective support is the equivalent number of equally weighted inflows. Inflow weight is $\log(1 + N_t)$. Stock mass is the surviving volume stock V_t. Current inflow share is the current inflow weight divided by V_t. Active-year mass discounts the history of years with positive effective inflow at the baseline depreciation rate. Effective support is V_t^2 divided by the discounted sum of squared inflow weights.

C.2 Stock-Law Sensitivity

The baseline recursion of Section 3.3 weights each year's inflow by $\log(1 + N_t)$. The volume and momentum accumulators initialize at zero at the earliest dated year in the sample, 700 BCE, so the first dated work supplies the first inflow, and the first defined stock value is that work's score. The checks here change one piece at a time and leave everything else in place. The raw-count rule replaces the damped weight with $v_t = N_t$, one unit of volume per dated book; an empty year still contributes no inflow, and because volume and momentum depreciate at the same rate, the inherited composition is unchanged in such a year.

Table 21 reports the checks in three panels. Panel A of Table 21 recomputes the 1000–1920 endpoint change of both wedges across the maintained depreciation rates and, at baseline depreciation, across the five inflow weights. The decline survives every rule, and its size is non-monotone in δ_K because depreciation reweights the entire book history rather than decaying a fixed initial stock. Panel B of Table 21 holds depreciation at the baseline and reruns the primary Europe/North America battery of Table 1 under five inflow weights, from the raw count to heavily damped logs. Every rule preserves the agreement, which strengthens mildly as damping increases, and the preferred $\log(1 + N_t)$ sits in the middle

of that range. Panel C of Table 21 re-estimates the primary urbanization specification of Section 7 with the cultural regressor built three ways—annual flows, cumulative averages, and inherited stocks. Annual flows do not form a stable, directionally coherent pattern across the headline components on the much thinner sample in the displayed Bartlett–Conley coefficient-to-standard-error comparisons; cumulative averages agree with the inherited stock on every sign, and the inherited stock remains preferred on interpretation because culture is a persistent state.

Online Appendix Section OA.F reports the depreciation grid and period support; the complete record—signed correlations, specificity checks, stock-year and weighting variants, support-threshold checks, and benchmark-family summaries—is in the replication package's stock-robustness extension.[29] A separate negative-control extension shows that the Pantheon calibration collapses under exposure shuffles, while the main validation and prediction links weaken under score, origin, and exposure shuffles. None of this substitutes the text stocks for modern surveys. It shows the main patterns do not come from one arbitrary stock convention or from generic persistence.

C.3 Corpus Composition

Table 22 reports three checks on corpus composition, and all three leave the decline in place. Panel A recomputes the wedges holding the genre or subject mix fixed at its full-sample shares, so the decline cannot come from a changing mix of genres or subjects; under both fixes the endpoint changes stay close to the observed-composition baseline. Panel B splits the corpus by provenance—strict originals, other original-like works, translations and adaptations, and mixed or mediated records—and rebuilds each inherited path. Levels differ across subsets, but every subset declines to 1920, so the fall is not carried by translations or editorial mediation. Panel C replaces the baseline one-source-work-one-weight rule with the Buringh and van Zanden (2009) edition-production weights on the covered 1000–1800 sample; the weighted decline is somewhat steeper than the baseline, so the maintained weighting is, if anything, conservative.

[29] As a check on the stock logic, we simulate data where the truth is known. Books arrive unevenly and carry noisy scores of a slowly moving latent state. The inherited stock recovers that state better than annual flows; the advantage disappears when the latent state is pure noise; and rising publication volume alone does not manufacture a trend.

Table 21: Stock-law sensitivity

Panel A. Global endpoint changes

		Endpoint change	
Variation	Rule	Innovation Wedge	Total Friction
Depreciation rate	$\delta_K = 0.0025$	-0.917	-0.272
Depreciation rate	$\delta_K = 0.0050$	-1.038	-0.329
Depreciation rate	$\delta_K = 0.0100$	-0.925	-0.333
Annual inflow weight	Raw count	-1.055	-0.341
Annual inflow weight	$\log(1 + 0.25n)$	-1.013	-0.323
Annual inflow weight	$\log(1 + n)$ (preferred)	-1.038	-0.329
Annual inflow weight	$\log(1 + 4n)$	-1.075	-0.337
Annual inflow weight	$\log(1 + 16n)$	-1.102	-0.343

Panel B. Count dampening in external validation

Annual inflow weight	Tests	Mean Spearman ρ	Predicted-sign share (%)
Raw count	138	0.046	60.1%
$\log(1 + 0.25n)$	138	0.049	63.0%
$\log(1 + n)$ (preferred)	138	0.061	65.2%
$\log(1 + 4n)$	138	0.063	65.9%
$\log(1 + 16n)$	138	0.066	66.7%

Panel C. Current-design urbanization gradients

Construction	PD	UA	INDIV	LTO	MAS	INDUL	TF	IW
Annual flows	-0.023	-0.012	-0.045	-0.089	-0.089	-0.034	0.485	2.380
	(0.034)	(0.036)	(0.039)	(0.039)	(0.024)	(0.032)	(1.324)	(0.692)
Cumulative averages	-0.458	-0.659	0.501	0.446	-0.392	0.488	-17.003	-6.090
	(0.075)	(0.094)	(0.111)	(0.126)	(0.121)	(0.088)	(3.039)	(1.029)
Inherited stock	-0.373	-0.535	0.388	0.054	-0.328	0.393	-12.100	-3.069
	(0.057)	(0.077)	(0.075)	(0.082)	(0.083)	(0.078)	(3.031)	(0.850)

Notes. Panel A of Table 21 reports corpus-wide Innovation-Wedge and Total-Friction endpoint changes under the displayed stock rules. Panel B of Table 21 reports the maintained external-validation battery under the displayed inflow weights; the maintained grid summary remains in Online Appendix Section OA.E.1.5. Panel C of Table 21 reports each coefficient on the estimate line and its Bartlett–Conley standard error in parentheses on the blank-stub line. The no-signal description for annual flows means that the headline components do not form a stable, directionally coherent pattern in the displayed coefficient-to-standard-error comparisons. It does not mean that every annual-flow coefficient is individually indistinguishable from zero. PD, UA, INDIV, LTO, MAS, INDUL, TF, and IW denote Power Distance, Uncertainty Avoidance, Individualism, Long-Term Orientation, Masculinity, Indulgence, Total Friction, and the Innovation Wedge.

Table 22: Corpus-composition checks

Panel A. Fixed corpus composition

	Innovation Wedge			Total Friction
Series	1000	1920	Change	Change
Observed composition	2.051	1.013	-1.038	-0.329
Fixed genre mix	1.929	0.996	-0.933	-0.293
Fixed subject mix	2.200	0.988	-1.212	-0.361

Panel B. Originals and translations

	Book counts		Innovation Wedge	
Subset	All years	Pre-1600	1000	1920
All books	23064	504	2.051	1.013
Strict originals	17433	255	3.437	1.027
Other original-like	2339	23	2.791	0.999
Translations/adaptations	1529	122	4.115	0.856
Mixed or mediated	1737	97	1.614	0.936

Panel C. Edition weighting on the covered sample

	Innovation Wedge			Total Friction
Path	1000	1800	Change	Change
One source work, one weight	2.051	1.108	-0.942	-0.296
Edition-production weights	2.442	1.121	-1.321	-0.348

Notes. Panel A fixes the full-sample genre or subject mix while preserving within-group annual scores. Panel B classifies source works by original, translation, adaptation, or mediated status and rebuilds each inherited path. Panel C compares the baseline one-source-work weighting with Buringh–van Zanden edition-production weights on the covered 1000–1800 sample. The Online Appendix reports endpoint summaries, provenance categories, and the edition-weighted accounting comparison; complete annual paths remain in the replication package.

C.4 Transition Timing

The dimensions do not all shift at once; some move centuries before others. Table 23 dates the sequence by the century holding each series' largest hundred-year change in the direction of its 1920 value.

Table 23: The post-1000 transition has an internal sequence.

Series	1000	1920	Largest 100-year move	Bootstrap start-year percentiles (5/50/95)	Direction
Long-Term Orientation	42.2	67.5	1074–1174	1067–1074–1510	up
Individualism	36.8	50.5	1200–1300	1010–1036–1176	up
Power Distance	77.4	55.2	1514–1614	1008–1044–1514	down
Uncertainty Avoidance	84.5	64.4	1130–1230	1024–1130–1428	down
Innovation Wedge	2.05	1.01	1130–1230	1013–1130–1157	down
Total Friction	1.19	0.86	1512–1612	1021–1130–1512	down

Notes: The table uses the annual log-damped inherited stock series from 1000 to 1920. For Power Distance, Uncertainty Avoidance, Total Friction, and the Innovation Wedge, the endpoint direction is downward. For Individualism and Long-Term Orientation, the endpoint direction is upward. The largest 100-year move is the window with the largest movement in the endpoint direction. The exercise is descriptive; it is not a formal lead-lag or causal time-series test. The bootstrap dates the largest-move windows to roughly a century and a half for the Innovation Wedge and individualism, and to four to five centuries for the slower-moving components. Half-completion years are not stable under resampling and are not treated as table evidence.

The constructive margins move early—Long-Term Orientation has its largest century move in 1074–1174, and Individualism in 1200–1300. The blocking margins split—Uncertainty Avoidance records its largest decline early, in 1130–1230, overlapping the constructive movement, while Power Distance falls latest, centered on 1514–1614. Half-completion years point the same way but are too noisy to pin down, so the timing rests on the window evidence.

C.5 A Worked Country Stock: England, 1500–1800

The England case shows how dated books become a country wedge for England from 1500 to 1800. It uses the production scores and the baseline recursion with $\delta_K = 0.005$. It is a measurement illustration; the separate England productivity comparison in Appendix D.3 tests the model's Western-aggregate path against measured English TFP and does not use this country stock.

The sample is defined by the place evidence in the records rather than by a fixed country list. The main case keeps scored books with explicit England place evidence in the country, region, or location label and excludes explicit Scotland and Ireland evidence, because the historical object is England rather than the post-1707 British state or the wider British Isles; a variant pools the England, Scotland, Ireland, Wales, United Kingdom, and Great Britain

labels. Table 24 reports both. The England-only stock draws on 660 scored books dated 1500–1800, and its Innovation Wedge falls from 2.920 in 1500 to 1.407 in 1700 and 1.155 in 1800—moving from a high wedge toward near parity—a log change of −0.928; the pooled variant gives −0.905, so the sample rule is not driving the movement.[30]

Table 24: England worked case: samples, wedge path, and top-book concentration

Sample and wedge path

Sample	All books	Books dated 1500–1800	τ_{1500}	τ_{1700}	τ_{1800}	$\Delta \log \tau$
England only	6893	660	2.920	1.407	1.155	-0.928
UK/British Isles pooled	9234	739	2.888	1.425	1.169	-0.905

Top-15 concentration

Sample	Top-15 share, 1500	Top-15 share, 1700	Top-15 share, 1800
England only	0.630	0.098	0.050
UK/British Isles pooled	0.568	0.090	0.046

Notes: The England-only sample keeps books with explicit England place evidence and excludes explicit Scotland and Ireland evidence; the pooled variant adds Scotland, Ireland, Wales, United Kingdom, and Great Britain labels. τ columns report the Innovation Wedge of the country stock at the indicated year under the baseline log-damped recursion with $\delta_K = 0.005$. The Top15 columns report the share of the surviving denominator stock accounted for by the fifteen books with the largest surviving discounted weight at that date.

The Top15 rows of Table 24 make the support structure visible—the fifteen highest-weight books carry 63 percent of the 1500 denominator of the wedge, the constructive stocks, falling to 10 percent by 1700 and 5 percent by 1800. The ranked title lists and British-Isles variants are reported in Online Appendix Section OA.F.

Table 25 reports each component's contribution to the fall. Each row freezes one dimension's stock at its 1500 value and recomputes the 1800 wedge, so the row's entry is the part of the fall that disappears. Patience does the most work (freezing Long-Term Orientation removes −0.290 of the −0.928 baseline movement), followed by Individualism, Power Distance, and Uncertainty Avoidance; the contributions do not sum mechanically because the wedge is a log of a ratio, and the interaction remainder is −0.071. The ordering matches the aggregate timing evidence of Section 5. The constructive margins carry the early movement.

[30]Values in this appendix are read from the generated tables below and from `england_summary.csv` in the replication package's bridge target, which recomputes the case from raw text to final score.

Table 25: England worked case: one-at-a-time component freezes, 1500–1800

Component	τ_{1500}	τ_{1800}	$\Delta \log \tau$	Freeze contribution
Baseline (all components move)	2.920	1.155	-0.928	-0.928
PD frozen at 1500 value	2.920	1.387	-0.744	-0.184
UA frozen at 1500 value	2.920	1.337	-0.781	-0.147
INDIV frozen at 1500 value	2.920	1.462	-0.692	-0.236
LTO frozen at 1500 value	2.920	1.543	-0.638	-0.290
Interaction remainder	2.920	–	–	-0.071

Notes. Each freeze row holds the indicated component's stock at its 1500 value while the other components follow their realized paths, and recomputes the 1800 Innovation Wedge. The freeze contribution is the amount of the baseline log movement removed by the freeze. Contributions do not sum to the baseline because the wedge is a log of a ratio; the interaction remainder row reports the gap. PD denotes Power Distance, UA Uncertainty Avoidance, INDIV Individualism, and LTO Long-Term Orientation.

D Growth Identification and Accounting

The growth appendix extends Section 6 by reporting the model solution, the panel estimation with its aggregate validation, an external comparison between the model paths and measured productivity in England and Great Britain, accounting robustness, and the split between research productivity and talent allocation.

D.1 Model Solution

The idea-production block is

$$\dot{A}_t = \delta_t H_{A,t}^{\lambda} A_t^{\phi}, \qquad \delta_t = \bar{\delta} \exp(-\psi \tau_t^I), \qquad 0 < \phi < 1. \tag{17}$$

Multiplying by $A_t^{-\phi}$ and integrating yields the transition solution in equation (12). Under the same long-run effort path, two permanent wedge values satisfy

$$\frac{A_t(\tau^L)}{A_t(\tau^H)} \longrightarrow \exp\left\{-\frac{\psi}{1-\phi}(\tau^L - \tau^H)\right\}. \tag{18}$$

The convergence rate is $(1-\phi)g_A^{\infty}$, so the half-life of the remaining log-level gap is $\log 2/[(1-\phi)g_A^{\infty}]$. The annual recursion evaluates the same transition from a common initial idea stock. The realized and fixed-wedge paths start from the same idea stock. Online Appendix Section OA.G.1 gives the full derivation, special cases, balanced-growth limits, and numerical implementation.

D.2 Panel Estimation and Aggregate Validation

This subsection explains the estimation choices behind the panel elasticity and then checks what the estimate implies in the aggregate.

Births are counts, and many country-windows contain none. A log regression would drop those zeros and select the sample on the outcome. PPML instead fits the conditional mean directly, choosing coefficients so that expected births match observed counts. The Poisson label refers only to the fitting criterion; the estimates are consistent whenever the conditional mean is right, whatever the distribution of the counts. Country fixed effects absorb level differences, period effects absorb common shocks, and standard errors cluster by country. The coefficient on the wedge is ψ, estimated on 118 observations across the 13 countries of Section 6.

The wedge also beats the two rival series—the broader Total-Friction composite and the progress vocabulary of Appendix A.6. Substituting either for the wedge in this estimation panel attracts no significant coefficient, and the wedge fits best (Online Appendix Table OA.39).

The aggregate record—the Western-wide series of innovator births—tests timing rather than identifying ψ. Table 26 shows why. Estimating ψ freely from the aggregate series alone leaves the objective nearly flat between 0.20 and 1.02, so the aggregate data cannot distinguish values inside that range; the panel estimate remains the benchmark, and the free fits are diagnostics.

Table 26: Aggregate timing fits under panel-imposed and freely fitted cultural elasticity

Sample	Fit	ψ	η	R^2	Log RMSE	Objective	N
1000–1899, overlapping	Panel ψ imposed	0.631	0.506	0.846	0.443	160.723	818
1000–1899, overlapping	Free objective fit	0.128	0.482	0.847	0.421	144.825	818
1000–1899, non-overlapping	Panel ψ imposed	0.631	0.542	0.863	0.441	6.613	34
1000–1899, non-overlapping	Free objective fit	0.200	0.519	0.861	0.426	6.159	34

Notes. Every row uses the same validation base—the corpus-wide all-books Innovation Wedge, Western accounting population, and Pantheon science and invention births over the same Western country set. The wedge is anchored at its year-1000 value, and the model scale is normalized so realized productivity growth equals 1 percent in 1920. Free rows minimize the log-SSR objective over ψ, the intercept, and the Pantheon proxy elasticity; fixed rows impose the country-panel ψ and re-estimate the two nuisance parameters. Free aggregate rows are validation diagnostics, not benchmark elasticity estimates. R-squared is descriptive because changing ψ changes model-implied log idea flows, the dependent variable of the proxy regression; log RMSE and the objective are the coherent cross-ψ criteria. All rows use identical masks within sample, 25-year birth windows, a 25-year model lag, and positive Pantheon counts. A post-1500 subsample row was retired from this table; its record remains in the replication package. The symbol ψ is the cultural elasticity, η is the Pantheon proxy elasticity, Log RMSE is the root-mean-square error of log idea flows, Objective is the corresponding log sum of squared residuals, and N is the number of fitted birth windows.

Imposing the panel ψ costs almost nothing in aggregate fit—the fixed and free log RMSEs are 0.443 and 0.421. R^2 is descriptive across values of ψ because changing ψ changes the model-implied log idea flows being fit. The objective profile, alternative windows, proportional-proxy mappings, and historical aggregate rows remain in Online Appendix Section OA.G.6 as diagnostics or stress tests.

D.3 England Productivity as an External Benchmark

The historical England and Great Britain record provides an external comparison for the model's productivity path. The model is calibrated to Western aggregates, and the Broadberry and de Pleijt (2021) accounts never enter it, so measured English TFP is an outside series that can adjudicate between the model's realized path and its fixed-wedge counterfactual. England is the natural place to look because it is the Western economy with a productivity record long enough to span the transition. The comparison is a shape check on a Western-aggregate path, not an England-specific prediction, and the worked England country stock of Appendix C plays no role here.

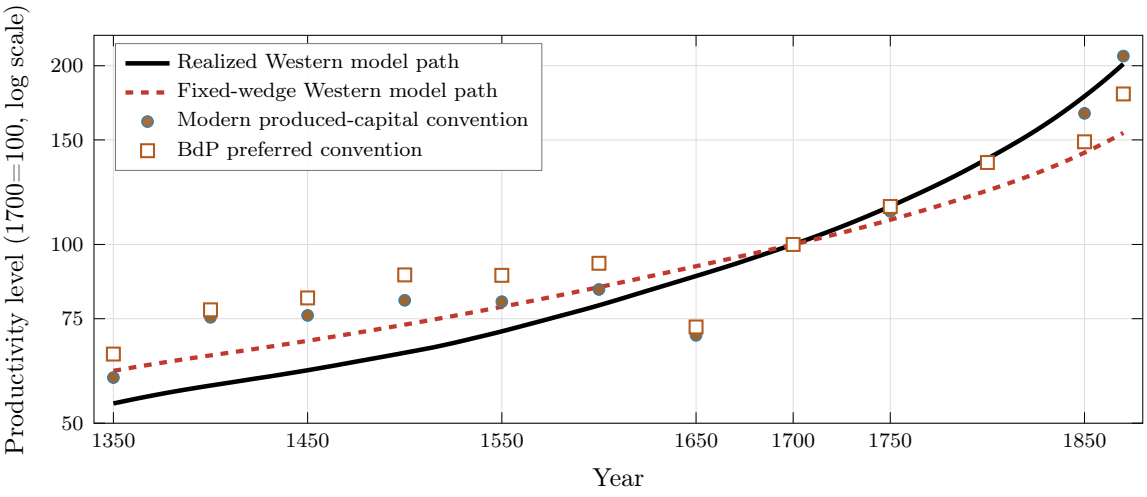


Figure 9: Historical England productivity shares the model's broad long-run rise.

Notes: All paths are indexed to 1700=100. The filled circles use domestic reproducible capital with a capital elasticity of 0.3, which Broadberry and de Pleijt report as their modern produced-capital robustness measure. The hollow squares use their preferred fixed capital excluding dwellings with a capital elasticity of 0.4. Output and point-year population come from Broadberry et al. (2015); capital comes from Broadberry and de Pleijt (2021), Tables 1B and 3A. England is used through 1699 and Great Britain from 1700. This is a one-country shape check on a Western-aggregate model path, not an England-specific prediction.

Figure 9 plots measured TFP in levels against the realized and fixed-wedge paths from a common anchor, with the measured series built from the Broadberry and de Pleijt (2021)

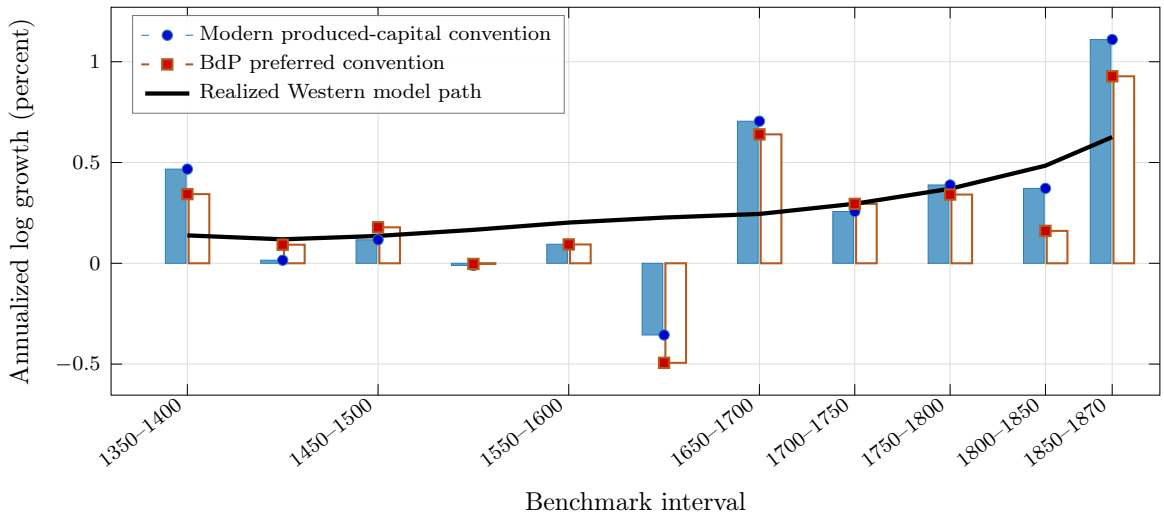


Figure 10: Measured benchmark growth is less smooth than the model path.

Notes: Bars are annualized log changes between consecutive published capital markers; the line applies the same calculation to the realized model path. All paths are indexed to 1700=100. The filled bars use domestic reproducible capital with a capital elasticity of 0.3, which Broadberry and de Pleijt report as their modern produced-capital robustness measure. The open bars use their preferred fixed capital excluding dwellings with a capital elasticity of 0.4. Output and point-year population come from Broadberry et al. (2015); capital comes from Broadberry and de Pleijt (2021), Tables 1B and 3A. England is used through 1699 and Great Britain from 1700. This is a one-country shape check on a Western-aggregate model path, not an England-specific prediction.

capital and output accounts following the accounting approach of Stefanski and Trew (2024b), and Figure 10 plots the corresponding benchmark growth rates. Under the standard produced-capital convention, measured TFP tracks the realized path across five centuries and ends far from the counterfactual; by 1870 it lies within 3.1 percent of the realized path and 29.8 percent from the counterfactual. Table 27 reports the distances both ways, anchored at a common year and anchor-free, so the reader can see that the tracking is not an artifact of where the paths are pinned together.

Under Broadberry and de Pleijt's own convention, which excludes dwellings and assigns capital a larger share, the measured series rises less steeply and lands between the two paths, so that convention does not distinguish them. The check is one-sided in that sense. The standard convention favors the realized path; no convention favors the counterfactual.

Table 27: Log-distance comparisons for the England productivity external benchmark

Panel A. Realized model path

Window	Scaling	Log RMSE	Endpoint gap
Produced capital, $\alpha = 0.3$			
1350–1870	1700 index	0.142	3.1
1350–1870	Scale profiled	0.131	2.5
1700–1870	1700 index	0.034	3.1
1700–1870	Scale profiled	0.031	4.3
BdP fixed capital, $\alpha = 0.4$			
1350–1870	1700 index	0.194	11.6
1350–1870	Scale profiled	0.177	19.5
1700–1870	1700 index	0.095	11.6
1700–1870	Scale profiled	0.072	5.5

Panel B. Fixed-wedge model path

Window	Scaling	Log RMSE	Endpoint gap
Produced capital, $\alpha = 0.3$			
1350–1870	1700 index	0.141	29.8
1350–1870	Scale profiled	0.130	24.3
1700–1870	1700 index	0.159	29.8
1700–1870	Scale profiled	0.105	17.8
BdP fixed capital, $\alpha = 0.4$			
1350–1870	1700 index	0.135	15.1
1350–1870	Scale profiled	0.110	7.3
1700–1870	1700 index	0.088	15.1
1700–1870	Scale profiled	0.053	8.0

Panel C. Supporting correlations

Window	Log level		Marker growth	
	Realized	Fixed	Realized	Fixed
Produced capital, $\alpha = 0.3$				
1350–1870	0.956	0.942	0.329	0.307
1700–1870	0.993	0.992	0.721	0.687
BdP fixed capital, $\alpha = 0.4$				
1350–1870	0.940	0.930	0.144	0.120
1700–1870	0.982	0.983	-0.795	-0.823

Notes. Endpoint gaps are absolute log-point differences multiplied by 100. Plain comparisons index every path to 1700=100. Scale-profiled comparisons choose a separate least-squares-optimal multiplicative scale for the measured England series against each model path and are therefore anchor-free. Correlations are supporting shape statistics; trending levels can produce high correlations even when benchmark growth timing differs. The produced-capital convention uses domestic reproducible capital. BdP denotes the Broadberry and de Pleijt fixed-capital convention, and Log RMSE is the root-mean-square log distance. The model path is Western aggregate, while the measured series is England through 1699 and Great Britain thereafter.

D.4 Accounting Robustness

This subsection stress-tests the accounting three ways—by the margin held fixed, by the parameters, and by the construction of the inputs. Table 28 starts with the benchmark and the counterfactuals. A component counterfactual freezes one dimension's stock at its year-1000 value while every other dimension follows its measured path. A block counterfactual freezes a pair that enters the wedge together—PD and UA in the numerator, or INDIV and LTO in the denominator. All rows share the same panel elasticity, population path, and normalization, with the scale set so realized productivity growth equals 1 percent in 1920, so differences across rows isolate the contribution of whatever is frozen.

Table 29 then varies the parameters rather than the paths, collecting the fixed-elasticity stress tests and parameter checks. The alternative-formula accounting, under both the benchmark elasticity and construction-specific re-estimation, is in Appendix B.5.2; implementation details and the classified robustness rows remain in Online Appendix Section OA.G.

Table 28: Benchmark growth accounting and component counterfactuals

			Contribution to annual productivity growth	
Counterfactual	ψ	Fixed share	1700	1920
Benchmark				
Full Innovation Wedge fixed	0.631	0.563	0.084	0.387
Component freezes				
PD fixed	0.631	0.880	0.022	0.079
UA fixed	0.631	0.902	0.017	0.075
INDIV fixed	0.631	0.927	0.016	0.060
LTO fixed	0.631	0.803	0.026	0.106
Block freezes				
PD+UA fixed	0.631	0.800	0.038	0.150
INDIV+LTO fixed	0.631	0.722	0.045	0.199

Notes. All rows use the maintained benchmark panel elasticity and accounting inputs. Fixed share is counterfactual productivity divided by realized productivity. The two growth columns report contributions to annual productivity growth in percentage points. The symbol ψ is the cultural elasticity. PD, UA, INDIV, and LTO denote Power Distance, Uncertainty Avoidance, Individualism, and Long-Term Orientation. Component and block rows freeze only the terms named in the table; the complete dimension matrix remains in the Online Appendix.

Table 29: Elasticity and parameter sensitivity

Variation	ψ	Fixed share	Growth contribution at 1920
Elasticity stress tests			
Fixed elasticity 0.2	0.200	0.801	0.133
Fixed elasticity 0.4	0.400	0.666	0.257
Fixed elasticity 0.6	0.600	0.574	0.370
Fixed elasticity 1.0	1.000	0.466	0.558
Fixed elasticity 1.5	1.500	0.407	0.726
Fixed elasticity 2.0	2.000	0.387	0.835
Parameter and functional-form robustness			
Feasible parameter grid (15 cells)	0.631	0.393–0.664	0.234–0.529
Baseline ratio	0.631	0.563	0.387
Log ratio	0.631	0.673	0.281
Additive signed index	0.631	0.646	0.321
Signed first principal component	0.631	0.474	0.958

Notes. Fixed share is counterfactual productivity divided by realized productivity, and the growth column reports the contribution to annual productivity growth in percentage points. The symbol ψ is the cultural elasticity. Rows in the first group vary that elasticity by assumption. Rows in the second group hold it at the maintained panel estimate while varying the named construction or the feasible parameter grid. The symbols λ and ϕ are the idea-production parameters defined in the model section.

Table 30 closes with the input constructions—anchor year, score scale, corpus composition, and resampling—all at the fixed panel elasticity, separating robustness rows from the unconditional-bootstrap stress test.

Table 30: Anchor, score-scale, composition, and bootstrap robustness

Role	Variation	ψ	Fixed share	Δg_{1920}
Benchmark	Wedge fixed in 1000	0.631	0.563	0.387
Robustness	Wedge fixed in 1300	0.631	0.698	0.323
Robustness	Wedge fixed in 1500	0.631	0.623	0.356
Robustness	Quantile-normalized scores	0.631	0.616	0.334
Robustness	Fixed genre mix	0.631	0.598	0.356
Robustness	Fixed subject mix	0.631	0.481	0.425
Stress test	Unconditional book bootstrap	0.631	0.387–0.826	0.253–0.545
Robustness	Year-stratified book bootstrap	0.631	0.365–0.786	0.277–0.535

Notes. Fixed share is counterfactual productivity divided by realized productivity, and Δg_{1920} is the contribution to annual productivity growth in percentage points. The symbol ψ is the cultural elasticity. The benchmark and robustness rows hold the elasticity at the maintained panel estimate and use the named reconstructed path or anchor. Bootstrap cells are the maintained fixed-elasticity intervals. The unconditional bootstrap is a stress test, while aggregate re-estimation remains a diagnostic in the Online Appendix.

Across implementations that hold the panel elasticity fixed, the 1920 contribution spans 0.23 to 0.96 percentage points. Re-estimated aggregate elasticities remain diagnostics because the aggregate objective does not identify the elasticity.

D.5 Research Productivity and Talent Allocation

The counterfactual identifies the composite elasticity of idea production with respect to culture. This subsection assesses how that elasticity divides between research productivity and the allocation of talent into research.

Pantheon births per million serve as a proxy for entry into idea work. Regressing their log-one-plus value on the 25-year mean Innovation Wedge with country and period fixed effects in the current 13-country benchmark panel gives an allocation elasticity of 0.455, with a 95 percent interval from 0.095 to 0.815. The residual research-productivity elasticity is 0.176, with an interval from -0.184 to 0.537. This is an imprecise upper-bound diagnostic. The counterfactual relies on the composite-elasticity invariance, not on the split. Formally, the calibration pins only the product $\bar{\delta} s_A^{\lambda}$ through the 1920 normalization, and the counterfactual moves only the wedge path. Any division of the composite elasticity between research productivity and the allocation of talent leaves that product unchanged, so the accounting is invariant to the split, exactly as in the main-text mechanism. Treating part of the effect as reallocation would matter for output only through the labor withdrawn from goods production, and because notable researchers are a vanishing share of the workforce, that correction is second

order. Full regression output and the invariance calculation remain in Online Appendix Section OA.G.7.

E Geography and Spatial Validation

The appendix reports the map and grid construction, the European and North American displays, and the complete urbanization checks cited in Section 7.

E.1 Geographic Construction

We use two spatial constructions. The one-degree descriptive grid admits point and capital origins and draws the maps; the three-degree regression grid admits point-resolved books only. Section 7 states the distinction; the complete origin rules, recursion, smoother, support mask, and HYDE construction are in Online Appendix Section OA.H.

The descriptive grid maps the production of durable written culture. We observe where a text originated rather than where it was read, so production and consumption can diverge within cells even though the aggregate Western series is less exposed to that gap. Book-level scoring noise also survives more readily in thin cells than in the pooled stock. These limits motivate a descriptive interpretation for the maps and a coarser point-only construction for prediction.

Under the map assignment rule, the cell-year count is the number of source-work rows assigned to cell g in year t. The log-damped volume and score-weighted momentum recursions from Section 3.3 then run separately within each cell, after which the component stocks and wedge ratios are formed. The result is an annual panel of fixed cells carrying surviving literary mass and cultural composition.

The prediction exercise rebuilds these stocks on three-degree cells using point-resolved books only. Country-level origins assigned to modern capitals therefore create no regression variation. Book inflows stop in 1920. Three-degree cells pool enough books to reduce thin-cell noise while retaining regional variation, whereas one-degree cells are noisier and five-degree cells leave fewer units for within-Europe inference. The one- and five-degree estimates appear in Table 35.

E.2 European Surface and Full Urbanization Results

Panel A of Figure 6 is the color-gradient companion to the lower-tail path in Panel B. Each year's smoothed values are converted to percentiles, which removes the common continental decline and displays spatial ordering. A region can therefore remain dark relative to other regions while its absolute wedge falls. The surface is clipped to Europe excluding Russia, with Turkey retained, and is shown only where the stocked-cell support rule is met.

The two displays answer different questions. The path follows the low tail as it moves from eastern Mediterranean and Byzantine support through the Italian maritime and urban system and toward the Low Countries, the Rhine corridor, and adjacent northwestern and central Europe. Its period wedge is the volume-stock-weighted average of each cell's annual wedges before the cells are rank weighted. The path uses no hard cutoff, does not select the raster minimum, and adds no cross-cell book-volume weights. The relative surface supplies spatial contrasts that this centroid cannot show, including darker Iberian and lighter French areas in later snapshots. The display alone does not attribute those contrasts to particular historical episodes. Appendix B.3 reports the separate event comparisons, and Appendix E.3 applies the same descriptive construction to North America and the pooled transatlantic stock.

Table 31 reports the full set of urbanization results referenced from Section 7. Panel A of Table 31 reports separate regressions for all eight measures. Panel B of Table 31 enters the gaps and own levels of the four Innovation-Wedge components jointly. Panel C of Table 31 weights neighbors by literary-stock volume, and Panel D adds country-by-period effects. The table therefore retains the weighting variants, full component set, and both ratio measures outside the compressed main text.

Table 31: Full urbanization specifications and cultural measures

	Cultural measures							
Specification	PD	UA	INDIV	LTO	MAS	INDUL	TF	IW
Panel A. Separate regressions								
Specification strip	Cell FE Yes	Decade FE Yes	Inference Conley	Outcome horizon 100 yr				
Local gap	-0.373^{***}	-0.535^{***}	0.388^{***}	0.054	-0.328^{***}	0.393^{***}	-12.100^{***}	-3.069^{***}
	(0.057)	(0.077)	(0.075)	(0.082)	(0.083)	(0.078)	(3.031)	(0.850)
Observations	712	712	712	712	712	712	699	697
Cells	43	43	43	43	43	43	43	43
Within R^2	0.169	0.166	0.134	0.003	0.103	0.071	0.061	0.053
Panel B. Joint regression, core four components								
Specification strip	Cell FE Yes	Decade FE Yes	Inference Conley	Outcome horizon 100 yr				
Local gap	-0.260^{**}	-0.081	0.131	0.175^{**}	–	–	–	–
	(0.123)	(0.183)	(0.134)	(0.078)				
Observations	712	712	712	712				
Cells	43	43	43	43				
Within R^2	0.196	0.196	0.196	0.196				
Panel C. Volume-weighted neighbors								
Specification strip	Cell FE Yes	Decade FE Yes	Inference Conley	Outcome horizon 100 yr				
Local gap	-0.146^{***}	-0.156^{**}	0.150^{**}	-0.078^{*}	-0.184^{***}	0.063	−0.500	0.343
	(0.047)	(0.078)	(0.073)	(0.045)	(0.050)	(0.091)	(2.743)	(0.288)
Observations	712	712	712	712	712	712	699	697
Cells	43	43	43	43	43	43	43	43
Within R^2	0.056	0.042	0.049	0.010	0.049	0.019	0.012	0.010
Panel D. Country-by-period fixed effects								
Specification strip	Cell FE Yes	Decade FE Yes	Inference Conley	Outcome horizon 100 yr				
Local gap	−0.066	−0.139	0.031	−0.040	−0.052	−0.032	−2.088	1.339^{**}
	(0.056)	(0.100)	(0.061)	(0.078)	(0.059)	(0.110)	(3.983)	(0.651)
Observations	712	712	712	712	712	712	699	697
Cells	43	43	43	43	43	43	43	43
Within R^2	0.139	0.146	0.148	0.127	0.164	0.129	0.125	0.133

Notes. Columns are cultural measures and panels are regression specifications. Panel A estimates each measure separately. Panel B enters the own levels and local gaps of the four Innovation-Wedge components jointly. Its populated coefficient cells are rotating focal coefficients from that one regression, and the remaining dashes are structural because non-core measures do not enter it. Panels C and D repeat the separate regressions using literary-stock neighbor weights and country-by-period fixed effects. Every regression includes cell and decade fixed effects, the cell's own cultural stock, and baseline urbanization. Parentheses contain Bartlett–Conley standard errors with 1,000-kilometer spatial and 100-year temporal cutoffs. Neighbor inverse-distance weights are floored at 250 km. Borderless cells in Panel D are assigned to their dominant country by book count. The outcome is the 100-year change in the HYDE urbanization share, and within R^2 is calculated after absorbing the listed fixed effects. The sample uses point-resolved books on fixed $3^\circ \times 3^\circ$ cells for Europe excluding Russia, with Turkey retained. Baseline decades run from 1000 to 1820 and book inflows end in 1920. PD, UA, INDIV, LTO, MAS, INDUL, TF, and IW denote Power Distance, Uncertainty Avoidance, Individualism, Long-Term Orientation, Masculinity, Indulgence, Total Friction, and the Innovation Wedge. ***, **, and * denote significance at 1, 5, and 10 percent.

The stricter panels qualify rather than reverse the component result. Hierarchy remains negative throughout and is usually precise, while uncertainty avoidance and individualism keep their signs. MAS is negative wherever it appears. The excluded margins are also informative. Competitive striving predicts less urbanization and indulgence more, each sharply estimated in the baseline. Indulgence's placement in Total Friction is therefore consistent with its competing interpretation as freedom from rigid social control. The Innovation Wedge ratio is negative in the baseline and changes sign under volume weighting, country-by-period effects, and alternative grids. Total Friction is strongly negative in the baseline and collapses under the tighter specifications. The own-trait coefficients usually point in the same direction as the corresponding gaps. These patterns support the component stocks as the more stable spatial objects and leave the pooled Innovation Wedge as the input to the aggregate growth exercise, where grid-level score noise is averaged over the full corpus.

E.3 North America and the Pooled Transatlantic Path

The North American maps apply the European machinery of Section 7 to the United States and Canada. The North American record begins on the English seaboard, and its center of gravity then moves steadily west through the nineteenth century—by roughly 10 degrees of longitude in the corpus mean. Within the United States the sharpest contrast runs between South and West. The highest wedges sit in the South, where the written record of a slave society carries hierarchy and high uncertainty avoidance, and the lowest in the West. These are within-year descriptive contrasts, not causal claims.

Figure 11 reports the North American Innovation-Wedge surface after converting each year's displayed smoothed values into within-year percentiles, so it shows relative low- and high-wedge areas within North America in that year. The figure uses Gaussian cell smoothing with bandwidth 450 km, square-root volume-stock weights, physical-land clipping, and the display rule that at least three active stocked cells lie within 900 km. This is the same display rule used for Europe—show locations with at least three active stocked cells within twice the smoothing bandwidth—but with a larger bandwidth because North American support is later, sparser, and displayed at a continental scale. There is no additional year-specific cutoff based on relative kernel support.

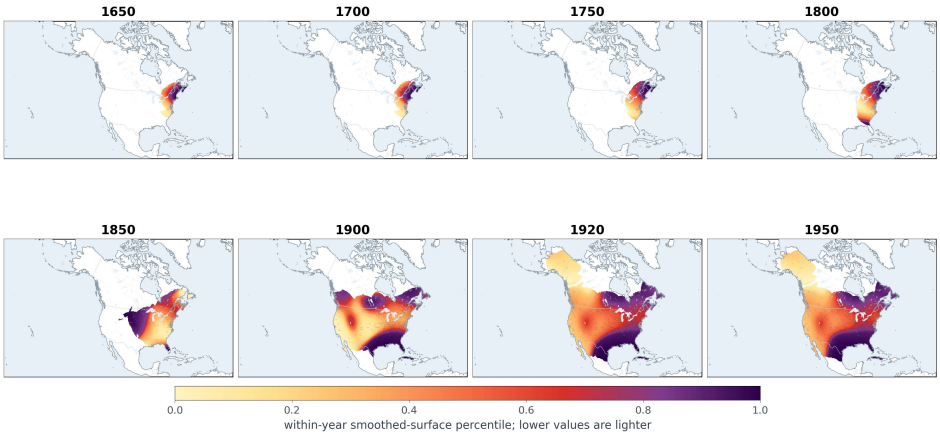
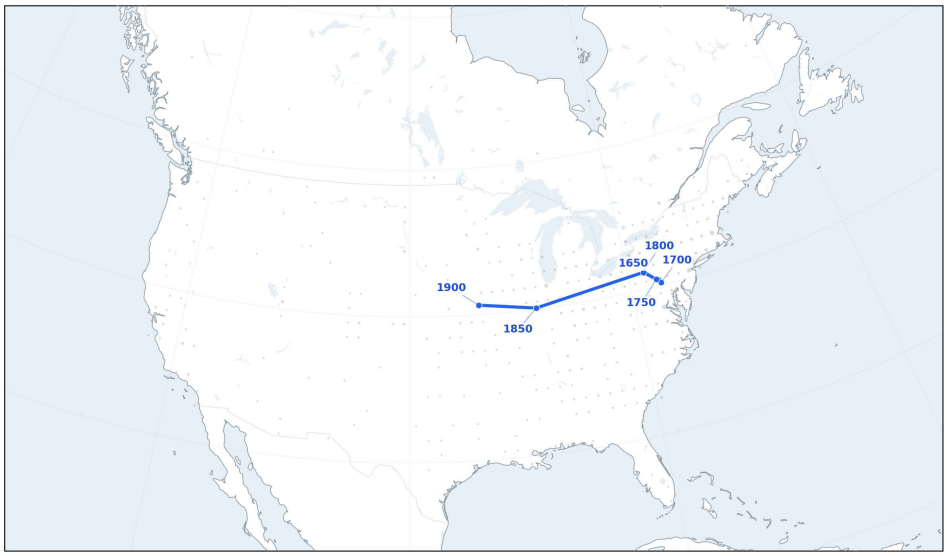


Figure 11: North American within-year relative Innovation-Wedge surface.

Notes: The figure uses $1° \times 1°$ modern-capital cell stocks for the United States and Canada. Displayed smoothed values in each year are converted into within-year percentiles. Lighter areas are lower-ranked Innovation-Wedge areas within the same year. Colors order cells within a year; the map is informative about spatial ordering, not magnitudes. The level of the wedge and its decline are reported in the aggregate-path figures and accounting tables. The smoothing bandwidth is 450 km, cell weights are $\sqrt{V_{g,t}}$, the display is clipped to physical land, and locations are shown only where at least three active stocked cells lie within 900 km. The same 900 km support rule is used for every displayed year.

Figure 12 reports the corresponding North American soft low-tail path and the pooled Europe–North America path. The North American path uses 50-year bins from 1650 through 1949 because the North American stock becomes informative much later than the European stock. The pooled path uses 100-year bins from 500 through 1949 and ranks all positive-stock cells in Europe and North America together. The pooled path is therefore a transatlantic centroid of the low-wedge cell distribution, not a physical route of diffusion across the ocean.



A North America.

B Pooled Europe and North America.

Figure 12: Soft low-tail Innovation-Wedge paths for North America and the pooled transatlantic stock.

Notes: Both panels use the same rank-weighted low-tail centroid as Figure 6. The North American panel uses 50-year nodes from 1650 to 1949. The pooled panel ranks Europe excluding Russia and the United States and Canada cells together in 100-year bins from 500 to 1949. The centroid is not constrained to lie on land and should be read as a summary of low-wedge mass in the pooled stocked cell distribution.

E.4 Spatial Robustness and Timing

The baseline local exposure averages observed neighbors within 1,000 kilometers, weights them by inverse distance with weights floored at 250 kilometers, and renormalizes over available neighbors when a value is missing. The literary-volume variant further weights neighbors by the size of their inherited stocks so that thinly stocked neighbors cannot drive the exposure.

The four tables test sensitivity to spatial noise, timing and regional composition, grid and support choices, and the measurement design. Table 32 varies the regression specification and neighbor weights. Table 33 checks non-overlapping windows, reverse timing, the two-region split, and leave-one-region-out ranges. Table 34 reports the spatial-noise draws, split-book exercise, and planted-signal calibration. Table 35 varies grid, geographic perimeter, and book-support thresholds. Power Distance, Uncertainty Avoidance, and Individualism each exceed all 500 simulated absolute statistics, which bounds rather than zeroes the tail probability at the simulation's resolution. Long-Term Orientation exceeds 166 draws and is null in the primary specification; Total Friction exceeds 495 draws and does not support a headline interpretation. The complete leave-one-region-out matrix for the estimated measures and the remaining variants are in Online Appendix Section OA.I.

Adjacent observations share most of their 100-year outcome window and are mechanically correlated; the covariance estimator uses Bartlett–Conley kernels with 1,000-kilometer and 100-year cutoffs, and a non-overlapping-window specification checks this directly.

The hierarchy, uncertainty-avoidance, and individualism gradients retain their signs in non-overlapping windows and when any one macro-region is removed. They do not predict urbanization changes that precede the measured cultural gaps. Total Friction fails these sharper timing and regional checks. It predicts the preceding century as well as the following one, with reverse coefficient -5.724 and $p = 0.027$, and changes sign across the two-region split, from 4.450 in the northwestern core to -13.806 in Mediterranean Europe. Consistent with Conley and Kelly (2025), persistent spatial gradients are not read as timing or mechanism.

Table 32: Alternative specifications for later urbanization

Specification	PD	UA	INDIV	LTO	MAS	INDUL	TF	IW
Joint core components	-0.260^{**} (0.123)	-0.081 (0.183)	0.131 (0.134)	0.175^{**} (0.078)	–	–	–	–
Volume-weighted neighbors	-0.146^{***} (0.047)	-0.156^{**} (0.078)	0.150^{**} (0.073)	-0.078^{*} (0.045)	-0.184^{***} (0.050)	0.063 (0.091)	-0.500 (2.743)	0.343 (0.288)
Country-by-period effects	-0.066 (0.056)	-0.139 (0.100)	0.031 (0.061)	-0.040 (0.078)	-0.052 (0.059)	-0.032 (0.110)	-2.088 (3.983)	1.339^{**} (0.651)
Point-or-capital assignment	-0.440^{***} (0.062)	-0.382^{***} (0.086)	0.333^{***} (0.092)	0.154 (0.097)	-0.245^{***} (0.070)	0.048 (0.095)	-12.574^{***} (2.661)	-5.816^{***} (0.997)

Notes. Each populated cell reports a local-gap coefficient with its Bartlett–Conley standard error in parentheses. The joint row is one regression containing the own levels and local gaps of PD, UA, INDIV, and LTO, and its remaining dashes are structural. Other rows estimate each measure separately. All rows use 100-year HYDE outcomes and 1000–1820 baselines. The benchmark uses point-only three-degree European cells, excluding Russia and retaining Turkey. Volume weights use surviving literary stock. Country-by-period effects assign each cell to its dominant country by book count. Point-or-capital assignment adds country-level source origins at modern capitals. PD, UA, INDIV, LTO, MAS, INDUL, TF, and IW denote Power Distance, Uncertainty Avoidance, Individualism, Long-Term Orientation, Masculinity, Indulgence, Total Friction, and the Innovation Wedge. ***, **, and * denote significance at 1, 5, and 10 percent.

Table 33: Timing and regional stability of the urbanization gradients

Specification	Cultural measures							
	PD	UA	INDIV	LTO	MAS	INDUL	TF	IW
Panel A. Timing checks								
Non-overlapping windows	-0.367^{***}	-0.503^{***}	0.404^{***}	-0.024	–	–	-9.872^{***}	–
	(0.081)	(0.104)	(0.093)	(0.075)			(3.300)	
Reverse timing	-0.013	-0.004	0.032	-0.006	–	–	-5.724^{**}	–
	(0.026)	(0.059)	(0.061)	(0.036)			(2.586)	
Panel B. Regional checks								
Northwestern core	-0.198	-0.288	0.466^{***}	0.073	–	–	4.450	–
	(0.134)	(0.198)	(0.125)	(0.128)			(6.191)	
Southern & eastern Mediterranean	-0.241^{***}	-0.315^{***}	0.239^{***}	0.048	–	–	-13.806^{***}	–
	(0.073)	(0.090)	(0.092)	(0.091)			(4.305)	
Leave-one-region-out, weakest $\lvert t\rvert$	-0.206^{***}	-0.599^{***}	0.134^{*}	-0.000	–	–	-5.792	–
	(0.054)	(0.115)	(0.078)	(0.090)			(3.813)	

Notes. Each specification is a two-line record with the coefficient on the named line and its Bartlett–Conley standard error on the blank-stub line below. The covariance estimator uses 1,000-kilometer spatial and 100-year temporal cutoffs. The leave-one-region-out row reports the omission producing the smallest absolute Conley statistic for each estimated measure. The complete omission matrix for the estimated measures remains in the Online Appendix. Non-overlapping rows retain every tenth baseline decade. Reverse timing uses the preceding rather than following century. The regional rows split the sample into the northwestern core and southern and eastern Mediterranean. MAS, INDUL, and IW were not estimated in this battery, so their dashes are structural. PD, UA, INDIV, LTO, MAS, INDUL, TF, and IW denote Power Distance, Uncertainty Avoidance, Individualism, Long-Term Orientation, Masculinity, Indulgence, Total Friction, and the Innovation Wedge. ***, **, and * denote significance at 1, 5, and 10 percent.

Table 34: Measurement and placebo diagnostics for the urbanization design

Specification	PD	UA	INDIV	LTO	TF
	500/500	500/500	500/500	166/500	495/500
	$\lvert t\rvert = 6.503$	$\lvert t\rvert = 6.964$	$\lvert t\rvert = 5.163$	$\lvert t\rvert = 0.662$	$\lvert t\rvert = 3.992$
Spatial-noise placebo	$\widehat{\beta} = -0.373$	$\widehat{\beta} = -0.535$	$\widehat{\beta} = 0.388$	$\widehat{\beta} = 0.054$	$\widehat{\beta} = -12.100$
	Share = 1.00	Share = 1.00	Share = 1.00	Share = 0.75	Share = 0.70
Split-book cross-fit	$\widetilde{\beta} = -0.324$	$\widetilde{\beta} = -0.425$	$\widetilde{\beta} = 0.389$	$\widetilde{\beta} = 0.108$	$\widetilde{\beta} = -2.503$
	0.488	0.493	0.488	0.475	0.284
Planted-signal recovery	$[0.393, 0.544]$	$[0.418, 0.581]$	$[0.414, 0.638]$	$[0.387, 0.537]$	$[0.024, 0.634]$

Notes. These cells are diagnostics rather than regression estimates. Placebo cells report, from top to bottom, the draw count supplied by the maintained diagnostic source, the observed absolute Conley statistic, and the observed coefficient. Split-book cells report the expected-sign share and the median coefficient across deterministic half-corpus assignments. Planted-signal cells report the median and percentile interval supplied by the maintained diagnostic source. PD, UA, INDIV, LTO, and TF denote Power Distance, Uncertainty Avoidance, Individualism, Long-Term Orientation, and Total Friction. MAS, INDUL, and IW are omitted because these diagnostics were not estimated for them.

Table 35: Grid, scope, and support checks for later urbanization

Specification	PD	UA	INDIV	LTO	MAS	INDUL	TF	IW
Europe, one-degree grid	-0.608^{***} (0.079)	-0.889^{***} (0.135)	0.657^{***} (0.077)	-0.269^{*} (0.152)	-0.723^{***} (0.119)	0.610^{***} (0.182)	-21.522^{***} (5.304)	1.492^{***} (0.559)
Europe, three-degree grid	-0.373^{***} (0.057)	-0.535^{***} (0.077)	0.388^{***} (0.075)	0.054 (0.082)	-0.328^{***} (0.083)	0.393^{***} (0.078)	-12.100^{***} (3.031)	-3.069^{***} (0.850)
Europe, five-degree grid	-0.191^{***} (0.054)	-0.176^{***} (0.050)	0.163^{***} (0.061)	-0.172^{***} (0.046)	-0.071 (0.045)	0.167^{***} (0.052)	-4.621^{***} (1.754)	-4.005^{***} (1.044)
Global, three-degree grid	-0.335^{***} (0.038)	-0.553^{***} (0.076)	0.343^{***} (0.052)	0.039 (0.075)	-0.337^{***} (0.066)	0.478^{***} (0.082)	-14.317^{***} (3.144)	-3.104^{***} (0.992)
Contemporaneous support at least 5	-0.183^{*} (0.109)	0.092 (0.132)	0.026 (0.091)	-0.322^{***} (0.094)	-0.639^{***} (0.140)	-0.264^{*} (0.149)	7.533^{***} (2.889)	1.028^{***} (0.350)
Contemporaneous support at least 15	0.102^{*} (0.059)	0.136^{**} (0.067)	0.010 (0.076)	0.017 (0.110)	0.356^{***} (0.101)	-0.123^{**} (0.054)	-4.956 (6.261)	-1.662 (2.343)
Log-ratio specification	–	–	–	–	–	–	-30.276^{***} (4.300)	-17.001^{***} (2.246)

Notes. Each populated cell reports a local-gap coefficient with its Bartlett–Conley standard error in parentheses. Grid rows rebuild point-only stocks and exposures at the stated resolution. Europe excludes Russia and retains Turkey, and the global row changes only the perimeter. Support rows require the stated contemporaneous cell-book count. The log-ratio row is defined only for Total Friction and the Innovation Wedge, so component dashes are structural. All specifications retain the 100-year outcome, cell and decade effects, the own cultural measure, baseline urbanization, and 1,000-kilometer and 100-year Bartlett–Conley covariance. PD, UA, INDIV, LTO, MAS, INDUL, TF, and IW denote Power Distance, Uncertainty Avoidance, Individualism, Long-Term Orientation, Masculinity, Indulgence, Total Friction, and the Innovation Wedge. ***, **, and * denote significance at 1, 5, and 10 percent.

References

Daron Acemoglu, Davide Cantoni, Simon Johnson, and James A. Robinson. The consequences of radical reform: The French revolution. *American Economic Review*, 101(7):3286–3307, 2011. doi: 10.1257/aer.101.7.3286.

Alberto Alesina and Paola Giuliano. Culture and institutions. *Journal of Economic Literature*, 53(4):898–944, 2015. doi: 10.1257/jel.53.4.898.

Yann Algan and Pierre Cahuc. Inherited trust and growth. *American Economic Review*, 100(5):2060–2092, 2010. doi: 10.1257/aer.100.5.2060.

Ali Almelhem, Murat Iyigun, Austin Kennedy, and Jared Rubin. Enlightenment ideals and belief in progress in the run-up to the industrial revolution: A textual analysis. *Quarterly Journal of Economics*, 141(1): 263–314, 2026. doi: 10.1093/qje/qjaf054.

Sascha O. Becker and Ludger Woessmann. Was Weber wrong? A human capital theory of Protestant economic history. *Quarterly Journal of Economics*, 124(2):531–596, 2009. doi: 10.1162/qjec.2009.124.2.531.

Sascha O. Becker, Erik Hornung, and Ludger Woessmann. Education and catch-up in the industrial revolution. *American Economic Journal: Macroeconomics*, 3(3):92–126, 2011. doi: 10.1257/mac.3.3.92.

Roland Bénabou, Davide Ticchi, and Andrea Vindigni. Forbidden fruits: The political economy of science, religion, and growth. *Review of Economic Studies*, 89(4):1785–1832, 2022. doi: 10.1093/restud/rdab069.

Alberto Bisin and Thierry Verdier. The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298–319, 2001. doi: 10.1006/jeth.2000.2678.

Nicholas Bloom, Charles I. Jones, John Van Reenen, and Michael Webb. Are ideas getting harder to find? *American Economic Review*, 110(4):1104–1144, 2020. doi: 10.1257/aer.20180338.

Stephen Broadberry, Bruce M. S. Campbell, Alexander Klein, Mark Overton, and Bas van Leeuwen. *British Economic Growth, 1270–1870*. Cambridge University Press, Cambridge, 2015.

Stephen N. Broadberry and Alexandra de Pleijt. Capital and economic growth in britain, 1270–1870: Preliminary findings. CEPR Discussion Paper DP15889, CEPR, 2021.

Christopher Leslie Brown. *Moral Capital: Foundations of British Abolitionism*. University of North Carolina Press, Chapel Hill, 2006.

Eltjo Buringh and Jan Luiten van Zanden. Charting the "rise of the West": Manuscripts and printed books in Europe, a long-term perspective from the sixth through eighteenth centuries. *Journal of Economic History*, 69(2):409–445, 2009. doi: 10.1017/S0022050709000837.

Christopher Clark. *Iron Kingdom: The Rise and Downfall of Prussia, 1600–1947*. Harvard University Press, Cambridge, MA, 2006.

Gregory Clark. *A Farewell to Alms: A Brief Economic History of the World*. Princeton University Press, Princeton, 2007.

Timothy G. Conley and Morgan Kelly. The standard errors of persistence. *Journal of International Economics*, 153:104027, 2025. doi: 10.1016/j.jinteco.2024.104027.

N. F. R. Crafts. *British Economic Growth during the Industrial Revolution*. Clarendon Press, Oxford, 1985.

Paul A. David. The historical origins of open science: An essay on patronage, reputation and common agency contracting in the scientific revolution. *Capitalism and Society*, 3(2), 2008. doi: 10.2202/1932-0213.1040.

David de la Croix, Matthias Doepke, and Joel Mokyr. Clans, guilds, and markets: Apprenticeship institutions and growth in the preindustrial economy. *Quarterly Journal of Economics*, 133(1):1–70, 2018. doi: 10.1093/qje/qjx026.

Jeremiah E. Dittmar. Information technology and economic change: The impact of the printing press. *Quarterly Journal of Economics*, 126(3):1133–1172, 2011. doi: 10.1093/qje/qjr035.

Seymour Drescher. *Abolition: A History of Slavery and Antislavery*. Cambridge University Press, Cambridge, 2009. doi: 10.1017/CBO9780511770555.

Raquel Fernández. Cultural change as learning: The evolution of female labor force participation over a century. *American Economic Review*, 103(1):472–500, 2013. doi: 10.1257/aer.103.1.472.

Oded Galor and Ömer Özak. The agricultural origins of time preference. *American Economic Review*, 106 (10):3064–3103, 2016. doi: 10.1257/aer.20150020.

Yuriy Gorodnichenko and Gérard Roland. Which dimensions of culture matter for long-run growth? *American Economic Review*, 101(3):492–498, 2011. doi: 10.1257/aer.101.3.492.

Luigi Guiso, Paola Sapienza, and Luigi Zingales. Does culture affect economic outcomes? *Journal of Economic Perspectives*, 20(2):23–48, 2006. doi: 10.1257/jep.20.2.23.

Joseph Henrich. *The WEIRDest People in the World: How the West Became Psychologically Peculiar and Particularly Prosperous*. Farrar, Straus and Giroux, New York, 2020.

Geert Hofstede, Gert Jan Hofstede, and Michael Minkov. *Cultures and Organizations: Software of the Mind*. McGraw-Hill, New York, 3 edition, 2010.

Geert H. Hofstede. *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Sage Publications, Thousand Oaks, CA, 2 edition, 2001.

Robert J. House, Paul J. Hanges, Mansour Javidan, Peter W. Dorfman, and Vipin Gupta, editors. *Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies*. Sage Publications, Thousand Oaks, CA, 2004.

Charles I. Jones. R&D-based models of economic growth. *Journal of Political Economy*, 103(4):759–784, 1995. doi: 10.1086/262002.

Kees Klein Goldewijk, Arthur Beusen, Jonathan Doelman, and Elke Stehfest. Anthropogenic land use estimates for the Holocene—HYDE 3.2. *Earth System Science Data*, 9(2):927–953, 2017. doi: 10.5194/essd-9-927-2017.

Deirdre N. McCloskey. *Bourgeois Equality: How Ideas, Not Capital or Institutions, Enriched the World*. University of Chicago Press, Chicago, 2016.

Stelios Michalopoulos and Melanie Meng Xue. Folklore. *Quarterly Journal of Economics*, 136(4):1993–2046, 2021. doi: 10.1093/qje/qjab003.

Joel Mokyr. *The Enlightened Economy: An Economic History of Britain 1700–1850*. Yale University Press, New Haven, 2009.

Joel Mokyr. *A Culture of Growth: The Origins of the Modern Economy*. Princeton University Press, Princeton, NJ, 2016.

Douglass C. North and Barry R. Weingast. Constitutions and commitment: The evolution of institutions governing public choice in seventeenth-century England. *Journal of Economic History*, 49(4):803–832, 1989. doi: 10.1017/S0022050700009451.

Nathan Nunn. Culture and the historical process. *Economic History of Developing Regions*, 27(sup1): S108–S126, 2012. doi: 10.1080/20780389.2012.664864.

Nathan Nunn and Leonard Wantchekon. The slave trade and the origins of mistrust in Africa. *American Economic Review*, 101(7):3221–3252, 2011. doi: 10.1257/aer.101.7.3221.

Project Gutenberg. Project Gutenberg. `https://www.gutenberg.org`, 2026. Public-domain digital library; English-language collection.

Paul M. Romer. Endogenous technological change. *Journal of Political Economy*, 98(5, Part 2):S71–S102, 1990. doi: 10.1086/261725.

Jonathan F. Schulz, Duman Bahrami-Rad, Jonathan P. Beauchamp, and Joseph Henrich. The church, intensive kinship, and global psychological variation. *Science*, 366(6466):eaau5141, 2019. doi: 10.1126/science.aau5141.

Shalom H. Schwartz. Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in Experimental Social Psychology*, volume 25, pages 1–65. Academic Press, 1992. doi: 10.1016/S0065-2601(08)60281-6.

Shalom H. Schwartz. *Cultural Value Orientations: Nature and Implications of National Differences*. State University–Higher School of Economics Press, Moscow, 2008.

Mara P. Squicciarini. Devotion and development: Religiosity, education, and economic progress in nineteenth-century France. *American Economic Review*, 110(11):3454–3491, 2020. doi: 10.1257/aer.20191054.

Mara P. Squicciarini and Nico Voigtländer. Human capital and industrialization: Evidence from the age of enlightenment. *Quarterly Journal of Economics*, 130(4):1825–1883, 2015. doi: 10.1093/qje/qjv025.

Radek Stefanski and Alex Trew. Selection, patience, and the interest rate. *Journal of Political Economy Macroeconomics*, 2(1):149–181, 2024a. doi: 10.1086/727798.

Radoslaw Stefanski and Alex Trew. Accounting for interest: Understanding 700 years of declining real interest rates. Working paper, 2024b.

Guido Tabellini. Culture and institutions: Economic development in the regions of europe. *Journal of the European Economic Association*, 8(4):677–716, 2010. doi: 10.1111/j.1542-4774.2010.tb00537.x.

Jan Teorell, Aksel Sundström, Sören Holmberg, Bo Rothstein, Natalia Alvarado Pachon, Victor Saidi Phiri, Chuwei Chen, and Zhen Liu. The quality of government standard dataset, version Jan26, 2026.

United Nations, Department of Economic and Social Affairs, Population Division. World population prospects 2024. Online database, 2024. URL `https://population.un.org/wpp/`.

Nico Voigtländer and Hans-Joachim Voth. Persecution perpetuated: The medieval origins of anti-semitic violence in Nazi Germany. *Quarterly Journal of Economics*, 127(3):1339–1392, 2012. doi: 10.1093/qje/qjs019.

Amy Zhao Yu, Shahar Ronen, Kevin Hu, Tiffany Lu, and César A. Hidalgo. Pantheon 1.0, a manually verified dataset of globally famous biographies. *Scientific Data*, 3:150075, 2016. doi: 10.1038/sdata.2015.75.