Predicting Protestantism: Writing, Biography and Conversion during the Reformation

Marcel Caesmann* Hans-Joachim Voth†

November 2026

Abstract

Who will join a revolution? We examine a paradigmatic ideological rupture with major repercussions: the Protestant Reformation. For 176 authors who printed before 1517, 48 of whom later broke with Rome, we ask whether what they had published predicts the break. Embeddings of the text rank a later convert above an author who stayed with Rome in 65 of 100 draws (background facts such as year of birth and language: 71). A language model that compares masked passages two at a time, never seeing a name or a label, gives each author a score: a one-standard-deviation higher score is associated with a 9 percentage point higher probability of a later break. The score separates the sides about as well among authors of the same birth cohort or the same clerical status. The signal sits in religious writing, on the quarrels of the fifteenth century and in intellectual outlook, not on the doctrines that would define the movement. After 1517 the trace becomes much clearer. On what each convert printed before his own break, a one-standard-deviation higher score is associated with a 22 percentage point higher probability; converts' texts move toward reform before they break, and the two sides divide along Luther's profile, converts toward it and Catholics away from it. While other factors such as personal connections, economic interests and patronage links may well have played a role, published text predicts one of the most consequential decisions open to the learned of the age: to side with Luther or not.

Keywords: Reformation, ideological alignment, text as data, humanism, large language models.
JEL codes: N33, Z12, C55, D74.

*WU Vienna. Email: marcel.caesmann@wu.ac.at.
†University of Zurich and CEPR. Email: voth@econ.uzh.ch.

> "Virtually no one in 1500 could have imagined that within a century, Western Europe would be split between mutually exclusive versions of Christianity."
>
> — Diarmaid MacCulloch (1987)

1 Introduction

On October 31, 1517, an obscure professor of theology at a provincial university nailed a piece of paper to a church door in Wittenberg. It contained 95 theses. The Catholic Church initially treated the incident as a nuisance: Pope Leo X is said to have dismissed Luther as "a drunken German" who "will feel different when he is sober."[1] Within four years, the narrative had changed: In February 1521, the papal nuncio Aleander reported from Worms that "all Germany is in great turmoil; nine tenths are raising the field cry 'Luther!,' and for the remaining tenth, if they are indifferent to Luther, the slogan is at least: 'Death to the Roman court!' "[2] Within a decade, roughly half of Germany had left the Roman Church. Ever since, a schism divides Western Christianity, altering the relationship between political and ecclesiastical authority, and reorganising the institutions through which Europeans worship, learn, and are governed (MacCulloch, 2004; Cameron, 2012). The puzzle is how sudden it was: "perhaps the most surprising aspect of the Reformation, was that it spread beyond Wittenberg in the first place" (Becker et al., 2016).

How could an obscure German scholar suddenly split Western Christendom? Kuran (1991, 1995) argues that preference falsification can make revolutions unpredictable *in principle*. When dissent is costly, people profess loyalty they do not feel; private opposition can therefore accumulate invisibly until a small shock tips a cascade. In its strict form, the argument implies that 1517 was unforeseeable because the information relevant to the coming rupture had left no public trace.

In this paper, we examine whether joining the Reformation was predictable. Using only

[1] Attributed; Bainton (1950) judges the remark "in all likelihood" inauthentic—"yet revealing." It is revealing precisely because it preserves what informed contemporaries thought the theses were worth.

[2] Aleander's dispatch of February 8, 1521, *Die Depeschen des Nuntius Aleander vom Wormser Reichstage 1521*, trans. P. Kalkoff (Halle, 1886), p. 43, quoted in Dieter (2021).

what people had already written and put into print before 1517, could we have predicted who would later break with Rome? Once the controversy had begun, the same question can be asked of each future convert's writing before his own break. This is a prediction exercise about individual choices: it asks whether later alignment could be predicted from earlier public writing. If these writings carry no signal of later confession, even to modern machine learning, then later individual alignment was illegible ex ante. If they do carry signal, the coming divide was at least partly legible at the level of individuals, and preference falsification had not erased every public trace of later alignment.

If the coming schism did leave a public trace, where should we look for it? Moeller argued that "without humanism, no Reformation": if any trace of the coming schism was visible before 1517, it should run through humanism (Moeller, 1959).[3] Other aspects—such as long-standing criticisms of the sale of indulgences—may be similarly predictive. To our knowledge, ours is the first quantitative test of the predictive power of pre-existing ideas for joining a revolution, at the level of the individual.

The Reformation suits this test. It was an ideological revolution, and many of its protagonists were writers by profession: humanists, jurists, theologians, physicians, and clerics who left a dense trail of printed Latin and German. The invention of movable type had, by 1500, produced a large and datable corpus. The outcome is recorded, person by person, in biographical and archival sources: who later broke with Rome and who stayed Catholic.

We use all surviving pre-1517 printed output of every author whose works we can both locate in the German bibliographic record (VD16) and recover as digitised books. We then identify the authors who lived long enough to face the confessional choice and whose later alignment can be established from biographical evidence. For every such author, we transcribe pages sampled across the surviving works. With the authors a language model can name left out, this gives 176 authors, 48 of whom later broke with Rome.

We mask every name of a person, a place and an institution in these texts and measure them in two ways. The first turns each author's text into an embedding, a numerical representation of the text, and predicts the break from it with a model fitted on other authors. It needs no language model. The second requires a comparison: a language model evaluates two passages by different authors, without a name or a label, and says which of the two leans further toward reform. A Bradley–Terry model turns its verdicts

[3]Moeller's argument concerned humanism's enabling role in the early Reformation—especially the humanists' "productive misunderstanding" of Luther—not a rule predicting which individuals would ultimately leave Rome.

into one scale, and an author's score is the mean strength of his passages.

First, what these authors printed before 1517 predicts who would later break with Rome.[4] The embeddings rank a randomly drawn later convert above a randomly drawn Catholic with an AUC of 0.65, almost as well as the background facts about an author—his year of birth, the language, length and date of his works, and where they were printed—which reach 0.71. The tournament reaches 0.63. Ranked by the score and cut into quintiles, 16 of the 36 authors in the most reform-leaning quintile later broke with Rome, and 6 of the 35 in the most Roman. A one-standard-deviation higher score is associated with an 8.7 percentage point higher probability of a later break, over a base of 27.3 percent, and the association remains with controls for language, the length of the text, year of birth, distance to Wittenberg and fame. The score also orders the authors by degree: the further an author's pre-1517 text leans toward reform, the earlier he breaks.

The result holds even if we eliminate authors that the language model can identify or the well-known figures of the age, and it holds among authors of the same education, the same stake in the old church, the same cohort and the same language. Within birth cohorts the AUC is 0.59 against 0.61 over all pairs, and within clerical status 0.60 against 0.62 (Appendix Table C6).

Second, the signal reflects a broad intellectual and religious outlook that predates the Protestant programme. The separation sits in religious writing, where the score ranks the two sides with an AUC of 0.70. The Catholics wrote more on religion, but this accounts for 15 percent of the gap between the two sides' mean scores; how the two sides wrote about religion accounts for 77 percent. To see on what they differ, the same language model codes every text on twelve axes: eight on the doctrines over which the two parties would contradict each other in writing, four on outlook, devotion, grievance and the demand for reform. The axes that separate the sides are the quarrels of the fifteenth century, over vows and the religious life, the pope's power, penance and indulgences, and with them outlook: inward piety and humanist against scholastic method. The two articles on which the movement would stand, justification by faith and Scripture against tradition, do not separate them. Much of the signal comes from the Catholics, who defend doctrine more firmly, whereas the later converts take no pronounced stance. In 1516 the future Protestants do not yet sound like Protestants; they sound like humanists who do not rise to defend the church.

Third, after 1517 the texts separate the two sides far more clearly. We run the same

[4] All 176 authors of the sample before 1517 are men, and we use masculine pronouns throughout. No woman who wrote in this milieu has text printed under her name before 1517 that meets our criteria. One, the Nuremberg abbess Caritas Pirckheimer, is among the 224 authors of 1517 to 1530.

tournament on what was printed from 1517 to 1530 and count a convert's text only to the year before his documented break, so that nothing he printed as a Protestant enters. The score now ranks a later convert above a Catholic with an AUC of 0.80, against 0.63 before 1517, and a one-standard-deviation higher score is associated with a 22.4 percentage point higher probability of a later break. We follow each author through his own texts: a convert's text leans further toward reform three to five years before his break than his own earlier text did, and further still in the two years before it, net of what moved every author in the same years.

In these years a programme appears. In what the converts printed between 1517 and their break, their stance is on the reform side on Scripture against tradition and on justification by faith, while the Catholics hold the Roman side on every question of doctrine. Stances that had been unrelated also come to hang together. Before 1517 an author's position on one question said little about his position on another; by the mid-1520s an author who took the reform side on one question took it on others too, and one who defended on one defended on others. The common line is Luther's: the converts' profiles over the twelve axes come to resemble his own profile of 1518 to 1521, and the Catholics' turn away from it; agreement with Luther's profile separates the sides with an AUC of 0.64 before 1517 and of 0.87 in 1522 to 1525. Before 1517 the positions were already in print, but nothing yet sorted them along a common line. A public event that supplies such a line is what coordination models describe (Chwe, 2001; Pinker, 2025).

Our findings relate, first, to the economics of the Reformation (Becker et al., 2016). The existing literature largely focuses on the supply and diffusion of the Reformation: distance to Wittenberg as the geography of adoption (Becker and Woessmann, 2009), the political economy of territorial adoption (Cantoni, 2012), printing as the transmission technology (Rubin, 2014; Dittmar, 2011), competition among media as the accelerant (Dittmar and Seabold, 2019), religious competition and the reallocation it unleashed (Ekelund et al., 2002; Cantoni et al., 2018), external shocks that bought the movement time (Iyigun, 2008), Luther's personal network and influence in its early spread (Becker et al., 2020), and universities that transmitted ideas and were themselves reorganised by confessional division (de la Croix et al., 2025; de la Croix and Morault, 2026). This work studies where and how the Reformation spread. We study which individuals were already inclined toward it. Closest to the individual level, Kim and Pfaff (2012) show that ideologically mobilised university students acted as bridge actors, carrying evangelical ideas from Wittenberg and Basel into their home towns, so that a town's exposure to Protestant- versus Catholic-aligned universities predicts whether it adopted the Reformation. Their

unit is the town, and a student's leaning is inferred from where he studied. We predict the choice of named individuals from what they printed before 1517. The conversion historiography reaches individuals only once evangelical preaching and confessional pressure were upon them (Pettegree, 2005); we step one stage earlier, before there was anything to join. We establish two facts. Later alignment was partly visible, author by author, in the public record before 1517. The texts of 1517 to 1530 then show the divide being drawn, each convert's before his own break.

Second, we also contribute to the literature on preference falsification, thresholds, and cascades. Two classes of models disagree about what a pre-revolutionary public record should contain. In Kuran (1991, 1995), public expression is falsified and the distribution of latent dissent is invisible: revolutions are unforecastable in principle because the relevant information exists in no public record. In threshold and coordination models (Granovetter, 1978; Morris and Shin, 1998, 2003; Chwe, 2001), private dispositions leave noisy, dispersed public traces, and the discontinuity comes from coordination—behaviour jumps when a public event turns dispersed signals into common knowledge (Lohmann, 1994; Heinemann et al., 2004; Pinker, 2025). Telling the classes apart requires individual-level evidence from before a rupture, which is what we supply. While positions may have been partly concealed, traces remained: who would later break with Rome was already apparent, author by author. We can rule out the claim that later alignment could not be forecast. The texts of 1517 to 1530 are then consistent with the second class of models: a public event turned positions that had been dispersed across separate questions into one question with two sides.

Finally, we relate to the text-as-data literature and the predict-then-interpret programme (Gentzkow et al., 2019a; Ash and Hansen, 2023; Dell, 2025; Mullainathan and Spiess, 2017; Ludwig and Mullainathan, 2024). Our setting has a feature most applications lack: the texts predate the categories they are used to predict—no author writing in 1515 could have called himself a Protestant, because the word did not yet exist. We use machine learning as the programme prescribes, first to establish that a signal exists and then to ask where it lies—in intellectual outlook and devotion, in doctrine, in grievance against the institution, or in an explicit demand for reform. The tournament brings comparative judgement, known from the scaling of political texts (Carlson and Montgomery, 2017; Benoit et al., 2019) and from ranking with language models (Qin et al., 2024), to historical text. We also guard against memorisation, the main danger of large pretrained models on historical text (He et al., 2025): every name is masked, a passage whose author the language model names with confidence is dropped, and the model is asked to continue

passages it may know.

The remainder of the paper is organised as follows. Section 2 sets the historical scene: the reforming world before 1517, what breaking with Rome meant, and Luther as a public event. Section 3 describes the corpus and the samples, the coding of each author's later side, and the three measures of the text. Section 4.1 predicts the break and its timing from the texts printed before 1517, Section 4.2 asks where in the text the signal sits, and Section 4.3 follows the two sides through 1517 to 1530. Section 5 weighs alternative explanations: memorisation, biography and stakes, language and genre, survival and inference. Section 6 concludes.

2 Historical Background

2.1 A Reforming World before 1517

The Church that Luther challenged was already unsettled. The half-century before 1517 produced persistent calls for reform *in capite et membris*, in head and members. Conciliarists had contested papal supremacy. The *devotio moderna* cultivated an interior, Christ-centred piety across the Low Countries and the Rhineland. Humanism trained a generation to prize the Latin of the ancients and the Greek of the New Testament over the technical Latin of the schools. Erasmus's 1516 Greek New Testament implied that the Vulgate, and the theology built upon it, rested in places on mistranslation (Ozment, 1980; Cameron, 2012).

The quarrels of these decades were separate, and each had its defenders. Doctors of theology wrote for and against indulgences and the penance behind them. Every estate complained of the wealth and conduct of the clergy, and the clergy answered. The orders fought over vows and the strict rule, Observants against Conventuals. The cult of the saints, of Mary and of relics drew new devotion and open mockery in the same years. The pope's power over councils, benefices and money was the standing grievance of the imperial diets (Ozment, 1980; Cameron, 2012). An author could criticise and argue for change on one of these questions and defend the Church on the next. No party, confession or list of articles bound his answers together.

None of this was yet Protestantism. Justification by faith, the article on which the movement would stand, was a question for the schools, not a quarrel of the Church, and few yet set Scripture against the tradition of the Church as rival authorities. Humanists and scholastics quarrelled over method and style well before they divided over doctrine,

and the currents of reform overlapped and cut across the line 1517 would later draw: Erasmus and many leading humanists remained in the Roman Church, and not every reformer had begun as a humanist (Moeller, 1959; Ozment, 1980). This was a cultural fault line before it carried confessional meaning, and one that still ran through what future converts and future Catholics had written before 1517.

What such dispositions had to do with individual conversion is among the oldest questions in the field. Febvre (1929) rejected the view that the Reformation arose from clerical abuses, locating its origins instead in *un immense appétit du divin*—an immense appetite for the divine—among the laity; Ozment (1975) found its appeal in relief from the burden of late-medieval religion. Both are claims about *individual* motivation, and both imply that what distinguished the future convert should be religious disposition rather than grievance. Section 4.2 tests that implication.

2.2 What Breaking with Rome Meant

Breaking with Rome was a sequence of acts by which a person made a confessional choice public and hard to reverse. The forms are unusually concrete: a priest married in defiance of celibacy; a cleric resigned a benefice, or was deprived of one, when his territory changed religion; an author subscribed to the Augsburg Confession or accepted an evangelical office; others faced formal proceedings for heresy, exile or—in the years after the Edict of Worms placed Luther's followers under the imperial ban—execution (Cameron, 2012; MacCulloch, 2004). Section 3.2 dates a convert's break by the earliest of these acts.

These acts could be costly, and the cost fell hardest on the authors our sample can observe. Many established authors held clerical offices, benefices or teaching positions that a break with Rome put directly at risk; the first evangelical martyrs were burned at Brussels in 1523, and heresy remained a capital crime across much of the Empire for decades thereafter (MacCulloch, 2004). The stakes were not uniform—they turned on a person's office, his territory's politics and the decade in which he chose—but for an author old and established enough to have published before 1517, with a benefice and a career to lose, the choice was rarely cheap. The clerical estate and the benefice are the stakes that Section 3.4 codes.

2.3 Luther as a Public Event

From the autumn of 1517 the quarrels had one name. Luther's theses on indulgences were in print within weeks. The Leipzig disputation of 1519 moved the question to the

pope's authority; the three treatises of 1520 set out a programme, on the sacraments, the priesthood of all believers and the freedom of a Christian; the hearing at Worms in 1521 and the ban that followed made adherence a public and punishable act (Oberman, 1989; MacCulloch, 2004). The press carried each step. Luther was the most printed author of his age, his German pamphlets reached towns across the Empire within weeks of their first printing, and the other side answered in the same medium (Scribner, 1981; Pettegree, 2005). By 1530 the Augsburg Confession and the Confutatio set the two positions down article by article (Cameron, 2012). What had been separate questions, each with its own defenders, was now one question with two sides.

How fast the event spread is well understood; who was ready for it is not. It helps to separate two things the Reformation literature often studies together. The *supply* of reform—who preached it, where the presses stood, how fast printed ideas travelled—is comparatively well identified, by geography and by technology (Rubin, 2014; Dittmar, 2011; Cantoni, 2012). The *demand* for reform—the disposition that may have made some readers receptive and others hostile—is harder to observe. We do not observe that private state. Our object is its possible public textual trace: the positions an author put into print, before 1517 and after it.

Contemporaries on both sides took the demand to be individual and fed by text. Johannes Cochlaeus, Luther's most persistent Catholic opponent, complained that once Luther's German New Testament circulated, "shoemakers and women and every kind of unlearned person" read it "most eagerly as the font of all truth," and within months "did not blush to dispute about the faith and the Gospel… with Masters and Doctors of Sacred Theology" (Vandiver et al., 2002, pp. 106–107). The remark is hostile testimony, but it locates the movement's energy where we do: in individual readers responding to what they read.

The individual level is also where the prediction question is cleanest. Whether a whole territory adopted reform depended on who held authority to decide: in the imperial free cities, councils and burgher pressure could carry a change of religion; in the principalities, the rule later codified as *cuius regio, eius religio* left the choice to the prince, and popular disposition could be overridden or imposed (Moeller, 1972; Cameron, 2012). A city's confession thus mixed the demand of its inhabitants with the decision of whoever ruled it. A person's observed alignment was an individual public act but remained constrained by territorial institutions and available opportunities. We therefore predict individuals, and the measured supply side—the distance from the town where most of an author's works were printed to Wittenberg—enters as a control (Section 4.1), not as the object of study.

3 Data

3.1 Corpus, Risk Set, and Sampling

We need the text that people put into print before 1517, transcribed from the original editions so that no later knowledge enters it. We begin from the German bibliographic record: every main author of an edition printed between 1501 and 1530 and catalogued in VD16, a frame of 2,118 authors. We then digitise the books themselves—page images drawn primarily from the Bavarian State Library and supplemented by a systematic search for digitised editions held in other libraries, transcribed by a vision model into running text. Every word of running text we analyse is our own transcription of an edition of the time; we import no modern edition that could carry later editorial knowledge back into the sample. For the prediction of Section 4.1 only editions dated before 1517 count, so that a later reprint cannot introduce a revised preface.

Prediction requires a population that faced the confessional choice at all. We fix this at-risk set before consulting text availability: an author must have been alive in 1517, must not be a worldly ruler whose catalogued works do not represent personal authorship, and must have been active in the German lands; where the record is silent on the year of death or the place of activity, the author stays in. This leaves 1,100 authors. The analysis sample then adds two requirements: a documented later side, stayed Catholic or broke with Rome (Section 3.2), and at least 2,000 characters of transcribed text. These reduce it to 178 authors with text printed before 1517, 48 of whom later broke with Rome and 130 of whom stayed Catholic (Table 1). The 2,000-character floor removes empty or failed transcriptions; it is a minimum, not a tuned threshold.

The measures of the text do not all use the same authors. The embeddings use all 178. The tournament uses 176 of them, since 2 Catholics are dropped as recognised authors (Section 3.3);[5] we call these 176 the pre-1517 sample. The texts printed from 1517 to 1530 give a second sample of 224 authors, 64 of whom later broke with Rome; a convert's text counts only to the year before his documented break. Luther is not among them: his own texts serve as the reference of Section 4.3.

Selection into the sample is substantial. Entry depends on a work surviving, being digitised and transcribing legibly, and on the reference works recording the author's later

[5]The two measures differ here for a reason. An embedding is a numerical representation of the text itself: whether a model could recall the author has no bearing on the vector his text receives, so every author with text is kept. A language model that judges a passage is different: it may base its verdict on what it knows, or believes it knows, about the author. We therefore drop the passages whose author it names.

Table 1. From the catalogue to the two samples

Step	Authors	Stayed Catholic	Later converts
Main authors of a VD16 edition printed 1501–1530	2,118		
… alive in 1517	1,314		
… not a worldly ruler	1,268		
… active in the German lands	1,100		
… with a documented later side	496	226	270
… with at least 2,000 characters of text available	301	215	86
Of these, with text printed before 1517	178	130	48
Pre-1517 sample (recognised authors dropped)	176	128	48
Of these, with text printed 1517–30, converts before their break	227	161	66
1517–30 sample (recognised authors dropped; without Luther)	224	160	64

Notes: The frame is every main author of a VD16 edition printed 1501–1530 (2,118 authors). The steps are applied in this order, each a subset of the one before: alive in 1517; not a worldly ruler; active in the German lands; a documented later side, stayed Catholic or broke with Rome (Protestant, or switching), established from reference works and documented acts; and at least 2,000 characters of transcribed text available in our corpus. The two lower blocks count, of these 301 authors, those with at least 2,000 characters of text available from before 1517 and those with at least 2,000 characters available from 1517–30, a convert's text counted up to the year before his documented break. Recognised authors dropped: a passage leaves the set when the model names its author at confidence 90 or more, and an author with no passage left leaves the sample; 2 authors leave the pre-1517 sample and 3 the 1517–30 sample, Luther among them, whose texts are the reference of Section 4.3.

side. It also depends on having been in print before 1517, which most later converts were not: of the 496 authors with a documented side, 54.4 percent later broke with Rome; of the 176 authors of the pre-1517 sample, 27.3 percent. The corpus is therefore the observable part of the pre-1517 writing population. It is not a census.

Finally, we sample the text rather than read every page. We transcribe a sample of pages, as a rule up to 200 per author: a short work in full, a long one in blocks of consecutive pages spread through the book. From an author's transcribed text, before 1517 and in 1517 to 1530 separately, one seeded draw then takes up to ten chunks of up to 18,000 characters, going round his works in turn. Of the measures of Section 3.3, the stance measure reads these chunks, the tournament reads passages of 400 words drawn from the text, and the embeddings read all of it. Sampling keeps a single long book from dominating an author's profile, at the cost that a position stated only on an unsampled page can be missed—one reason a text may be coded zero on a question it never had the chance to address. Appendix A describes the frame, the transcription and the sampling in detail.

3.2 Later Confessional Alignment

We code each author's eventual side so that it cannot have been read off the very texts we use to predict it. We take the outcome from biography, never from the writing: it records a later public alignment, not an enduring private "type." Where the De Gruyter *Verfasser-Datenbank* online (VDBO)—whose *Deutscher Humanismus 1480–1520* author lexicon was built for exactly this population—supplies a classification, it takes priority (De Gruyter, 2026).[6] Otherwise we use a documented act or the best available biographical judgement, drawing on Deutsche Biographie (Deutsche Biographie, 2026).[7] The acts are clerical marriage, an evangelical sermon, subscription to the Augsburg Confession, flight from or forfeiture of a post when a territory changed religion, or heresy proceedings. An author who broke with Rome and later returned counts as a convert. Of the 176 labels of the pre-1517 sample, 171 rest on a reference work, or on a database entry confirmed by a second source or by a dated break; 5 do not.

Pre-1517 text never enters the outcome. Even so, a biographer summarising an author's intellectual life may lean on the same writings we read, so the label is not a perfect firewall against circularity. The language model that reads the texts is never shown a label, and every text behind the prediction of Section 4.1 was printed before 1517, before any of these authors broke.

The timing of the break matters as well as the side. A convert's texts of 1517 to 1530 count only to the year before his break, the hazard model of Section 4.1 asks when the break came, and Section 4.3 follows each convert through his own texts toward it. Using grounded biographical search we record the earliest documented breach with Rome, whether an evangelical sermon, a marriage, or another public act; the reviews of the labels correct these dates and add dates where the search found none. A date is known for 40 of the 48 converts of the pre-1517 sample, with a median of 1522 (Table 2), and for every convert of the sample of 1517 to 1530, since his text is cut at it. A convert without a date stays in the pre-1517 sample and is left out wherever the date is needed. Appendix A gives the sources behind the labels and the dates.

[6] VDBO aggregates several standard German biographical author lexicons; the component relevant to our window is the *Deutscher Humanismus 1480–1520* Verfasserlexikon (ed. F. J. Worstbrock), alongside the *Verfasserlexikon des deutschen Mittelalters* and the early-modern (*Frühe Neuzeit*) and *Killy* lexicons.

[7] *Deutsche Biographie* compiles the *Neue Deutsche Biographie* (NDB) and the older *Allgemeine Deutsche Biographie* (ADB), which we therefore cite through it rather than separately.

3.3 Measuring Textual Orientation

A model that reads these texts should judge the text and not recognise its author. Before any model reads a text, we therefore replace every name of a person, a place and an institution in it by a tag: the author's own name, every person alive after about 1450, towns, territories, dioceses and universities. Biblical and classical names, saints, the Church Fathers, the scholastic and humanist authorities, Rome and the names of nations stay. Three measures of the masked text follow.

The first measure turns the text into numbers. An embedding model makes each author's text a vector, and a classifier fitted on the other authors ranks him, so that no author is scored by a model that saw his label.

The second measure is a tournament. To read the position a text argues rather than the words it uses, we ask a language model (gpt-6.1-sol) to compare: it sees two masked passages of 400 words by different authors and is asked which of the two leans further toward reform. The prompt defines reform broadly. It lists the reform positions on doctrine, from justification by faith and Scripture over tradition to the rejection of indulgences, vows and the cult of the saints, and the dispositions that went with them: humanist over scholastic method, inward Christ-centred devotion over external observance, attacks on clerical abuses and calls for reform. A passage that takes the traditional side leans less toward reform than one that says nothing. The model is shown no name and no label. A Bradley–Terry model turns its verdicts into one scale, and an author's score is the mean strength of his passages (Section 4.1). We call it the score.

The third measure puts questions to the text. The same model codes every chunk on twelve axes: eight on doctrine, with the two ends of each scale taken from what the parties later put in writing, above all the Augsburg Confession and the Confutatio of 1530, and four on outlook, devotion, grievance and the demand for reform, on which authors of both later sides took positions before 1517. A code runs from -2, the Roman position, to $+2$, the reform position, with 0 when a chunk takes no side or does not address the question. The prompt asks for the position argued, not the subject mentioned: a defence and a criticism of indulgences receive opposite signs, though both are about indulgences. The model also returns a single overall lean on the same scale: a judgement of the chunk as a whole, not the average of the twelve axes. We call this third measure the stance measure. Section 4.2 uses the axes to take the score apart, and Section 4.3 follows the axes and the lean through the 1520s.

Masking does not rule out that a language model knows a passage. We therefore put every passage to the model with the question who wrote it. A passage whose author

it names at confidence 90 or more is dropped, and another passage of the same author takes its place unless the model names that one too; the stance measure leaves the chunks concerned out. An author with no passage left leaves the sample (Table 1). Section 5.1 tests what this leaves. Appendix A gives the details of the masking and of the dropping of recognised authors, and Appendix B the twelve axes and the full prompts.

3.4 Biographical Covariates

We collect education, career, office and family variables for each author from VDBO and Deutsche Biographie, with Wikipedia, Germania Sacra, the Repertorium Academicum Germanicum and Wikidata where these are silent, and keep an exact source quotation behind each value (see Appendix A for the sources and the extraction). Degree, teaching and office are cut at 1517, so that a later Protestant appointment cannot enter as a pre-Reformation covariate. The coding covers 170 of the 176 authors of the pre-1517 sample.

The items we use are a doctorate by 1517, teaching at a university before 1517, the clerical estate in 1517 (priest, member of an order, or in clerical office), a benefice, the father's standing, and the place of activity. A benefice counts only as documented personal possession before 1517 of a prebend, canonry, endowed parish or comparable endowed office. The father's standing sets the urban elite, office-holders and nobles against the rest; it is the thinnest item, coded for 26 converts and 49 Catholics, because the sources name no father for 92 of the 170. An item counts when a source records it, and silence counts as no: "no benefice" means no such holding found in the sources, not proof that none existed. The clerical estate and the benefice measure an author's stakes in the church of 1517: a cleric or a benefice holder had an income and a standing to lose.

3.5 Descriptives

Table 2 describes the pre-1517 sample by later side, and two facts preview the argument. The later converts were, on average, younger—median birth year 1482 versus 1470—and wrote a smaller share of their work in Latin: both consistent with the humanist, vernacular-friendly milieu. They also had less in print, 2 works at the median against 3, and less transcribed text, so age, language and text length enter the analysis as controls (Table 3).

The later converts were also less established in the church and the university of 1517. A doctorate is recorded for 26 percent of them against 43 percent of the Catholics, a benefice

Table 2. Authors by later side

	Later converts	*N*	Stayed Catholic	*N*
Authors	48		128	
Birth year, median	1482	46	1470	112
Year of the documented break, median	1522	40	—	
Mean print year of the works, mean	1510	48	1509	128
Works in the corpus, median	2	48	3	128
Log characters of text (std)	−0.37	48	0.14	128
Latin share of the works (std)	−0.23	47	0.09	127
Doctorate by 1517	26 %	46	43 %	124
Taught at a university before 1517	48 %	46	59 %	124
Benefice documented before 1517	17 %	46	38 %	124
Cleric in 1517	30 %	46	60 %	124
Father of elite standing	50 %	26	67 %	49
Famous authors	19 %	48	15 %	128
Recognised authors	34 %	47	45 %	128
Distance of the imprint town to Wittenberg (std)	−0.19	43	0.07	117
Tournament score, mean (+ reform, − Catholic)	0.80	48	−0.55	128

Notes: Sample: the 176 authors of the pre-1517 sample (48 later converts). Each row uses the authors for whom the item is known; the *N* columns give their number by side. Works, print years, characters and Latin share refer to the author's works printed before 1517 that are in the corpus; characters count the transcribed text without spaces, and the Latin share is the share of works the catalogue records as Latin. (std): the variable is standardised over the authors of the sample (mean 0, standard deviation 1), and the row gives each side's mean. The biography items come from a coding of each author's position in 1517 from reference works, which covers 170 of the 176 authors; an item counts when a source records it, and silence counts as no. Cleric: priest, in an order, or in clerical office. Father of elite standing: urban elite, office-holder or noble; the sources name no father for 92 of the 170 authors and code no standing for 95. Famous authors: on a hand-curated list of 46 well-known figures. Recognised authors: shown up to two masked chunks of the author's text, the model names him for at least one. Distance is from the imprint town of most of the author's works to Wittenberg. The tournament score is the mean strength of the author's passages (Section 4.1); positive values lean toward reform.

for 17 against 38 percent, and the clerical estate for 30 against 60 percent. Section 5.2 holds these differences against the score.

We measure geography by the town where most of an author's works were printed, the supply-side geography of the press. The later converts' works were printed somewhat closer to Wittenberg on average (Table 2), and the distance enters the analysis as a control.

4 Empirical Analysis

4.1 Predicting Conversion from Pre-1517 Text

Luther's own account of the movement's spread was purely textual: "I simply taught, preached, wrote God's Word; otherwise I did nothing… I did nothing. The Word did it

all."[8] If the Word did it, the words should be observable—before conversion as well as after. We therefore ask whether an author's eventual break with Rome can be predicted out of sample from his pre-1517 text: the question is whether the information was there to be read before the event.

We rank the authors in two ways from their pre-1517 text, with every passage stripped of its author's name and of the names of contemporaries and places (Section 3). The first measure turns each author's text into an embedding, and a logistic model turns the embedding into a predicted probability of a later break, always fitted on other authors. The second is a tournament in which a language model is asked which of two passages leans further toward reform; it is never shown a label. We score both by the area under the ROC curve (AUC): the chance that a random later convert is ranked above a random Catholic. A ranking without information scores 0.5; a ranking that puts every convert above every Catholic scores 1.[9]

Embeddings. The measure is model-free. Authors, never individual documents, are held out together (Appendix A). The embeddings reach an AUC of 0.645 (95 percent confidence interval 0.554 to 0.736; Figure 1).

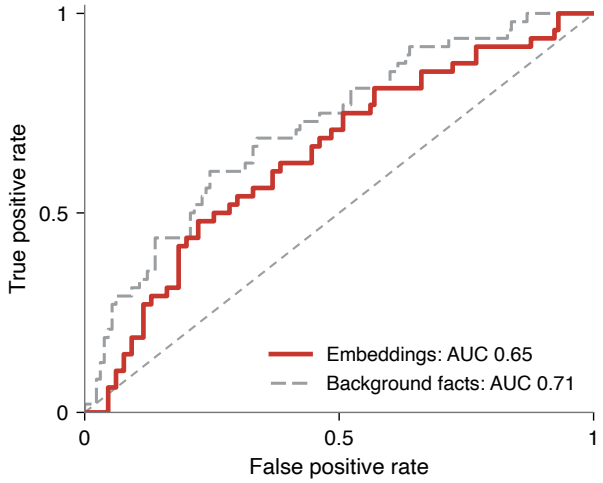
The text alone predicts almost as well as the background facts about an author: his year of birth, the mean print year of his works, their Latin share and length, and the region where most of them were printed. Fitted the same way, these facts give the second curve of Figure 1. On the 176 authors of Table C1 the embeddings reach 0.652, the facts 0.711 and the two together 0.736; measured from chance, the embeddings account for 65 percent of what the two reach together.

Tournament. The second measure rests on an old insight about judgement: it is easier to say which of two things is more than to say how much either is (Thurstone, 1927). No yardstick existed in 1510 for how far a text leaned toward reform, and a model asked to grade a scholastic commentary and a vernacular Passion song on one scale must invent one. Asked instead which of two passages leans further toward reform, the model needs no scale, only a comparison. Comparisons of this kind are used to measure positions in political texts (Carlson and Montgomery, 2017; Benoit et al., 2019), and language models rank more reliably when they compare two texts than when they score one (Qin et al.,

[8] Second Invocavit Sermon, Wittenberg, March 10, 1522 (Doberstein, 1959, pp. 77–78); WA 10/3:18–19.

[9] The embeddings use the full sample of 178 authors, 48 of whom later broke with Rome; the tournament uses 176 of them, since 2 Catholics are dropped as recognised authors (Section 3.1).

Figure 1. ROC curves before 1517



Notes: Receiver operating characteristic curves on the full pre-1517 sample of 178 authors (48 later converts). Each curve ranks the authors by a score and plots, for every cut, the share of later converts above the cut (true positive rate) against the share of Catholics above it (false positive rate); the dashed diagonal is a ranking without information. Embeddings: each author's masked pre-1517 text is embedded (text-embedding-3-large, vectors of width 1,024), reduced to 20 principal components and entered into a class-balanced logistic regression; the model is fitted in five folds split by author and repeated 100 times, and an author's score is his mean rank out of fold, so no author is scored by a model that saw his label. Background facts: the same procedure on what the catalogue and reference works record without reading a page: year of birth (with its square and a flag when missing), mean print year of the author's works, Latin share, log characters of text and eight indicators of the main print region. AUC, the area under the curve, is the probability that a randomly drawn later convert ranks above a randomly drawn Catholic; the 95 percent confidence intervals (DeLong) are 0.55 to 0.74 and 0.63 to 0.80. Facts and embeddings together reach 0.74 (Appendix Table C1).

2024).[10]

To fix ideas, consider two Latin books printed at Cologne: Hermann von dem Busche's sermon to the clergy at the synod of 1509 (VD16 B 9946) and Jacobus Magdalius's exposition of Christ's Passion of 1504 (VD16 B 4682). The model sees a masked passage of each, without the names, and is asked which of the two leans further toward reform and to give its reason in one sentence. It answers twice, with the passages in either order, and picks Busche's both times. Its reason: Busche "contrasts Scripture's commands with contemporary conduct and attacks avarice and trust in earthly treasure, whereas [Magdalius] affirms Peter's power to bind and absolve within an authority-based scholastic exposition." Busche later broke with Rome; Magdalius stayed Catholic.

The contrast is not always this plain. Johannes Oecolampadius's sermons on the Pas-

[10] A language model asked to score a single text has to calibrate an absolute scale, and its scores tend to bunch on a few values. Asked which of two texts ranks higher, it makes only a relative judgement, and rankings built from such comparisons are more accurate (Qin et al., 2024; Liusie et al., 2024). The model gives its reason with every verdict; the reasons quoted below are these. The same model, asked for an absolute lean on a five-step scale, separates the sides with an AUC of 0.59 (Table C1).

sion (Strasbourg 1512; VD16 O 304), devotional preaching by the later reformer of Basel, meet Hieronymus Dungersheim's epitome of the four books of the Sentences (Leipzig 1514; VD16 D 2950), a textbook of scholastic theology. Asked to choose, the model still decides, in both orders and for a finer reason: Oecolampadius "commends inward meditation on Christ's passion and polished letters, whereas [Dungersheim] explicitly makes baptismal efficacy for heretics depend on Catholic unity and ordinarily reserves baptism to priests." Oecolampadius later broke with Rome; Dungersheim stayed Catholic.

A ranking of all passages and authors needs many meetings that overlap: if one passage is set against a second and the second against a third, the three can be placed on one scale. Each passage, 400 words drawn from what an author printed before 1517, therefore meets 8 passages by other authors printed within five years of it, whatever their language, their subject or their later side.[11] These meetings link every passage to every other through a chain of comparisons.

A Bradley–Terry model (Bradley and Terry, 1952; Davidson, 1970) turns the verdicts into that scale. A passage wins a meeting only when the model picks it in both orders (Prompt B1); a split verdict is a tie. Each passage i has a strength θ_i, and the further its strength lies above that of the passage it meets, the more likely it is to win. With $\pi_i = e^{\theta_i}$,

$$P(i \text{ beats } j) = \frac{\pi_i}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}}, \qquad P(\text{tie}) = \frac{\nu\sqrt{\pi_i \pi_j}}{\pi_i + \pi_j + \nu\sqrt{\pi_i \pi_j}}, \tag{1}$$

where $\nu \geq 0$ governs how often a meeting ends in a tie. The strengths are estimated by maximum likelihood: they are the values under which the model's verdicts are most likely, and nothing else enters. A win over a strong passage raises a passage's strength by more than a win over a weak one, so the scale reflects whom a passage beat and not only how often it won. An author's score is the mean strength of his passages:

$$\text{score}_a = \frac{1}{n_a} \sum_{i \in a} \theta_i. \tag{2}$$

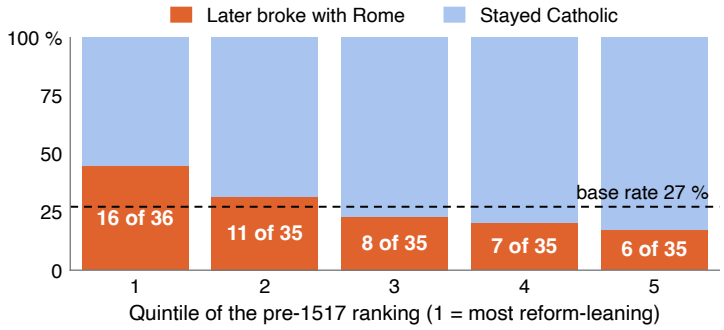
A gap of 1 between two passages means that the model calls the first more reform-leaning in 73.1 percent of the meetings it decides.

[11] A stricter comparison sets a convert's work against a Catholic's work in the same language, of about the same year and by an author of about the same age, on up to four passages of each (Appendix A gives the limits). Of the 122 pairs, 7 end in a tie; in the others the convert's work is ahead in 62 percent (p 0.015; Table C7). Such pairs share no work and cannot be placed on one scale. On the same pairs the tournament score puts the convert above the Catholic in 75 percent (p 0.007).

The tournament reaches an AUC of 0.633 (95 percent confidence interval 0.541 to 0.725), close to the embeddings' 0.645. The two measures share nothing but the masked text: the embeddings are fitted on other authors' labels, the tournament never sees one. Their rankings of the authors correlate at 0.52 (Spearman; Figure C1).

Ranked by the score and cut into quintiles, 16 of the 36 authors in the most reform-leaning quintile broke with Rome, and 6 of the 35 in the most Roman; in the sample as a whole 27.3 percent did (Figure 2).

Figure 2. Converts by quintile of the pre-1517 score



Notes: The 176 authors of the pre-1517 sample are ranked by the tournament score of Equation (2) and cut into quintiles, the most reform-leaning first (36, 35, 35, 35 and 35 authors). Each bar is the share of the quintile who later broke with Rome; the label gives the count. The line marks the share in the whole sample, 48 of 176. The AUC of the score is 0.633 (Appendix Table C1).

The model picks the same passage in both orders in 0.93 of the meetings, and the split-half reliability of the score is 0.89 (Table C4). No single author carries the AUC (Figure C2), and reshuffling the labels across authors reaches the tournament's AUC in 2.6 of every 1,000 draws (Table C1). The model's language tilt and the works at the two ends of the scale are in Appendix C.2.

Next, we estimate a linear probability model. For author a,

$$\text{broke}_a = \alpha + \beta\,\text{score}_a + X_a'\gamma + \varepsilon_a, \tag{3}$$

where broke_a is one if the author later broke with Rome and zero if he stayed Catholic, and X_a collects the controls added one block at a time in columns (2) to (6) of Table 3. The score is standardised over the 176 authors. The coefficient in column (1) implies that a one-standard-deviation higher tournament score is associated with an 8.7 percentage point higher probability of a later break, over a base of 27.3 percent (statistically significant at the 1 percent level).

Table 3. Break with Rome and the pre-1517 score

	Dependent variable: later broke with Rome (1 = yes)					
	(1)	(2)	(3)	(4)	(5)	(6)
Tournament score (std)	0.087	0.100	0.075	0.088	0.089	0.112
	(0.031)	(0.031)	(0.034)	(0.036)	(0.036)	(0.034)
Latin share of the text (std)		−0.065	−0.095	−0.087	−0.090	−0.110
		(0.035)	(0.036)	(0.040)	(0.040)	(0.040)
Log characters of text (std)		−0.114	−0.115	−0.109	−0.118	−0.095
		(0.033)	(0.035)	(0.040)	(0.041)	(0.039)
Year of birth			0.006	0.006	0.006	0.005
			(0.002)	(0.002)	(0.002)	(0.002)
Mean print year			0.009	0.009	0.009	0.012
			(0.009)	(0.009)	(0.009)	(0.009)
Distance to Wittenberg (std)				−0.066	−0.072	−0.069
				(0.038)	(0.039)	(0.037)
Famous authors					0.122	
					(0.091)	
Recognised authors						−0.166
						(0.070)
Authors	176	174	156	141	141	140
Later converts	48	47	45	40	40	39
Mean of dep. var.	0.273	0.270	0.288	0.284	0.284	0.279
R^2	0.038	0.121	0.197	0.203	0.213	0.239

Notes: Estimates of Equation (3), a linear probability model, by OLS on the pre-1517 sample; the controls X_a are the variables in the rows below the score. The dependent variable equals one if the author later broke with Rome and zero if he stayed Catholic (mean in the last rows). The tournament score is the mean strength of the author's passages (Equations (1) and (2)). (std): the variable is standardised over the 176 authors of the sample (mean 0, standard deviation 1), so its coefficient is the change in the probability of a later break per standard deviation. Latin share: share of the author's pre-1517 works that the catalogue records as Latin. Log characters: logarithm of the characters of his transcribed pre-1517 text. Year of birth and the mean print year of his pre-1517 works enter in years. Distance is from the imprint town of most of his works to Wittenberg. Famous authors: on a hand-curated list of 46 well-known figures. Recognised authors: shown up to two masked chunks of the author's text, the model names him for at least one. Columns (5) and (6) add each of the two to column (4). Columns (2) to (6) keep the authors for whom the added controls are known, which is why the number of authors falls. Heteroskedasticity-robust (HC1) standard errors in parentheses.

Later converts and Catholics differ in observable characteristics such as age, language and the amount they wrote (Table 2). Including controls for the Latin share and the length of the text, year of birth, the mean print year and the distance from the imprint town to Wittenberg leaves the coefficient at 0.088 in column (4), statistically significant at the 5 percent level. It also remains so when we add whether an author is on a list of well-known figures (column 5, 0.089) or whether the model, shown masked passages of his text, can name him (column 6, 0.112; Appendix A). These are conditional associations, not causal effects.

Table 3 asks who broke, not when, and it counts an author who died in 1520 as a Catholic like one who lived through the 1520s and stayed one. A hazard model asks when the break came (Table C3). An author is at risk from 1517 until he breaks, dies

or reaches 1530; a death ends his time at risk and does not count as a decision to stay. A one-standard-deviation higher score raises the yearly odds of a break by 47 percent (permutation p 0.025), and by 43 percent with year of birth and years in print held fixed (p 0.042). The score orders the authors by degree and not only by side: the further an author's pre-1517 text leans toward reform, the earlier he breaks.

4.2 Unpacking the Black Box: Genre, Outlook, Doctrine, and Reform

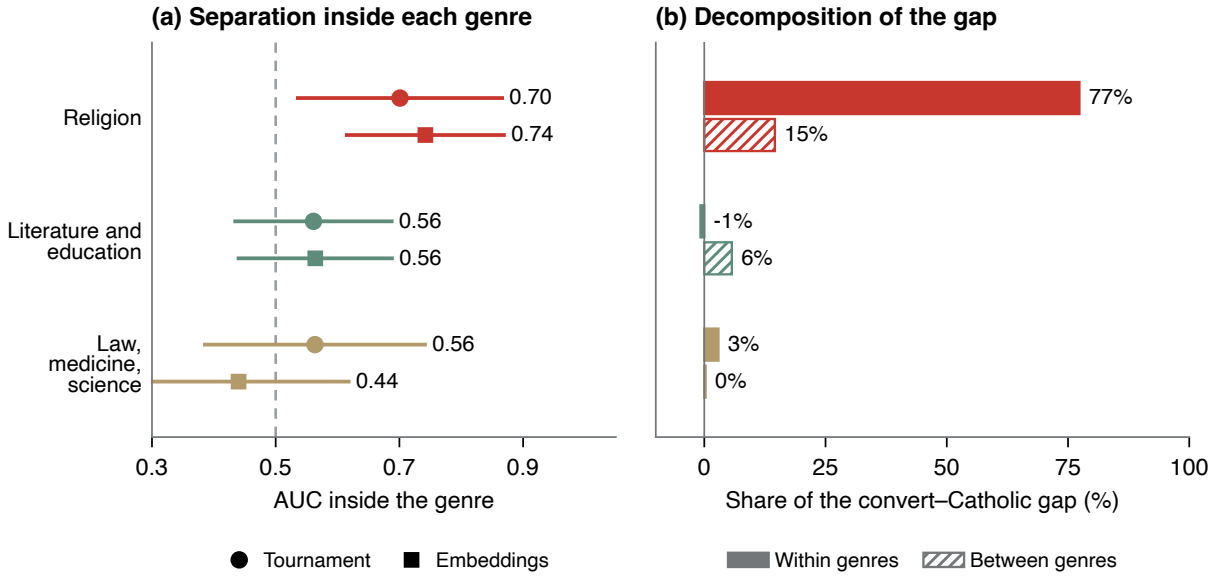
Prediction is the first step of the analysis (Ludwig and Mullainathan, 2024). The score of Section 4.1 ranks the authors, but it does not say what in their text it rewards: a confessional programme already written down, or an orientation older than any confession. We consider four candidates: an author's intellectual outlook and the cast of his devotion; his stance on doctrine; his grievance against the institution; and his demand that the church be reformed. Rival accounts rest on different candidates. If the trace is grievance over fees, indulgences and the wealth of prelates, the Reformation reads as a revolt against an extractive church. If it is the demand for reform, it reads as a movement already formed and waiting for its moment. If it is outlook and devotion, it reads as a divide in learning and piety that the quarrel later cut along.

Genre. We first decompose the score by genre: the two sides may differ in what they wrote about, or in how they wrote within the same genre. A fixed prompt reads each work's title and two extracts of its masked text and assigns one subject code (Prompt B4), and the codes form three genres: religion, from theology and sermons to canon law and religious controversy; literature and education, from grammars and school texts to humanist poetry and history; and law, medicine and science. An author's score in a genre is the mean strength of his passages from works of that genre.

The separation sits in religious writing. On religious text alone the score ranks a later convert above a Catholic with an AUC of 0.70 (95 percent confidence interval 0.54 to 0.87; Figure 3, panel (a)). In literature and education and in law, medicine and science it reaches 0.56: only a weak signal, with confidence intervals that include one half, so that we cannot reject pure chance. The embeddings, fitted on each genre's text alone, show the same pattern.

The choice of genre accounts for little of the gap between the sides. The Catholics wrote more on religion, 37 percent of their text against 25 percent of the converts', and religious text scores nearer Rome. We split the gap between the converts' and the Catholics'

Figure 3. Separation by genre



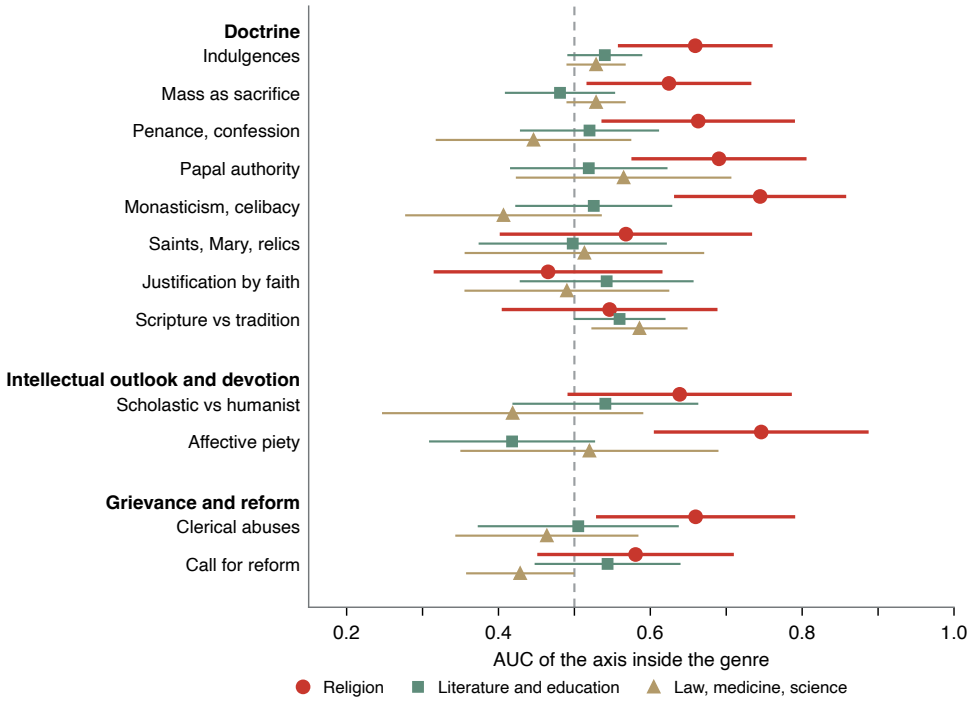
Notes: The 176 authors of the pre-1517 sample (48 later converts) and their works printed before 1517. Each work carries one subject code, read by a fixed prompt from its title and two extracts of its masked text (Prompt B4). The codes form three genres. Religion: theology and Bible commentary; sermons; devotion, prayer and saints' lives; church administration (synodal statutes, indulgence instructions); canon law; religious controversy (the dispute over Jewish books, witchcraft, the defence of clerical privilege). Literature and education: grammar, rhetoric and school texts; philosophy and logic; humanist poetry, orations, satire and drama; editions of classical authors; history, chronicles, news and political verse. Law, medicine, science: civil law and procedure; government ordinances; medicine and plague tracts; astronomy, calendars, prognostications and cosmography. Colours mark the genre, as in Figure 4. Panel (a): AUC of each author's mean score over his text in the genre, with 95 percent DeLong intervals; circles the tournament score (mean strength of his passages in the genre); the tournament counts 81 authors (16 later converts) in religion, 102 (30) in literature and education and 50 (14) in law, medicine, science; squares the embeddings, the classifier of Figure 1 fitted on the genre's text alone, on the authors of the sample with 2,000 characters in works of the genre (82, 102 and 51 authors). Panel (b): the gap between the converts' and the Catholics' average tournament score (1.35 strength units) split into two parts per genre, which add up to the gap; solid bars the within part, hatched bars the between part. Between genres: the difference between the two sides' share of text in the genre, times the distance of the genre's mean score from the mean over genres, each genre weighted by its average share of text. Within genres: the difference between the two sides' mean score in the genre, times the genre's share of all text (Appendix C.5 gives the equations). The gap is 1.35 (95 percent interval 0.41 to 2.25, from 1,000 resamples of authors within each side); the within part of religion is 1.05 (0.29 to 1.74), its between part 0.20 (−0.01 to 0.48); every other part lies within 0.08 of zero, with an interval that spans it.

mean score into a part that comes from which genres the authors wrote in and a part that comes from how each side wrote within a genre (Kitagawa, 1955; Oaxaca, 1973; Blinder, 1973); Appendix C.5 gives the equations. The Catholics' larger share of religious writing accounts for 15 percent of the gap. How the two sides wrote about religion accounts for 77 percent (panel (b)).

Stance. We next ask on which questions the sides differ in their religious writing. The stance measure of Section 3.3 puts twelve questions to every chunk of text, its twelve axes, and codes each from −2, the Roman position, to +2, the reform position, with 0 when

the chunk takes no side or does not address the question (Prompt B2). Eight axes are on doctrine: justification, indulgences, papal authority, Scripture and tradition, the mass as sacrifice, monasticism and celibacy, the saints, and penance. These are the articles on which the Augsburg Confession and the Confutatio of 1530 contradict each other, with indulgences, which neither treats, anchored in Luther's Ninety-five Theses and the papal reply; each end of the scale is thus a position one party put in writing (Appendix Table B1). The other four axes cover the remaining candidates, on which authors of both later sides took positions before 1517. Scholastic against humanist method and affective piety measure outlook and devotion, clerical abuses measures grievance, and the call for reform measures the demand that the church be reformed. An author's stance on an axis within a genre is his mean code over his chunks of that genre.

Figure 4. Twelve axes by genre



Notes: Stance chunks printed before 1517 by the authors of the pre-1517 sample, sorted by the subject code of their work into the three genres of Figure 3: religion (80 authors, 16 later converts), literature and education (100, 29) and law, medicine, science (48, 13). The model codes each chunk on twelve axes from −2, the Roman position, to +2, the reform position, with 0 when the chunk takes no side or does not address the question (Prompt B2). An author's position on an axis inside a genre is his mean code over his chunks of that genre; the dots show the AUC of that position, with 95 percent DeLong intervals. The eight doctrine axes are the articles on which the Augsburg Confession and the Confutatio of 1530 contradict each other (Appendix Table B1).

On religious text the axes that separate the sides are the old quarrels (Figure 4). Monasticism and celibacy separates them with an AUC of 0.75, and papal authority,

penance and indulgences follow. Outlook also separates: affective piety at 0.75 and scholastic against humanist method at 0.64. Grievance over clerical abuses separates too, the call for reform less. At the bottom stand the two positions that would define the movement, Scripture against tradition and justification by faith (0.47). In the other two genres no axis separates the sides in this way (Appendix C.5).

Much of the predictive power of these axes comes from the Catholics, who defend doctrine more firmly, whereas the later converts take no pronounced stance. On indulgences, the mass, penance and the pope, between 25 and 41 of the 64 Catholics with religious text take the Roman side; most of the 16 later converts do not address these questions at all, and none takes the reform side. Over all text the same holds: before 1517 both sides' mean code lies below zero on every one of the eight doctrine axes, and the Catholics' lies further below, -0.24 against -0.13 over the eight (Figure 7, panel (a), in Section 4.3). The model's own reasons say the same (Appendix Table C13): what loses a meeting is the defence of saints, school, monastic life, pope and indulgences; what wins is humanist learning and inward piety.

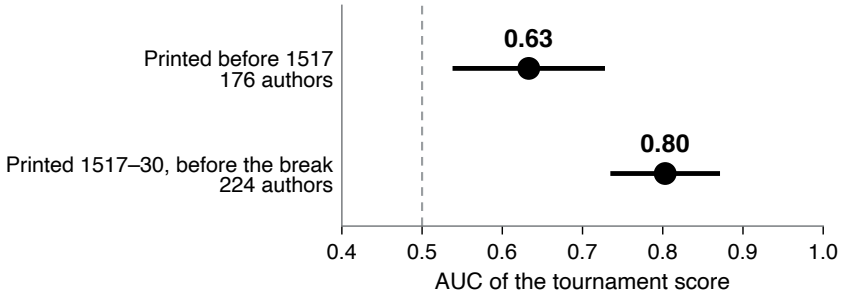
The clearest markers of the predictive signal are the quarrels of the fifteenth century (Section 2), over vows and the religious life, indulgences and penance, and the pope's power, on which the Catholics defend and the later converts mostly stay silent, and with them intellectual outlook and inward devotion. Grievance separates less: the gap between the two sides' mean codes is 0.08 on clerical abuses and -0.02 on the call for reform. The articles on which the movement would stand, justification by faith and Scripture against tradition, do not separate the sides. The score therefore picks up a broad intellectual outlook, spread across many questions and strong on none: how an author argued, how inward his piety was, and how readily he defended the existing church. This is close to Moeller's "without humanism, no Reformation" (Moeller, 1959), and to the view of Febvre (1929) and Ozment (1975) that religious disposition rather than grievance set the future convert apart. It also connects the result to the economics of humanist education and the print-borne circulation of humanist learning (de la Croix et al., 2025; Dittmar and Seabold, 2019).

4.3 Coordination

After 1517 the trace of an author's later side becomes much clearer. We read what was printed from 1517 to 1530, by these authors and by those who entered print only then, with the same model and the same twelve axes. Luther is not among the authors the model ranks; his own texts serve as a reference.

The tournament after 1517. We run the tournament of Section 4.1 again on the texts of 1517 to 1530, with the same model and the same prompt, and count a convert's text only to the year before his documented break, so that nothing he printed as a Protestant enters (Appendix A). The score now ranks a later convert above a Catholic with an AUC of 0.80 (95 percent confidence interval 0.74 to 0.87), against 0.63 before 1517 (Figure 5). The separation does not wait for the late 1520s: on the texts of 1517 to 1521 alone the AUC is 0.78. The embeddings, fitted on the texts of the same years, reach 0.70 (Appendix Table C8).[12]

Figure 5. Prediction before and after 1517



Notes: AUC of the tournament score before 1517 and in 1517–30. Before 1517: 176 authors (48 later converts), 2,348 passages, 9,392 meetings. In 1517–30: the texts printed from 1517 to the year before the author's own documented break, 224 authors (64 later converts), 1,434 passages, 5,736 meetings; Luther's texts are the reference of Section 4.3 and not in the sample. Each period is a separate tournament: passages of about 400 words are drawn from the author's works (up to 15 per author before 1517, 10 in 1517–30, fewer when his text is short), every passage meets 8 others, each meeting is read in both orders by the same model with the same prompt (Prompt B1), and a Bradley–Terry model with ties (Section 4.1) gives each passage a strength; the author's score is the mean strength of his passages. The model sees no label and no name. Bars: 95 percent confidence intervals (DeLong). The shorter windows 1517–21 and 1517–25 and the embeddings are in Appendix Table C8.

In the regression of Equation (3) on these texts, on the 191 authors for whom every control is known, 54 of them later converts, a one-standard-deviation higher score is associated with a 22.4 percentage point higher probability of a later break (statistically significant at the 1 percent level). Before 1517 it was 8.7. Including controls for the Latin share and the length of the text, year of birth and the distance to Wittenberg leaves it at 20.5 (Appendix Table C8).[13]

[12] The stricter comparison of Section 4.1, a convert's work against a Catholic's work matched as there and on the genre of the work as well, gives the same picture. On the works of 1517 to 1530, the convert's printed before his break, the model puts the convert's work nearer reform in 77 percent of the decided pairs among 128; before 1517 it was 62 percent among 122 without the match on genre and 58 percent among 112 with the genre matched (Appendix Table C7).

[13] The mean print year is not among the controls here: a convert's text stops at his break, so his works are earlier by construction.

Event study. Next, we follow each author through his own texts. A convert's text may have moved toward reform in the years before his own break, or minds may have been made up in 1517. Take Johannes Oecolampadius, whose Passion sermons of 1512 appeared in Section 4.1. The model codes each chunk of text for its overall lean, from −2 (Roman) to +2 (reform) (Section 3.3; Prompt B2). Of the 9 chunks of the Passion sermons it places 5 on the reform side and 4 on the Roman side, 0.0 on average. Oecolampadius broke with Rome in 1522. Of what he printed from 1517 to 1521, his own tracts and sermons and his translations of Gregory of Nazianzus, it places all 10 chunks on the reform side, 1.5 on average.

An event study asks whether this holds for the converts as a group. It takes each author's own texts, those printed after 1517 included, and regresses the lean of a chunk on an effect for its author, an effect for its calendar period and indicators of the years between its print year and its author's break:

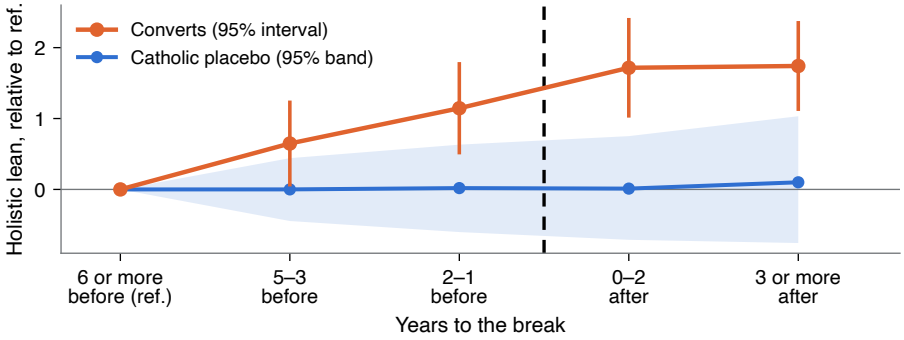
$$\text{lean}_{ca} = \alpha_a + \gamma_{t(c)} + \sum_k \beta_k \, \mathbf{1}\Big[\text{years to the break of } a \in k\Big] + \varepsilon_{ca}. \tag{4}$$

The bins k are three to five years before the break, one to two before, the break year and the two after, and three or more after; six or more years before is the reference. The author effect α_a compares each author with his own earlier text, and the period effect γ_t takes out what moved everyone, converts and Catholics alike, in the same years. A Catholic has no break, so all his chunks sit in the reference and help identify the period effects (Appendix A). If the model scored the author and not the text, every chunk of an author would carry the same lean, the author effect would absorb it, and the coefficients β_k would be zero.

A convert's slant increases before his break. It stands 0.65 above his own reference level three to five years before the break and 1.15 above that level in the two years before (Figure 6). Before 1517 the two sides' mean leans were −0.07 for the later converts and −0.40 for the Catholics (Figure 8, panel (a)); the rise in the two years before the break is larger than that gap.

Dates alone do not produce this: when each Catholic is given a break year drawn at random from the converts' break years, 18 of the 5,000 draws reach the converts' coefficient three to five years before the break, 0 in the two years before it and 0 in the break year and the two after (Table C9). The move toward reform precedes the documented act by years, within the author and net of what moved everyone in the same years.

Figure 6. Event study around the break



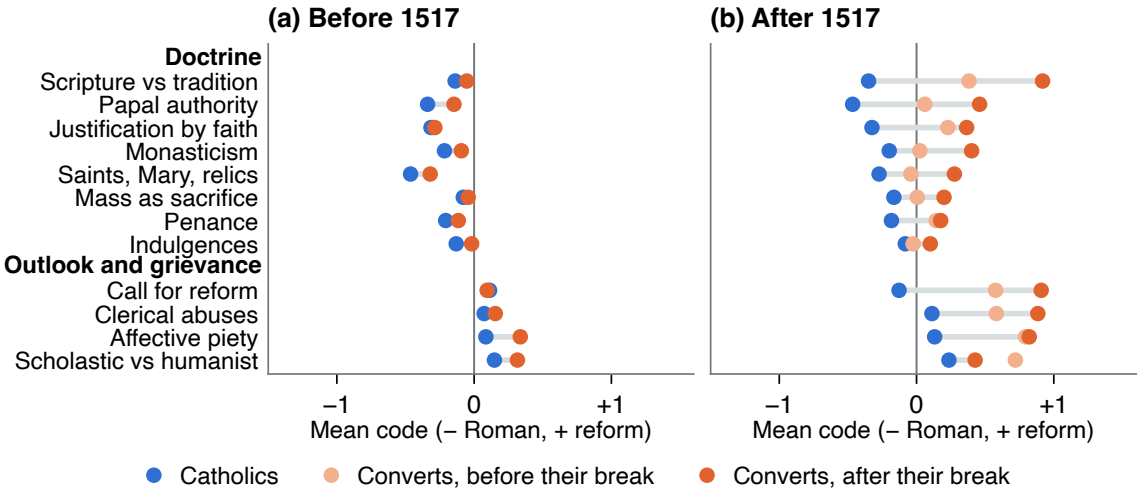
Notes: Estimates of Equation (4): the coefficients β_k on four indicators of the years between the chunk's print year and the convert's documented break (5 to 3 years before, 2 to 1 before, the break year to 2 after, 3 or more after; 6 or more years before is the reference), in a regression of the overall lean of a chunk of text, coded by the model from −2 (Roman) to +2 (reform), on author fixed effects α_a, calendar-period fixed effects γ_t (before 1510, 1510–16, 1517–19, 1520–21, 1522–25, 1526–30) and the four indicators. Catholics have no event time and identify the period effects. Sample: every stance-read chunk left after recognised passages are dropped, with a lean and a print year, 2,342 chunks of 280 authors: the Catholics and the converts of the pre-1517 and the 1517–30 samples (Table 1) with such a chunk, a convert's chunks after his break included (Appendix A); Luther and the converts without a documented break year are left out, so 76 converts remain. Converts: the coefficients (0.65, 1.15, 1.72, 1.74); bars: 95 percent confidence intervals from standard errors clustered by author. Catholic placebo: every Catholic is given a break year drawn from the converts'; the line is the mean coefficient over 5,000 draws, the band their 2.5th to 97.5th percentiles. The dashed line marks the break. Appendix Table C9 gives the coefficients, the placebo and the same model on the 17 converts with text both before 1517 and after their break.

Stance before and after 1517. Figure 7 gives each side's mean code on the twelve axes, before 1517 and after it, with the converts' text split at each convert's own break. Before 1517 the differences are small. The largest gap between the two sides on any axis is 0.25, and on the doctrine axes both sides are neutral or on the Roman side. In what the converts printed from 1517 to the year before their break, a programme appears: their mean code is now on the reform side on Scripture against tradition (0.38) and on justification by faith (0.23), while the Catholics hold the Roman side on every doctrine axis. After the break the picture is clearer still. Over the eight doctrine axes the mean gap between the sides is 0.10 before 1517, 0.35 before the break and 0.62 after it.

Lean, correlation and Luther's profile. Figure 8 follows the two sides through five periods, from before 1517 to 1526 to 1530, on all of a side's text (Appendix A). Panel (a) gives each side's mean lean. The two sides start close together, at −0.07 for the later converts and −0.40 for the Catholics. By the 1520s they stand far apart: the converts at 1.43 in 1522 to 1525, the Catholics at −0.78 in 1526 to 1530.

Panel (b) asks whether an author's stances hang together. For each side and period we correlate the authors' positions on every pair of the twelve axes and average over the pairs. Before 1517 an author's position on one question said little about his position on

Figure 7. Twelve axes before and after 1517



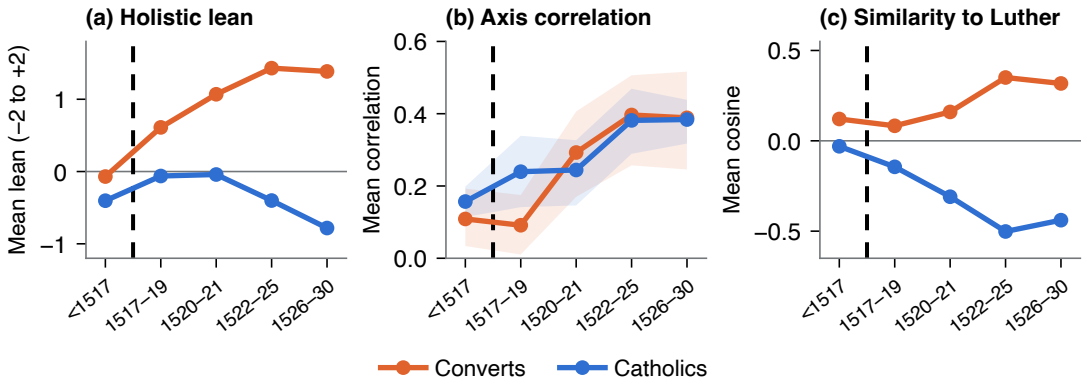
Notes: Mean code on each of the twelve stance axes, from −2 (Roman position) to +2 (reform position), with 0 when a chunk takes no position or does not address the question; averaged first over an author's chunks and then over the authors of a group. Catholics: texts printed before 1517 (128 authors) and 1517–30 (159). Converts: texts printed before 1517 (47), from 1517 to the year before their documented break (64), and after it (22). Chunks left after recognised passages are dropped; Luther's texts excluded. The eight doctrine axes and the four axes on outlook and grievance are defined in Appendix Table B1 and read with Prompt B2.

another: the mean correlation is 0.11 among the converts and 0.16 among the Catholics. By 1522 to 1525 it is 0.40 and 0.38. An author who takes the reform side on one question now takes it on others too, and one who defends on one defends on others: on both sides the bundle extends across more of the twelve axes. That widening carries the later rise: among the authors who take a position on both axes of a pair, the correlation does not keep rising through the 1520s (Appendix A). Appendix Figure C4 shows the correlations pair by pair.

Panel (c) asks whether this common line is Luther's. Luther's profile is his own mean code on each axis in 18 chunks of his writings of 1518 to 1521, read by the same model under the same prompt; it is highest on affective piety, the call for reform, clerical abuses, Scripture and justification. An author's profile is his mean code on each axis in a period, and we take the cosine between it and Luther's, each measured from the midpoint between the two sides. The converts' mean cosine rises from 0.12 before 1517 to 0.35 in 1522 to 1525; the Catholics' falls from −0.03 to −0.50. Author by author, agreement with Luther's profile separates the sides with an AUC of 0.64 before 1517 and of 0.87 in 1522 to 1525 (Appendix Figure C5).

Three checks show that Luther's profile is a natural yardstick. By the mid-1520s the

Figure 8. Lean, axis correlation and similarity to Luther



Notes: Authors with stance-read text in each period: converts 47, 31, 33, 32 and 22, Catholics 128, 55, 51, 75 and 93 before 1517, in 1517–19, 1520–21, 1522–25 and 1526–30; chunks left after recognised passages are dropped, all of a convert's text in the period including what he printed after his break (Appendix Table C10 splits it at the break). Before 1517 these are the authors of the pre-1517 sample with such a chunk; from 1517 the authors of the 1517–30 sample with a chunk printed in the period, together with the converts whose chunks in the period are from after their break, so every author of a period belongs to one of the two samples of Table 1. An author's profile in a period is his mean code on each of the twelve axes (−2 Roman to +2 reform, 0 when silent). Panel (a): the mean over the authors of a side of the author's mean overall lean. Panel (b): the mean, over the 66 pairs of axes, of the Spearman correlation between the two axes across the authors of a side; the band is the 2.5th to 97.5th percentile of 500 author bootstrap draws. Panel (c): the mean cosine between an author's profile and Luther's profile from his writings of 1518–21 (18 chunks), both measured from the midpoint between the two sides' mean profiles of the period; an author at the midpoint counts 0.

gap between the two sides' mean profiles points closely along it, with a cosine of 0.93, and the main direction of the authors' own profiles separates the sides about as well. The line also does not depend on our twelve axes, which are written from the confessions of 1530. The embeddings see no rubric, and there the direction from the texts of Johannes Eck, Luther's opponent at Leipzig, to Luther's own also separates the two sides in the 1520s. Finally, the result does not rest on what the converts wrote after they had broken: on their earlier texts alone, agreement with Luther's profile separates the sides just as well (Appendix Tables C10 and C11).

After 1517 each convert's text moved before his own break, and stances that had been unrelated came to hang together along Luther's line, the converts toward it and the Catholics away from it. Before 1517 the record held the authors' positions but no line along which to count them. This is consistent with coordination models, in which a public event turns positions dispersed across separate questions into common knowledge of who stood where, and of what it meant to stand on either side (Granovetter, 1978; Chwe, 2001; Pinker, 2025).

5 Robustness and Alternative Explanations

Mechanisms unrelated to what these authors thought could, in principle, produce the signal: a language model recognising its authors, text standing in for education, social origin or a stake in the old church, age and geography, the language and the kind of the works, a handful of famous authors, or the selective survival of books and of authors. We take each in turn. None of them, in the form we can measure, absorbs the signal.

5.1 Measurement and Memorisation

The leading concern with a language-model measurement of historical authors is that the model recognises the author and reports what it already knows about him. We guard against this by construction and then test for it. By construction: every name of a person, a place and an institution is masked before any model reads a text, and a passage whose author the model nonetheless names at confidence 90 or more is dropped, 96 passages of 37 authors before 1517; another passage of the same author takes its place unless the model names that one too (Section 3.3; Appendix Table A2). The result does not depend on where this cut is set: with every passage named at confidence 75 or more dropped the AUC is 0.636, and at 50 or more 0.634, against 0.633 in the main set (Appendix Table C2).

The first test leaves out the authors the model may know. Shown up to two masked chunks of each author's text and asked whether it knows the work or its author, the model names 74 of the 176 authors at least once. Without them the tournament reaches an AUC of 0.620 and the embeddings 0.619. A handful of famous authors do not carry the result either: without the 28 authors on a list of well-known figures of the Reformation era, written down before the probe, the AUC is 0.635 (Appendix Table A3). Table 3 holds both fixed in columns (5) and (6).

The second test asks whether the model knows the texts themselves. Asked to continue a passage from its first 250 words, it reproduces almost nothing; the few passages it can continue come mostly from editions of Cicero, Seneca, the Bible and the liturgy and from grammars (Table A2). The event study of Section 4.3 speaks to the same concern from another side. It compares each author with his own earlier text, so a model that merely recalled the author would give every chunk of his the same lean, and the coefficients of Equation (4) would be zero; instead, a convert's lean rises in the years before his break (Figure 6).

The result does not depend on a single model. Four language models from four

developers—OpenAI, Anthropic, Moonshot and DeepSeek—code identical masked chunks under the same stance prompt. Their author-level leans correlate at 0.81 to 0.88, and each ranks the later converts above the Catholics, with AUCs between 0.60 and 0.67 (Appendix Figure C1). This comparison keeps every chunk, recognised passages included, which is why the main model stands at 0.60 here and at 0.59 in Table C1. Performance varies across models, as measurement noise predicts, but the direction is stable under every one. We keep one model as the main measure for transparency.

Three sources of measurement error remain: the model's own noise, the transcription and the catalogue's attributions. Corrected for the attenuation that a split-half reliability of 0.89 implies, the AUC would be 0.641 (Appendix Table C4). The text is a machine transcription of scanned prints, with character error rates of 2.2 percent on Latin and 0.7 percent on German reference pages. Leaving out the works with the most doubtful pages, those on which a second transcription engine disagrees most, gives an AUC of 0.627 (Appendix Table A1). The catalogue credits some authors with reprints of another's text; without these works the AUC is 0.622 (Table C2).

5.2 Biography: Social Background, Education, and Stakes

The score could stand in for who an author was rather than what he argued: his schooling, his social origin, his stake in the old church. The covariates of Section 3.4 let us hold each of these against the signal. We do so by comparing like with like: the AUC is recomputed over pairs of a later convert and a Catholic who belong to the same group, both with a doctorate or both without, and set against the AUC over all pairs (Appendix Table C6, Figure C3). None of the groupings moves it much: a doctorate by 1517 gives 0.61 against 0.62, university teaching 0.64 against 0.62, and the father's standing, known for 75 authors, 0.65 against 0.65. Coverage differs across rows and every item is measured with error. This is a first pass and not a causal decomposition. It shows that the measured versions of these stories do not absorb the association. It does not show that education or origin were unimportant.

We examine institutional stake through the clerical estate and the benefice. Authors with an office or an income from the old church were less likely to leave it: 16 percent of those with either later broke with Rome, against 41 percent of those with neither, a descriptive fact and not an estimate of what a benefice deterred. The concern for the score is that it reads the office rather than the author: a cleric defends the church because he is a cleric, and the model places him at the Roman end. Among authors of equal stakes the tournament's AUC is 0.60, against 0.62 over all pairs, and the embeddings' 0.60 against

0.65. Stakes lower the separation a little and do not absorb it.

Age and geography are real compositional differences: the later converts were younger, and their works were printed closer to Wittenberg on average (Table 2), so the score could merely proxy for either. Of the authors born in 1480 or later, 43 percent broke with Rome; of those born before 1470, 13 percent. Within birth cohorts the AUC is 0.59, against 0.61 over all pairs; within the nearer and the farther half of the distances to Wittenberg it is unchanged (Table C6). Table 3 adds year of birth and distance as controls, and the coefficient on the score remains (columns 3 and 4). Coverage is incomplete and the distance is coarse, that of the town where most of an author's works were printed; what the estimates show is that the measured gradients of age and geography do not absorb the textual association.

5.3 Language and Genre

Language is a further compositional difference. The later converts printed more in German: of the authors who printed mainly in German, 48 percent later broke with Rome, against 23 percent of those who printed mainly in Latin (Table C6). A model that favoured German passages would then find the converts without judging a position. In decided meetings between a German and a Latin passage the German one wins 43.4 percent (95 percent confidence interval 40.9 to 46.0; Appendix Table C4): whatever tilt there is runs toward Latin, against the converts. Among authors who printed mainly in the same language the AUC is 0.65, against 0.63 over all pairs, and the coefficient on the score remains when Table 3 holds the Latin share fixed (column 2).

The kind of work is a different matter. The two sides may have written different kinds of works, and the model could rank the kind of work rather than the position taken in it. Section 4.2 addresses this at the level of the author. The Catholics wrote more on religion, but that accounts for 15 percent of the gap between the two sides' mean scores; and on religious text alone the score separates the sides with an AUC of 0.70 (Figure 3).

At the level of the single work the evidence is weaker. A stricter comparison matches a later convert's work to a Catholic's work in the same language, printed within a few years of it, by an author of the same cohort (Appendix A gives the limits), and asks the model which of the two leans further toward reform, on up to four passages of each and in both orders. The convert's work is nearer reform in 62 percent of the decided pairs among 122 (p 0.015). When the two works must also belong to the same genre, it is nearer reform in 58 percent among 112, which is not distinguishable from an even split (p 0.261; Appendix Table C7). Before 1517 one work set against another of its kind tells the sides apart less

reliably than an author's writing taken together. In 1517 to 1530 a comparison matched on the genre of the work gives 77 percent (Section 4.3).

5.4 Survival Bias and Inference

The corpus is what survived, and survival could correlate with confession—if, say, Protestant collectors preserved their forerunners' early works. We do not model the historical survival process, and whether preservation favoured either side we do not know. We can check only narrower things. The result does not depend on how much of an author's text the corpus holds: the coefficient on the score remains when the amount of text is held fixed (column 2 of Table 3). It also does not depend on the exact date at which we require an author's survival: restricting the sample to the authors known or plausibly alive in 1520, and again in 1525, leaves the AUC at 0.613 and 0.630 and the coefficient of column (4) of Table 3 positive and broadly similar in magnitude (Appendix Table C2).

Finally, the sample is set by what survived and by the limited number of people who were in print before 1517 and lived to face the choice: 48 later converts among 176 authors. With these numbers the confidence intervals are wide, however they are constructed: for the tournament's AUC, 0.541 to 0.725 by the method of DeLong et al. (1988) and 0.529 to 0.714 in a bootstrap over authors (Appendix Table C1). Neither chance nor a single author is a likely source of the result: labels reshuffled across authors rarely reach the tournament's AUC, and no author's removal moves it much (Section 4.1; Appendix Figure C2). Power is the binding constraint on finer cuts: the subsamples of this section hold fewer converts, and their estimates are correspondingly less precise.

6 Conclusion

Great political and religious upheavals catch contemporaries unaware. An Austrian minister is said to have asked who would lead a revolution in Russia—"perhaps Herr Bronstein from the Café Central?", meaning Trotsky; on New Year's Day 1933 the *Frankfurter Zeitung* declared the National Socialist assault on the democratic state repelled, a month before Hitler took power. An influential literature explains why the surprise may be inevitable. When dissent is costly, people falsify their preferences in public, and the information needed to foresee the rupture exists in no public record until the cascade has begun (Kuran, 1991, 1995). We have examined this claim for a paradigmatic ideological revolution, the Reformation that split the Western Church after 1517, at the level of the individual, and of what he had put into print before the rupture.

What 176 authors printed before 1517 predicts who among them would later break with Rome. Embeddings of the text rank a later convert above a Catholic with an AUC of 0.65; a tournament in which a language model compares masked passages two at a time, never seeing a name or a label, reaches 0.63. The embeddings establish the signal without a language model; the language model helps locate where the signal lies. A one-standard-deviation higher score in the tournament is associated with an 8.7 percentage point higher probability of a later break, over a base of 27.3 percent, and with an earlier break. The association holds among authors of the same education, stakes, cohort and language and without the authors the model can name, and four language models from four developers, coding the same texts for their stance, rank the two sides in the same direction.

Before 1517 the future Protestants were generally not writing against Rome. Before an articulated Protestant programme had formed, the public record carried a broader orientation that later aligned with conversion. It sits in religious writing: in the old disputes over vows and the religious life, the pope's power, penance and indulgences, and in intellectual outlook and inward devotion. Justification by faith and Scripture against tradition, the articles on which the movement would stand, do not separate the sides. The clearest asymmetry lies with those who stayed: the Catholics defend doctrine more firmly, and the later converts take no pronounced stance.

After 1517 the two sides come apart in print. The same tournament on the texts of 1517 to 1530, each convert's counted only to the year before his break, separates the sides with an AUC of 0.80. A convert's text moves toward reform years before his documented break, against his own earlier writing and net of what moved every author in the same years. A programme appears in the converts' texts, and stances that had been unrelated fall into line along Luther's profile, the converts toward it and the Catholics away from it. One interpretation—not established by these estimates—is that Luther supplied a programme and common knowledge to a previously diffuse public orientation (Chwe, 2001; Pinker, 2025).

The larger lesson concerns demand. Printing, competition among media, princely politics and distance to Wittenberg shaped the supply of the new faith; the demand has been harder to measure (Section 2). We measure its trace in public print. The evidence is bounded by what survived and by the limited number of people who were in print before the Reformation and lived to face the choice, and it cannot speak for the laity. But it covers the population at the centre of the event: the humanists, theologians, jurists and clerics among whom the Reformation was argued out, and who had to take a side in

public. Among them, the strong claim that a rupture leaves no public trace beforehand fails. For the study of revolutions, this suggests that future dissenters can sometimes be identified before the event: later individual alignment left a public textual trace long before the break with Rome.

References

Ash, E. and Hansen, S. (2023). Text Algorithms in Economics. *Annual Review of Economics*, 15:659–688.

Bainton, R. H. (1950). *Here I Stand: A Life of Martin Luther*. Abingdon-Cokesbury, New York.

Becker, S. O., Hsiao, Y., Pfaff, S., and Rubin, J. (2020). Multiplex Network Ties and the Spatial Diffusion of Radical Innovations: Martin Luther's Leadership in the Early Reformation. *American Sociological Review*, 85(5):857–894.

Becker, S. O., Pfaff, S., and Rubin, J. (2016). Causes and Consequences of the Protestant Reformation. *Explorations in Economic History*, 62:1–25.

Becker, S. O. and Woessmann, L. (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic History. *The Quarterly Journal of Economics*, 124(2):531–596.

Benoit, K., Munger, K., and Spirling, A. (2019). Measuring and explaining political sophistication through textual complexity. *American Journal of Political Science*, 63(2):491–508.

Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *Journal of Human Resources*, 8(4):436–455.

Bradley, R. A. and Terry, M. E. (1952). Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika*, 39(3/4):324–345.

Cameron, E. (2012). *The European Reformation*. Oxford University Press, Oxford, 2nd edition.

Cantoni, D. (2012). Adopting a New Religion: The Case of Protestantism in 16th Century Germany. *The Economic Journal*, 122(560):502–531.

Cantoni, D., Dittmar, J., and Yuchtman, N. (2018). Religious Competition and Reallocation: The Political Economy of Secularization in the Protestant Reformation. *The Quarterly Journal of Economics*, 133(4):2037–2096.

Carlson, D. and Montgomery, J. M. (2017). A pairwise comparison framework for fast, flexible, and reliable human coding of political texts. *American Political Science Review*, 111(4):835–843.

Chwe, M. S.-Y. (2001). *Rational Ritual: Culture, Coordination, and Common Knowledge*. Princeton University Press, Princeton, NJ.

Davidson, R. R. (1970). On Extending the Bradley-Terry Model to Accommodate Ties in Paired Comparison Experiments. *Journal of the American Statistical Association*, 65(329):317–328.

De Gruyter (2026). Verfasser-Datenbank (VDBO). Online database, De Gruyter, Berlin and Boston. https://www.degruyterbrill.com/database/vdbo/html, last accessed 17 July 2026.

de la Croix, D. and Morault, P. (2026). Winners and Losers from the Protestant Reformation: An Analysis of the Network of European Universities. *The Journal of Economic History*, 86(2):460–500.

de la Croix, D., Scebba, R., and Zanardello, C. (2025). Flora, Cosmos, Salvatio: Premodern Academic Institutions and the Spread of Ideas. Discussion Paper 20569, CEPR.

Dell, M. (2025). Deep Learning for Economists. *Journal of Economic Literature*, 63(1):5–58.

DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3):837–845.

Deutsche Biographie (2026). Online database of the Historische Kommission bei der Bayerischen Akademie der Wissenschaften and the Bayerische Staatsbibliothek. https://www.deutsche-biographie.de. Last accessed 17 July 2026.

Dieter, T. (2021). Another Quincentennial: The Diet and Edict of Worms (1521). *Lutheran Quarterly*, 35(1):1–17.

Dittmar, J. and Seabold, S. (2019). New Media and Competition: Printing and Europe's Transformation after Gutenberg. CEP Discussion Paper 1600, Centre for Economic Performance, LSE. Conditionally accepted, *Journal of Political Economy*.

Dittmar, J. E. (2011). Information Technology and Economic Change: The Impact of the Printing Press. *The Quarterly Journal of Economics*, 126(3):1133–1172.

Doberstein, J. W., editor (1959). *Luther's Works, American Edition, vol. 51: Sermons I*. Muhlenberg Press, Philadelphia.

Ekelund, Jr., R. B., Hébert, R. F., and Tollison, R. D. (2002). An Economic Analysis of the Protestant Reformation. *Journal of Political Economy*, 110(3):646–671.

Febvre, L. (1929). Une question mal posée: les origines de la Réforme française et le problème général des causes de la Réforme. *Revue historique*, 161:1–73.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition methods in economics. In Ashenfelter, O. and Card, D., editors, *Handbook of Labor Economics*, volume 4A, pages 1–102. Elsevier, Amsterdam.

Gentzkow, M., Kelly, B., and Taddy, M. (2019a). Text as Data. *Journal of Economic Literature*, 57(3):535–574.

Gentzkow, M., Shapiro, J. M., and Taddy, M. (2019b). Measuring Group Differences in High-Dimensional Choices: Method and Application to Congressional Speech. *Econometrica*, 87(4):1307–1340.

Granovetter, M. (1978). Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6):1420–1443.

He, S., Lv, L., Manela, A., and Wu, J. (2025). Chronologically Consistent Large Language Models. Technical report, arXiv:2502.21206.

Heinemann, F., Nagel, R., and Ockenfels, P. (2004). The Theory of Global Games on Test: Experimental Analysis of Coordination Games with Public and Private Information. *Econometrica*, 72(5):1583–1599.

Iyigun, M. (2008). Luther and Suleyman. *The Quarterly Journal of Economics*, 123(4):1465–1494.

Kim, H. and Pfaff, S. (2012). Structure and dynamics of religious insurgency: Students and the spread of the reformation. *American Sociological Review*, 77(2):188–215.

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, 50(272):1168–1194.

Kuran, T. (1991). Now Out of Never: The Element of Surprise in the East European Revolution of 1989. *World Politics*, 44(1):7–48.

Kuran, T. (1995). *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Harvard University Press, Cambridge, MA.

Liusie, A., Manakul, P., and Gales, M. (2024). LLM comparative assessment: Zero-shot NLG evaluation through pairwise comparisons using large language models. In *Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 139–151, St. Julian's, Malta.

Lohmann, S. (1994). The Dynamics of Informational Cascades: The Monday Demonstrations in Leipzig, East Germany, 1989–91. *World Politics*, 47(1):42–101.

Ludwig, J. and Mullainathan, S. (2024). Machine Learning as a Tool for Hypothesis Generation. *The Quarterly Journal of Economics*, 139(2):751–827.

MacCulloch, D. (1987). *Groundwork of Christian History*. Epworth, London.

MacCulloch, D. (2004). *The Reformation: A History*. Viking, New York.

Moeller, B. (1959). Die deutschen Humanisten und die Anfänge der Reformation. *Zeitschrift für Kirchengeschichte*, 70:46–61.

Moeller, B. (1972). *Imperial Cities and the Reformation: Three Essays*. Fortress Press, Philadelphia.

Morris, S. and Shin, H. S. (1998). Unique Equilibrium in a Model of Self-Fulfilling Currency Attacks. *American Economic Review*, 88(3):587–597.

Morris, S. and Shin, H. S. (2003). Global Games: Theory and Applications. In Dewatripont, M., Hansen, L. P., and Turnovsky, S. J., editors, *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*, volume 1, pages 56–114. Cambridge University Press, Cambridge.

Mullainathan, S. and Spiess, J. (2017). Machine Learning: An Applied Econometric Approach. *Journal of Economic Perspectives*, 31(2):87–106.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.

Oberman, H. A. (1989). *Luther: Man between God and the Devil*. Yale University Press, New Haven.

Ozment, S. (1980). *The Age of Reform, 1250–1550: An Intellectual and Religious History of Late Medieval and Reformation Europe*. Yale University Press, New Haven.

Ozment, S. E. (1975). *The Reformation in the Cities: The Appeal of Protestantism to Sixteenth-Century Germany and Switzerland*. Yale University Press, New Haven.

Pettegree, A. (2005). *Reformation and the Culture of Persuasion*. Cambridge University Press, Cambridge.

Pinker, S. (2025). *When Everyone Knows That Everyone Knows. . . : Common Knowledge and the Mysteries of Money, Power, and Everyday Life*. Scribner, New York.

Qin, Z., Jagerman, R., Hui, K., Zhuang, H., Wu, J., Yan, L., Shen, J., Liu, T., Liu, J., Metzler, D., Wang, X., and Bendersky, M. (2024). Large language models are effective text rankers with pairwise ranking prompting. In *Findings of the Association for Computational Linguistics: NAACL 2024*, pages 1504–1518, Mexico City.

Reimers, C. W. (1983). Labor market discrimination against hispanic and black men. *Review of Economics and Statistics*, 65(4):570–579.

Rubin, J. (2014). Printing and Protestants: An Empirical Test of the Role of Printing in the Reformation. *The Review of Economics and Statistics*, 96(2):270–286.

Scribner, R. W. (1981). *For the Sake of Simple Folk: Popular Propaganda for the German Reformation*. Cambridge University Press, Cambridge.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4):273–286.

Vandiver, E., Keen, R., and Frazel, T. D., editors (2002). *Luther's Lives: Two Contemporary Accounts of Martin Luther*. Manchester University Press, Manchester.

A Data and Methods

This appendix gives the construction of the data behind Section 3: the frame, the transcription, the mask and the dropping of recognised authors, the labels and the break year, the biography coding, and the background facts. It then gives the methods behind Section 4: the embeddings, the tournament, the matched work pairs and the hazard model (Section 4.1; the pairs also in Section 5.3), the genres (Section 4.2), and the texts of 1517 to 1530, the event study and the profiles (Section 4.3).

The frame. The frame is every record with a title as main author of a VD16 edition printed 1501 to 1530, 2,124 records. Of these, 6 are not authors, a wrong or a duplicate authority number on a print or a book title catalogued as a person, and leave before the first count of Table 1, 2,118. An author counts as dead by 1517 when the recorded death year says so; an unknown year keeps him. A worldly ruler is one the frame flags as such, and an author is outside the German lands when his sphere of activity lay elsewhere. Exclusions found in the review of the labels are counted at the step they belong to. The last two steps of the table require a documented side and at least 2,000 characters of transcribed text, whitespace not counted.

The corpus and the transcription. Digitised copies of the editions come from the Bavarian State Library's digital collections and from other libraries, with hand downloads where no copy was online. A vision model, Gemini 3.8 Flash, read 63,375 pages with a diplomatic-transcription prompt; 167 came back empty. A conventional engine, Tesseract with a Fraktur model, read 32,395 of these pages beside it, the ones transcribed first. 4,726 of these hold fewer than 20 words; on the 27,601 others the share of words on which the two engines disagree is 0.169 per page at the median and 0.613 at the 95th percentile. Table A1 gives the character error rates against reference transcriptions and a spot check of 80 pages of the corpus drawn across the disagreement bands, the two periods and the two languages and rated blind to the author's side: 74 of the 77 rated pages are good, with no difference by band, language or period. The last block of the table drops the five percent of pages on which the engines disagree most and refits the measures; Section 5.1 reports it. The language models read the transcription as it stands, with the mask applied; a normalised version of the text, with case, u/v and i/j folded and abbreviation marks removed, serves the word models of Table C1 only.

An author's text is every edition under his name as main author in the corpus. An edition with several authors stays with its main author, and a reprint of another author's

text stays under the name it was printed under; Table C2 gives the tournament without them. From the text one seeded draw takes up to ten chunks of 18,000 characters, counted without whitespace. A chunk is the unit of the stance measure; the tournament reads passages of 400 words from the text (*The tournament* and *The texts of 1517 to 1530*, below). The embeddings and the word models read all of the text.

Table A1. Transcription quality

	Latin	German
Character error rate of the transcription, percent		
Reference pages of the corpus (6 Latin pages of 10 lines, 2 German pages)	2.2	0.7
Published ground truth (GT4HistOCR, 16 pages; Deutsches Textarchiv, 12 pages)	2.2	0.6
Human spot check of pages of the corpus, rated blind to the author's side		
Pages rated good	74 of 77	
Measures without the worst five percent of pages by engine disagreement, AUC (all pages in parentheses)		
Tournament score	0.627 (0.633)	
Word n-grams, cross-validated mean	0.628 (0.625)	
Background facts, cross-validated mean	0.706 (0.703)	
Facts and word n-grams, cross-validated mean	0.657 (0.657)	

Notes: How well the prints are transcribed. Character error rates compare the transcription with a reference transcription of the same pages: pages of our own corpus transcribed by hand (6 Latin pages of 10 lines, 2 German pages) and two published sets, GT4HistOCR for Latin prints of 1471 to 1522 (16 pages) and the Deutsches Textarchiv for German (12 pages, the median over pages). Human spot check: 80 pages of the corpus drawn across five bands of the word disagreement between the two engines, the two periods and the two languages, rated blind to the author's side; 74 of the 77 rated pages are good, with no difference by band, language or period. The last block drops the five percent of pages on which the two engines disagree most (threshold 0.613, the 95th percentile over 27,601 scored pages; 801 of 34,611 text rows) and refits each measure on what is left, against the same measure on all pages. The tournament refit drops the 53 passages of works with a fifth or more of their text on such pages; the rank correlation of the author scores with and without them is 0.998. The word n-grams and the background facts are cross-validated means over the 178 authors of the full sample, as in Table C1.

Masking. Every text any model reads is masked first. The author's own name, in every form the authority file records, becomes [NAME]; every person alive after about 1450 becomes [PERSON]; towns, territories, dioceses and universities become [PLACE]. Initials and truncated names are masked with them, and a model pass catches capitalised tokens the lexicons missed. Biblical and classical names, saints, the Church Fathers, the scholastic and humanist authorities, Rome and the names of nations stay. In the text of 1517 to 1530 the names and labels of the parties, Luther's name and its derivatives, "papist" and "evangelical", become [PARTY].

Recognised authors dropped. Prompt B3 asks who wrote a passage, and how confident the answer is. Two models answered it for every passage of the tournament. gpt-6.1-sol names 24 percent of the passages and Claude Sonnet 5.5 13; a guess counts as naming when it contains the surname. Of the guesses at confidence 90 or more, 54 and

69 percent do. The rule drops a passage that the model (gpt-6.1-sol) names at confidence 90 or more and draws another from the author's text: first from a chunk not yet read, then a further window of 400 words in a chunk already read. The new passage is put to the same question, and it takes the place of the old one only if the model does not name its author in turn. Before 1517 a new passage counts as named from confidence 50, and it is also set aside when the model takes it to be almost entirely another author's text; after two draws of each kind the passage goes without a replacement. The rule there removes 96 passages of 37 authors and replaces 51 of them, and 2 Catholics have nothing left. In 1517 to 1530, with two draws in all, 129 passages are named, 93 replaced and 36 out without replacement; Dürer and Pauli have nothing left, and Luther is left out as the reference of Section 4.3. The stance measure applies the same rule to its chunks: of 1,309 chunks before 1517, 1,213 are kept, and 1,344 of 1,458 in 1517 to 1530. Other cutoffs of the rule are rows of Table C2.

Three checks stand behind the rule. In a separate probe the model was shown up to two masked chunks of each of the 178 authors and asked whether it knew the work or the author; it recognised 76, of whom 74 are in the sample. A list of 46 well-known figures of the Reformation era was written down before the probe. Table A3 gives the AUC of the three measures without either set. The name swap rereads 600 meetings with the true author's name, a false famous name or no name written into the passage and measures the change of the converts' win share against a masked reread. Under the false name, Johann Eck on a later convert's passage and Philipp Melanchthon on a Catholic's, the share changes by -6.3 percentage points; under the true names it changes by 1.4, which is not distinguishable from zero. The model follows a famous name when it is shown one, which is why none is shown. The continuation probe asks the model to write the next 40 words of 1,000 passages and measures the four-word overlap with the true continuation against another passage of the same work. Section 5.1 gives the results of the probe and of the continuation, Table A2 the numbers.

Labels. Behind a documented side stands one of these: a classification in a reference work, VDBO, as a documented act or as the lexicon's own judgement; a documented act found in the biographies; a stated confession; Deutsche Biographie; the religion and occupation fields of Wikidata and of the German authority file; or the review of the research assistants, with a source quotation. A label is strong when a reference work stands behind it, or a database entry confirmed by a second source or by a dated break: 171 of the 176 labels of the pre-1517 sample. The other 5 are weak. An author who broke

Table A2. Author recognition

Check	Design	Result
Recognised authors dropped, before 1517	Every passage of the tournament is put to the model with the recognition question (Prompt B3). A passage whose author it names at confidence 90 or more is dropped and replaced from the author's unused text.	96 passages of 37 authors replaced. AUC 0.633 under the rule, 0.635 with no passage dropped (Table C2).
Recognised authors dropped, 1517–30	The same rule on the passages of 1517–30.	129 passages named, 93 replaced, 36 out without replacement. AUC 0.803 under the rule, 0.799 without it.
Continuation probe	The model is given the first 250 masked words of a passage and writes the next 40. The four-word overlap with the true continuation is set against the overlap with another passage of the same work.	1,000 passages: overlap 0.0095 against 0.0001. 13 passages above 0.2, editions of Cicero, Seneca, the Bible and the liturgy.
Name swap	600 tournament meetings are reread with the true author's name written into the passage, with a false famous name, and with no name. The change in the share of meetings the convert's passage wins is measured against a masked reread.	True name 1.4 points (−0.7 to 3.3); false name −6.3 points (−8.9 to −4.0).

Notes: Checks that the model judges the text and not the author. Every text any model reads is masked first: names of persons, places and institutions are replaced by tags (Appendix A). Recognition: every tournament passage was put to two models, gpt-6.1-sol and Claude Sonnet 5.5, with the question who wrote it (Prompt B3); a guess counts as naming the author when it contains the surname, and the drop uses the answers of gpt-6.1-sol, the model of the tournament. Continuation: the model is given the first 250 masked words of a passage and asked for the next 40; the overlap is the share of four-word sequences of the answer that occur in the true continuation, against the same share for another passage of the same work. Name swap: 600 meetings of the pre-1517 tournament, 214 of them between a convert's and a Catholic's passage; the result is the change, in percentage points, of the share of these meetings won by the convert's passage, with the 2.5th to 97.5th percentile of 2,000 meeting bootstrap draws. Table A3 gives the AUCs without the authors the model can name.

Table A3. Without recognisable authors

Sample	Authors (converts)	Tournament	Embeddings (out of fold)	Stance
All authors	176 (48)	0.63 ± 0.09	0.65 ± 0.09	0.59 ± 0.09
Authors the model recognised left out	102 (32)	0.62 ± 0.12	0.62 ± 0.11	0.60 ± 0.12
Pre-specified list of well-known figures left out	148 (39)	0.64 ± 0.10	0.63 ± 0.10	0.59 ± 0.10
Both left out	92 (28)	0.61 ± 0.13	0.59 ± 0.12	0.58 ± 0.13

Notes: AUC of the three pre-1517 measures when the authors a model may know are left out. Recognised: the 74 authors of the sample whom the model named in a probe on up to two of their masked chunks (the model was asked whether it knew the work or the author; a hit is a guess that contains the surname). Well-known: a list of 46 figures of the Reformation era written down before the probe. Nothing is refitted: the tournament score is that of the fit, the embedding score the out-of-fold score of the classifier of Figure 1 (all 178 authors in the folds), the stance score the author's mean overall lean; each cell is the AUC of the existing scores on the authors that remain, with 95 percent confidence intervals (DeLong). The stance lean exists for 171 of the 176 authors (46 of the 48 converts), the others having no chunk with a lean left after recognised passages are dropped; the stance column of each row is on the row's authors who have one. The reduced samples have fewer converts, so the intervals widen.

with Rome and later returned counts as a convert.

The break year. The break year is the earliest documented public act against Rome. It comes from a grounded biographical search, whose dates a second search confirmed

to within one year, and from the reviews of the labels where the search found none. A year is known for 40 of the 48 converts of the pre-1517 sample and for all 64 converts of the sample of 1517 to 1530, whose text is cut at it. A convert without a year stays in the pre-1517 sample; he is out of the sample of 1517 to 1530, of the hazard model of Section 4.1 and of the event study of Section 4.3.

Biography coding. The coding is as of 1517. The sources, in this order: VDBO, Deutsche Biographie, Wikipedia, Germania Sacra, the Repertorium Academicum Germanicum, and Wikidata without its religion field. A language model, gpt-6.1-sol, extracted the values from the source texts with a verbatim quotation for each; every quotation was checked against its source, and a value whose quotation is not found is flagged and not used. The items: the clerical estate in 1517 (lay, priest, secular cleric, member of an order); a benefice documented before 1517, personal possession of a prebend, canonry, endowed parish or comparable endowed office, where ordination, monastic status, an academic deanship, an expectancy or another person's office do not count; a doctorate by 1517; university study and teaching before 1517; the father's name and standing in four tiers, elite being the urban elite, an office-holder or a noble; the place of activity in 1517 and its polity, a free city being an imperial city or a Swiss confederate city. An item counts when a source records it, and silence counts as no: the sources are silent on the estate for 23.5 percent of the entries, name no father for 92 of the 170 and give no standing for 95. The coding covers 170 of the 176 authors of the pre-1517 sample and every author of the sample of 1517 to 1530.

Background facts and the other author variables. From the corpus and the catalogue we take the works and characters of an author's text in a period, the Latin share among his works with a catalogued language, the mean print year, and the imprint town of most of his works, which gives the region of print and the great-circle distance to Wittenberg. A work is religious when its subject code (*The genres*, below) is theology, preaching, devotion or church administration; Table C2 and Figure C2 use this definition. The first year in print, from VD16 or the incunabula, gives the years in print of Table C3. The background facts of Section 4.1 are year of birth, the mean print year, the Latin share, the length of the text and the region of print.

Embeddings. An embedding model (text-embedding-3-large) turns each author's masked pre-1517 text into a vector of 1,024 numbers. We keep its first 20 principal components and fit a class-balanced logistic regression on them. The authors are split

into five folds; the model is fitted on four and scores the fifth; the split is repeated 100 times, and an author's score is his mean rank out of fold. On the 178 authors of Figure 1, the 2.5th and 97.5th percentiles of the AUC over 2,000 author-bootstrap draws, stratified by side, are 0.551 and 0.735. For the permutation test the model is refitted 1,000 times on labels reshuffled across authors; the statistic is the cross-validated mean AUC (0.62 on the 178 authors), and p is two-sided around the mean of the refits: 0.060 on the 178 authors and 0.049 on the 176 authors of the tournament (Table C1, which shows the latter).

The tournament. From everything printed under an author's name before 1517, reprints included, we draw up to 15 passages of 400 words: one from each of the up to ten chunks of the seeded draw, then one from each chunk the draw left unused and, where an author has none left, windows of the chunks already read that overlap no earlier passage. This gives 2,348 passages; 126 authors have all 15. Each passage meets exactly 8 passages by other authors, printed within five years of it.[14] This gives 9,392 meetings. The likelihood of the Bradley–Terry model carries a ridge penalty of 10^{-3} on the θ_i, which pins down the level of the scale, and the θ_i are centred at zero. Author scores run from -8.65 to 6.89, with a standard deviation of 3.11; a gap of 2 between two passages means that the model calls the first more reform-leaning in 88.1 percent of the meetings it decides. We reshuffle the labels with the scores held fixed, 10,000 times across all authors and 10,000 times within language and birth decade, so that German authors stay German and young authors stay young. The tournament's AUC is reached in 2.6 of every 1,000 draws of the first kind and in 19 of every 1,000 of the second, a one-sided count. Leaving out any one author keeps the AUC between 0.625 and 0.646 (Figure C2). Order consistency, the share of meetings in which both orders pick the same passage, is 0.925; the rest are ties. We split each author's passages at random into two halves and score each half; stepped up by the Spearman–Brown formula, the split-half reliability of the score is 0.89 across authors (Table C4). Quoted reasons are the model's evidence line without the references to its sentences.

Matched work pairs. A later convert's work printed before 1517 is matched to a Catholic's work with the same language in the catalogue, printed within two years of

[14] The catalogue credits 5 of these works to the author they attack, among them the Cologne articles against Reuchlin's *Augenspiegel*; their 8 passages are not in the tournament. The 64 passages that would have met one of them are paired among themselves instead, in 31 new meetings and one that already existed.

it and, where both birth years are known, by an author born within five years of the convert; a work without a partner within these limits is matched within three years and ten. Works whose language the catalogue leaves undetermined are matched among themselves. A second design repeats the matching with the genre of Section 4.2 as an exact match. In both, each Catholic work is used once. From each work up to four passages of 400 words are drawn, and the model compares every passage of the one work with every passage of the other under the prompt of the tournament (Prompt B1), in both orders. A comparison counts when both orders pick the same passage; comparisons with a passage the model names at confidence 90 or more are left out, and both authors are in the sample. The pair goes to the work that wins more comparisons, and equal counts are a tie. The percentages in the text are shares of the decided pairs. One convert can have many pairs, so the p-value comes from an exact two-sided sign-flip test in which all pairs of a convert change sign together, with 43 converts in the first design and 40 in the second (Table C7).

The hazard model. The clock starts in 1517. An author leaves the risk set in the year of his break or of his death, and 1530 censors. The sample is the 168 authors with a documented side and, for converts, a documented year of the break, 38 breaks by 1530; 8 converts without a year are left out. The hazard is a logit on author-years, linear in time, with the score standardised over these authors; its p-value is two-sided, from 1,000 reshuffles of the score across them. A Cox model gives 1.48 and 1.44 (Table C3).

The genres. One fixed prompt reads each work's title and two extracts of its masked text and returns a subject code (Prompt B4). The 644 works behind the passages of the pre-1517 tournament form three genres: religion (189 works), literature and education (353) and law, medicine and science (102); the note of Figure 3 lists the subjects of each. Of the 644 works, 53 coded as law or as polemic were sorted by hand: canon law and religious controversy to religion, secular polemic and news to literature and education, civil law to law, medicine and science. The genre religion is wider than the religious works of Table C2 and Figure C2, which leave out canon law and religious controversy. An author's score in a genre is the mean strength of his passages from works of that genre, and his lean in a genre the mean of the model's overall lean over his chunks from them; an author enters a genre when he has such a passage or chunk. The embeddings of Figure 3 are the classifier of Figure 1 fitted on each genre's text alone, on the authors with 2,000 characters in works of the genre. Figure 4 uses the chunks of the stance measure, sorted

by the genre of their work.

The texts of 1517 to 1530. The second tournament reads what the authors printed from 1517 to 1530, a convert's text only to the year before his documented break. The sample is 224 authors, 64 of whom later broke with Rome (Section 3.1). Each author has up to 10 passages of 400 words, one from each chunk of the seeded draw, 1,434 in all; each passage meets 8 passages by other authors printed within five years of it, 5,736 meetings. The model, the prompt, the two reading orders and the dropping of recognised passages are those of the tournament before 1517; the names and labels of the parties are masked as well (*Masking*, above); and Equations (1) and (2) give the score. Table C4 gives order consistency and split-half reliability of both tournaments. The shorter windows of Table C8 score the authors with text in 1517 to 1521 (137 authors, 51 converts) and in 1517 to 1525 with the strengths of this fit. The embeddings are fitted as before 1517, on the 227 authors of the full sample of these years. The regression of Table C8 standardises the score over the tournament's authors and leaves out the mean print year, since a convert's text stops at his break and his works are earlier by construction; the background facts leave it out for the same reason. The matched pairs of 1517 to 1530 (Table C7, rows 3 to 5) follow the rule of the pairs before 1517 (*Matched work pairs*, above) and match the genre of the two works as well; the model compares whole chunks.

The event study. The event study of Equation (4) uses every chunk left after recognised passages are dropped, with a lean and a print year: 2,342 chunks of 280 authors, Catholics among them, and for the 76 converts with a documented year of the break their writings after it as well. The periods t are before 1510, 1510 to 1516, 1517 to 1519, 1520 to 1521, 1522 to 1525 and 1526 to 1530. Standard errors are clustered by author. Most converts are not observed on both sides of their break, so the second row of Table C9 restricts the model to the 17 converts who have text both before 1517 and after their break. In the placebo each Catholic is given a break year drawn with replacement from the converts' break years, the model is refitted on the Catholics alone, and the draw is repeated 5,000 times (the third row of the table).

Profiles and coordination. The profiles of Section 4.3 use the chunks of the stance measure left after recognised passages are dropped, and all of a convert's text, what he printed after his break included; Table C10 splits it at the break. An author's profile in a period is his mean code on each of the twelve axes over his chunks of the period, a chunk that does not address an axis counted as 0. Among authors who take a position on both

axes of a pair (panel (b) of Figure 8), over the pairs with at least five such authors, the mean correlation in the five periods is 0.18, 0.18, 0.49, 0.27 and 0.17 among the converts and 0.22, 0.35, 0.43, 0.32 and 0.29 among the Catholics, on 7.3 to 15.2 converts per pair. Luther's profile is his mean code on each axis in 18 chunks of his writings of 1518 to 1521: 1.67 on affective piety, 1.39 on the call for reform, 1.00 on clerical abuses and on Scripture, 0.83 on justification, 0.11 on indulgences, 0.11 on monasticism and −0.22 on the saints. Table C11 sets beside the agreement with Luther's profile (Figure C5) the first principal component of the authors' centred profiles and the cosine between the gap of the two sides' mean profiles and a reference profile, Luther's or that of Erasmus's writings of 1517 to 1521 (5 chunks). The last rows of Table C10 repeat the measure without the rubric: each author's text of a period is embedded, the language of the text is partialled out, and the embedding is projected on the line from Eck's texts of 1517 to 1521 to Luther's.

B The Twelve Axes and the Prompts

This appendix holds the twelve axes of the stance measure with the texts that anchor each end (Table B1) and the four prompts, word for word.

B.1 The Twelve Axes

The eight doctrine axes are the articles on which the Augsburg Confession of June 1530 and the Confutatio of August 3, 1530 contradict each other. The Confession anchors the reform end of each axis and the Confutatio the Roman end, each with a quotation from the text. Neither treats indulgences: there Luther's Ninety-five Theses anchor the reform end, and Leo X's *Cum postquam* of 1518 and the decree of Trent the Roman end. The Confession does not state the limit of papal power, so the Theses anchor that end too. The real presence in the sacrament is not coded. The four axes on outlook and grievance are positions held on both sides before 1517. They come from the historians named in the table, Moeller, Hamm, Scribner and Jedin among them, and have no confessional text at the Roman end; the table says so where a cell has none.

The model codes every axis on one scale. A chunk is not addressed when it is silent on the question or names it in passing. It is addressed when it asserts, denies, prescribes, performs or argues something about the matter; naming is not addressing. An addressed chunk takes a strong or qualified Roman position, a qualified or strong reform position, or no lean. A strong code needs an explicit position that is central to the chunk, its thesis

or argued at length. The model codes the text's own voice. Where the text reports or refutes another's view, the stance of the reporting voice is coded, and where that voice is not visible, the chunk is marked as not attributable. Each axis is coded on its own; a code on one implies nothing about another. Criticism of the sale of indulgences or of the traffic in masses that leaves the doctrine intact is coded as clerical abuses, not on the doctrinal axis. In the profiles the codes become −2, −1, 0, +1 and +2; a chunk that is not addressed counts 0, and a chunk that is not attributable is left out. The unit is a chunk of masked text of up to 18,000 characters. Chunks whose passage the model names at confidence 90 or more are left out, as in the tournament (Section 3.3).

Table B1. The twelve axes

Axis	Roman end (−2, −1)	Reform end (+1, +2)	Text that anchors the reform end	Text that anchors the Roman end
Doctrine				
Justification	Merit or works contribute to salvation together with grace.	Salvation by grace and faith alone; works are fruits, not merit.	CA IV: "men cannot be justified before God by their own powers, merits or works, but are justified freely, for Christ's sake, through faith"	Confutatio (1530), To Article IV: "For it is entirely contrary to holy Scripture to deny that our works are meritorious." Later: Trent VI, can. 12.
Indulgences	Defends indulgences, the treasury of merits or purgatory.	Denies their basis or efficacy.	Luther, 95 Theses (1517), th. 36: "Any truly contrite Christian has plenary remission of penalty and guilt owed to him even without letters of indulgence."	Trent XXV (4 Dec 1563), Decretum de indulgentiis: "teaches and enjoins that the use of indulgences is to be retained in the Church, and condemns with anathema those who either assert that they are useless or deny that the Church has the power to grant them" Not treated in the Confession or the Confutatio; in the period: Leo X, *Cum postquam* (1518).
Papal authority	Pope, councils and hierarchy bind in faith; canon law binds.	+1: the pope can err or his spiritual power is limited, while councils bind. +2: councils can err too, or papal spiritual power is denied.	95 Theses, th. 5: "The pope neither wishes nor is able to remit any penalties except those he has imposed by his own authority or that of the canons."	Confutatio (1530), Part II, To Article XXVIII (VII. Of Ecclesiastical Power): "For it is most abundantly proved that ecclesiastical power in spiritual things has been founded upon divine right"
Scripture and tradition	Church, Fathers, councils or custom govern the reading of Scripture.	Scripture comes first and judges tradition.	CA XXII: "a custom introduced against God's commands is not to be approved"	Confutatio (1530), To Article XV: "they must be admonished that the princes and cities see to it that the ecclesiastical rites of the Church universal be observed in their dominions and districts" Later: Trent IV, Decretum de editione et usu.
Mass as sacrifice	The mass is a sacrifice offered for the living and the dead; private masses; communion under one kind.	Christ's sacrifice was once for all; the Supper is received in faith, not offered as a work; the lay chalice. The real presence is not coded.	CA XXIV: "For Christ's passion was an oblation and satisfaction not only for original guilt but for all other sins as well"	Confutatio (1530), Part II, To Article XXIV (III. Of the Mass): "Finally, the force of the word shows that the mass is a sacrifice" Later: Trent XXII (17 Sept 1562), cap. II heading.

Axis	Roman end (−2, −1)	Reform end (+1, +2)	Text that anchors the reform end	Text that anchors the Roman end
Monasticism and celibacy	Affirms vows, orders, clerical celibacy, the religious life as a higher state.	Rejects or relativises them.	CA XXIII (Pars II, art. II): "For no human law, no vow can abolish God's command and ordinance."	Confutatio (1530), Part II, To Article XXIII (II. Of the Marriage of Priests): "From these facts it is regarded sufficiently clear that the celibacy of the clergy is not an abuse" Later: Trent XXIV (11 Nov 1563), can. 10.
Saints, Mary, relics	Promotes invocation of saints and Mary, relics, pilgrimage, images.	Rejects invocation or the cult. Saints held up only as examples is B.	CA XXI: "But Scripture does not teach us to invoke the saints or to seek help from the saints, because it sets before us the one Christ as mediator, propitiation, priest and intercessor."	Confutatio (1530), To Article XXI: "for in favor of the invocation of saints we have not only the authority of the Church universal but also the agreement of the holy fathers" Later: Trent XXV (3 Dec 1563), De invocatione ... sanctorum.
Penance and confession	Contrition, complete confession of all sins, satisfaction; forgiveness through the priest's absolution.	Lifelong inward repentance; forgiveness received by faith; full enumeration unnecessary.	CA XII: "Penance properly consists of these two parts: one is contrition, or the terrors struck into the conscience on recognition of sin; the other is faith"	Confutatio (1530), To Article XII: "For when they ascribe only two parts to repentance, they antagonize the entire Church, which from the time of the apostles has held and believed that there are three parts of repentance - contrition, confession and satisfaction." Later: Trent XIV (25 Nov 1551), can. 4.

Axis	Roman end (−2, −1)	Reform end (+1, +2)	Text that anchors the reform end	Text that anchors the Roman end
Outlook and grievance (held on both sides)				
Scholastic and humanist	Scholastic method, school authorities, Aristotelian quaestiones.	Ad fontes, philology, eloquence, Greek or Hebrew, rejection of Aristotle in theology.	Luther, Disputatio contra scholasticam theologiam (Sept 1517), th. 43–44 (WA 1, 226): "It is an error to say: without Aristotle no one becomes a theologian. . . . Rather, no one becomes a theologian unless he does so without Aristotle." Luther, Disputatio contra scholasticam theologiam, th. 50 (WA 1, 226): "In short, the whole Aristotle is to theology as darkness to light." Historians: Moeller (1959): "Ohne Humanismus keine Reformation". Spitz (1963); Rummel (1995, 2000); Burnett (2021).	–
Affective piety	Devotion centred on external observance and ceremony.	Inward, affective, Christ-centred devotion; the soul's direct relation to Christ.	CA XX: "Formerly consciences were tormented by the doctrine of works; they heard no consolation from the Gospel." Historians: Hamm (1992, 1993): "normative Zentrierung". Hamm (1982); Oberman (1977); Steinmetz (1980); Ozment (1975).	Trent XXII, cap. V: "the nature of men is such that it cannot easily be raised to the meditation of divine things without external aids"
Clerical abuses	Defends the clergy's estate, privileges or income against criticism (immunity, tithes, mendicancy, benefices).	Attacks clerical greed, corruption or hypocrisy.	95 Theses, th. 28: "It is certain that when the coin clinks in the chest, gain and avarice can increase" Historians: Scribner (1987); Goertz (1987); Cohn (1979); Dykema and Oberman (1993).	–
Call for reform	Defends the church as it is against calls for change.	Calls for reform of church, clergy, doctrine or worship.	Giles of Viterbo, opening oration, Fifth Lateran Council, 1512 (Mansi 32, 669): "men must be changed by religion, not religion by men" Historians: Jedin (1946); Hamm (1993); Strauss (1971); O'Malley (2000).	–

Notes: The twelve axes of the stance measure, as the model is told them (Prompt B2). On each axis the model codes a chunk of text −2 or −1 (Roman position, strong or moderate), +1 or +2 (reform position), 0 when the chunk addresses the question without leaning, and marks it silent when the question is not addressed; silence counts as 0 in every profile of the paper, and a code that cannot be attributed to the text's own voice is left out. The eight doctrine axes are anchored at the reform end in the Augsburg Confession (CA, 1530) and, where it is silent, in Luther's Ninety-five Theses, and at the Roman end in the Confutatio (1530) or the decrees the Confession answers. The four axes on outlook and grievance are positions held on both sides before 1517; they are anchored in the historians who proposed them and, where one exists, in a text of the period, and a dash marks an end that no text anchors. Quotations are in the wording of the editions cited; English renderings are ours.

B.2 The Prompts

The four prompts follow. Prompt B1 is the question of the tournament (Section 4.1), put twice to every meeting with the order of the two passages swapped; its evidence line is the sentence Table C13 counts. Prompt B2 is the stance measure of Section 4.2 and of

the profiles of Section 4.3. Prompt B3 is the recognition question, put to every passage and chunk; confidence 90 is the cut (Section 3.3). Prompt B4 reads the subject of a work from its title and two extracts of its masked text; the codes form the three genres of Section 4.2.

Prompt B1. Tournament

```
TASK
Below are two excerpts, A and B, from printed books c. 1490 to 1530. Names are partly masked: each author's own name
appears as [NAME]; other people alive after about 1450 as [PERSON]; contemporary towns, territories, dioceses and
universities as [PLACE]. Not masked: biblical and classical names, saints, Church Fathers, scholastic and humanist
authorities, Rome, and nations or large regions. Decide which of the two leans more toward the reform side.
Reform-leaning, broadly: justification by faith, against indulgences, papal or conciliar authority limited, Scripture
over tradition, the mass not a sacrifice, the lay chalice, against vows and celibacy, against the cult of saints, inward
repentance over confession and satisfaction - and the dispositions that went with these: humanist over scholastic method,
inward Christ-centred devotion over external observance, attacks on clerical abuses, calls for reform. Weigh doctrine,
grievances and outlook together; if something else in a text moves your judgement, name it in the evidence line.
Ordering: a text that takes the traditional side leans less toward reform than one that says nothing; one that says
nothing leans less than one that takes the reform side. Do not prefer a text for being longer, more religious, more
learned or more polemical as such.
Judge each text in its own voice, from the text alone. Do not try to identify the authors; do not use what you may know
about them or about what happened later. Do not try to fill a mask.

DECISION
You must choose one. If the difference is small, choose anyway and mark the margin slight.

OUTPUT
Exactly three plain-text lines, nothing else:
EVIDENCE=one sentence naming what decided it, with span IDs from both texts (A-S004, B-S011)
MORE_REFORM=A or B
MARGIN=clear or slight

TEXT A
«<SEGMENTED_TEXT_A»>
TEXT B
«<SEGMENTED_TEXT_B»>
```

Prompt B2. Stance measure

```
TASK
Read the document below: an excerpt from a printed book, c. 1490 to 1530. Names are partly masked: the author's own
name appears as [NAME]; other people alive after about 1450 as [PERSON]; contemporary towns, territories, dioceses and
universities as [PLACE]. Not masked: biblical and classical names, saints, Church Fathers, scholastic and humanist
authorities, Rome, and nations or large regions. Do not try to fill a mask.
For each axis, judge the position THIS TEXT takes, in its own voice, from the text alone, and cite one to three span IDs
(S001, S002, ...). Do not try to identify the author; do not use what you may know about who wrote it or what happened
later. Ordinary knowledge of theology and of the period is expected.

CODES - one scale, used identically on every axis
- Z: not addressed. Silence and passing reference are Z.
- C2 / C1: strong / qualified traditional (Roman) position.
- B: addressed without lean: states or weighs both poles, or treats the matter without taking a side.
- R1 / R2: qualified / strong reform position.
- U: addressed, but the direction cannot be attributed to the text's own voice (reported speech, a dialogue partner, an
opponent quoted for refutation). U is about attribution, never about your uncertainty.

RULES
1. Addressed means the text asserts, denies, prescribes, performs or argues something about the matter. Naming it is not
addressing it.
2. R2 / C2 requires an explicit position central to this excerpt: its thesis, or argued at length. A position only
enacted (a prayer, a practice, a framework assumed without argument) is R1 / C1.
3. Code the text's own voice. Where it reports or refutes another's view, code the stance of the reporting voice; if
that voice is not visible in the excerpt, code U.
4. Axes are independent: a code on one axis implies nothing about another. On the disposition axes, C and R name the
poles only.
5. Criticism of the sale of indulgences or of the traffic in masses that leaves the doctrine intact is clerical_abuses,
not the doctrinal axis.

AXES (C = traditional pole, R = reform pole)
Doctrinal
```

- justification: C: merit or works contribute to salvation together with grace. R: salvation by grace and faith alone; works are fruits, not merit.
- indulgences: C: defends indulgences, the treasury of merits or purgatory. R: denies their basis or efficacy.
- papal_authority: C: pope, councils and hierarchy bind in faith; canon law binds. R1: the pope can err or his spiritual power is limited, while councils bind. R2: councils can err too, or papal spiritual power is denied.
- scripture_vs_tradition: C: church, Fathers, councils or custom govern the reading of Scripture. R: Scripture comes first and judges tradition.
- mass_sacrifice: C: the mass is a sacrifice offered for the living and the dead; private masses; communion under one kind. R: Christ's sacrifice was once for all; the Supper is received in faith, not offered as a work; the lay chalice. The real presence is not coded.
- monasticism_celibacy: C: affirms vows, orders, clerical celibacy, the religious life as a higher state. R: rejects or relativises them.
- saints_mary_relics: C: promotes invocation of saints and Mary, relics, pilgrimage, images. R: rejects invocation or the cult. Saints held up only as examples is B.
- penance_confession: C: contrition, complete confession of all sins, satisfaction; forgiveness through the priest's absolution. R: lifelong inward repentance; forgiveness received by faith; full enumeration unnecessary.

Disposition (outlook and grievances; held on both sides)

- scholastic_vs_humanist: C: scholastic method, school authorities, Aristotelian quaestiones. R: ad fontes, philology, eloquence, Greek or Hebrew, rejection of Aristotle in theology.
- affective_christocentric: C: devotion centred on external observance and ceremony. R: inward, affective, Christ-centred devotion; the soul's direct relation to Christ.
- clerical_abuses: C: defends the clergy's estate, privileges or income against criticism (immunity, tithes, mendicancy, benefices). R: attacks clerical greed, corruption or hypocrisy.
- call_for_reform: C: defends the church as it is against calls for change. R: calls for reform of church, clergy, doctrine or worship.

OVERALL

One code (C2, C1, B, R1, R2 or Z) for the lean of the text as a whole, reform understood broadly: doctrine, grievances against clergy and church practice, and intellectual outlook, weighed together. Z only if no axis is addressed in the text's own voice.

OUTPUT

Exactly thirteen plain-text lines, nothing else, in this order: justification, indulgences, papal_authority, scripture_vs_tradition, mass_sacrifice, monasticism_celibacy, saints_mary_relics, penance_confession, scholastic_vs_humanist, affective_christocentric, clerical_abuses, call_for_reform, OVERALL.

Each line: name=CODE|spans - one to three span IDs separated by commas, no brackets or ranges; empty only for Z. For B cite the spans that show the balance; for U the span whose speaker is unclear.

Example:

justification=R1|S004,S017

indulgences=Z|

SEGMENTED DOCUMENT

«<SEGMENTED_TEXT»>

Prompt B3. Recognition

Names are partly masked: the author's own name appears as [NAME]; other people alive after about 1450 as [PERSON]; contemporary towns, territories, dioceses and universities as [PLACE]. Not masked: biblical and classical names, saints, Church Fathers, scholastic and humanist authorities, Rome, and nations or large regions.

The excerpt below comes from a book printed c. 1490 to 1530. Say who you think wrote it and which work it is from. Guess even if you are unsure, and say how sure you are.

OUTPUT

Exactly three plain-text lines, nothing else:

AUTHOR=surname and first name of the author, or unknown

WORK=short title of the work, or unknown

CONFIDENCE=0 to 100, how likely your author guess is right (0 when unknown)

TEXT

«<SEGMENTED_TEXT»>

Prompt B4. Subject of a work

You are a historian of early printing. Below are the catalogue title, the opening pages and a passage from the middle (OCR, may be noisy; [...] marks the gap) of a book printed c. 1470-1530. Names of contemporary people and places are partly masked as [NAME], [PERSON] and [PLACE], and in books printed after 1517 the names of the religious parties as [PARTY]. Do not try to identify the author.

Classify the book by its main subject: exactly one class from the list below.

CONTENT BEFORE FORM: decide by what the book is about, not by its literary form. A saint's life or a prayer in verse is devotion; a dialogue or a letter about doctrine is theology; a poem or oration on law, medicine or politics goes to that subject; a sermon is preaching whatever its theme. Use literature only when the content itself is literary (poetry, drama, satire, letters or orations whose subject is not one of the other classes), and language_education only when the

```
book teaches language, rhetoric or school matter.  Each class names the classes of the Universal Short Title Catalogue
(USTC) it holds.  University publications (disputations, theses, academic orations, lectures) go to the class of their
faculty:  theology to theology, law to law, medicine to medicine, arts to philosophy or language_education.

CLASSES
- theology:  academic theology, Scripture commentary, doctrine, quaestiones, Bibles, catechisms (USTC Religious, Bibles,
Catechisms)
- preaching:  sermons, sermon collections, postils, funeral orations (USTC Sermons, Funeral Orations)
- devotion:  prayer, meditation, liturgy, hymns, saints' lives, pilgrimage, devotional and pastoral manuals (USTC
Religious)
- church_admin:  documents issued by or for the church:  synodal statutes, episcopal and papal mandates, canon law in
practice, indulgence documents (USTC Religious, Jurisprudence, Ordinances and Edicts)
- law:  civil and canon law as scholarship, legal procedure, legal opinions (USTC Jurisprudence)
- government_ordinances:  secular ordinances and edicts, government and political thought, petitions (USTC Ordinances and
Edicts, Government and Political Thought, Petitions)
- medicine:  medical texts, plague tracts, surgery, pharmacy, health regimens (USTC Medical Texts)
- science_prognostication:  science, mathematics, astrology, cosmography, almanacs, calendars, prognostications, botany,
zoology, agriculture, music (USTC Science, Mathematics, Astrology and Cosmography, Almanacs Calendars and
Prognostications, Botany and Zoology, Agriculture, Music)
- philosophy:  philosophy and morality, logic, Aristotle and his commentators (USTC Philosophy and Morality)
- language_education:  grammar, rhetoric and dialectics, letter writing, dictionaries, school texts, adages (USTC
Educational Books, Linguistics and Philology, Dictionaries, Dialectics and Rhetoric, Adages and Emblem Books, Multilingual
Book)
- classical_authors:  an edition, translation or commentary of an ancient Greek or Latin non-Christian author (USTC
Classical Authors)
- literature:  poetry, orations, dialogues, letters as literature, satire, drama, Latin or vernacular (USTC Poetry,
Orations, Literature, Drama)
- history_news:  chronicles, histories, biographies other than saints', news, travel, genealogy, antiquities (USTC History
and Chronicles, News Books, Newspapers, Periodicals, Travel, Heraldic Works and Genealogies, Antiquarianism, Commemorative
Events)
- politics_polemic:  political tracts, war and the Turks, the Jews, military, economics, witchcraft (USTC Political
Tracts, Military Handbooks, Economics, Witchcraft and Demonology)
- other:  occasional prints and anything else:  wedding or greeting pamphlets, art, cookery, etiquette, games, book trade
(state briefly what it is in NOTE)

OUTPUT
Exactly two plain-text lines, nothing else:
SUBJECT=<class>
NOTE=<at most ten words on what the book is>

CATALOGUE TITLE
«<TITLE»>

TEXT
«<TEXT»>
```

C Additional Results

This appendix holds the tables and figures behind Section 4 and Section 5: the prediction before 1517 and the tournament (Section 4.1), the alternative explanations (Section 5), the texts of 1517 to 1530 (Section 4.3) and the genre decomposition (Section 4.2).

C.1 Prediction before 1517

Table C1 gives every measure of the pre-1517 text on the same 176 authors with its intervals: the tournament, the embeddings, word n-grams (0.67) and the model's lean (0.59), then the background facts and the embeddings' share. Table C2 refits the tournament on other sets of passages, the rows that Section 5 refers to; the AUC stays between 0.613 and 0.641. Table C3 is the hazard model of Section 4.1.

Table C1. All pre-1517 measures

Measure	AUC ± 95 % CI	Cross-validated mean	Author bootstrap, 95 %	Permutation p
Tournament score (one fit, no labels)	0.63 ± 0.09	0.62	0.53–0.71	0.003
Embeddings	0.65 ± 0.09	0.62	0.56–0.74	0.049
Word n-grams	0.67 ± 0.09	0.63	0.58–0.75	0.011
Stance, overall lean	0.59 ± 0.09	0.57	0.48–0.66	—
Background facts alone	0.71 ± 0.08	0.70	0.63–0.80	—
Background facts and embeddings	0.74 ± 0.08	0.69	0.66–0.81	—
share of the embeddings: (0.65 − 0.5) / (0.74 − 0.5)	65 %			
Shuffled labels, word n-grams	0.50 ± 0.09	0.50	0.41–0.59	—

Notes: Every measure of the pre-1517 text on the same 176 authors (48 later converts). First column: the AUC of the author score (the tournament's is the score of Equation (2)); the figure after the ± is the half-width of its 95 percent confidence interval (DeLong); for the tournament and the stance lean this is the plain AUC of the score, for the fitted measures the AUC of the author's mean out-of-fold rank over 100 repeats of a five-fold cross-validation split by author, with only these authors in the folds (Figure 1 uses the full sample of 178). Cross-validated mean: the mean over the 100 repeats of the AUC of each repeat. Author bootstrap: 2.5th to 97.5th percentile of the AUC over 2,000 draws of authors with replacement, stratified by side. Permutation p: two-sided, around the mean of 1,000 refits with the labels permuted across authors; a dash where no permutation was run. The tournament's p is one-sided: the share of 10,000 draws of random labels on its fixed scores that reach its AUC (Appendix A). Word n-grams: a logistic regression on word counts with an L2 penalty chosen in an inner fold. Background facts: those of Figure 1. The share of the embeddings is the gain of facts plus embeddings over 0.5 that the embeddings alone account for. Shuffled labels: the word model refitted on labels shuffled once, which gives chance.

Table C2. Other passage sets

Set of passages	Passages	Authors (converts)	AUC (95 % CI)	AUC, religious / other	LPM, all controls
Main: named at 90+ dropped and replaced	2,348	176 (48)	0.633 (0.54–0.72)	0.69 / 0.58	0.088 (0.036)
No drop	2,297	176 (48)	0.635 (0.54–0.73)	0.70 / 0.58	0.093 (0.036)
Named at 75+ dropped and replaced	2,149	175 (47)	0.636 (0.54–0.73)	0.70 / 0.59	0.085 (0.037)
Named at 50+ dropped and replaced	2,060	174 (46)	0.634 (0.54–0.73)	0.70 / 0.59	0.074 (0.038)
Main, reprints of other authors' works and two works of unclear authorship out	2,077	162 (42)	0.622 (0.52–0.72)	0.67 / 0.59	0.064 (0.037)
Main, work named at 90+ out	2,323	175 (47)	0.641 (0.55–0.73)	0.69 / 0.60	0.085 (0.036)
Main, works of unclear print kind out	2,246	175 (47)	0.619 (0.52–0.71)	0.65 / 0.58	0.078 (0.037)
Main, authors alive in 1520	2,140	161 (48)	0.613 (0.52–0.71)	0.67 / 0.58	0.081 (0.040)
Main, authors alive in 1525	1,806	136 (43)	0.630 (0.53–0.73)	0.74 / 0.58	0.102 (0.043)

Notes: The headline of Table 3 and Figure 2 under other sets of passages, each row one separate fit of the tournament (Equation (1)) on the pre-1517 sample. Main: a passage the model names at confidence 90 or more is dropped and replaced from the author's unused text (Section 3.3). No drop keeps the named passages; the next two rows apply the rule at confidence 75 and 50. Reprints out drops the passages of works that are reprints of other authors' texts (56 author-works) and of the two works of unclear authorship credited to the artists Burgkmair and Dürer. Work named out drops passages whose work, not only author, the model names. Unclear print kind out drops the 42 author-works whose kind of print could not be settled. Alive in 1520 and alive in 1525 keep the main set and the main fit but only the authors recorded as alive in that year (an author with no recorded death year is kept): an author who died soon after 1517 had little occasion to break, so these rows ask whether the ranking rests on authors who could still choose. Authors are those of the sample with a passage left in the set. Columns: the AUC with its 95 percent confidence interval (DeLong); the AUC of the score computed from the author's religious and from his other passages; the coefficient of Equation (3) as in Table 3 column (4) on the row's score, standardised over the row's authors, with HC1 standard errors.

Table C3. Hazard of the break

Score and sample	Authors (breaks)	Controls	Hazard ratio per SD	95 % CI	p
Tournament, all labelled authors (who breaks and when)	168 (38)	none	1.47	1.03–2.10	0.025
		birth	1.42	0.99–2.04	0.042
		years in print	1.50	1.05–2.14	0.019
		birth + years in print	1.43	0.99–2.05	0.042
Cox model (Efron ties)					
Tournament, all labelled authors (who breaks and when)	168 (38)	none	1.48	1.04–2.11	0.020
		birth	1.44	1.01–2.05	0.035
		years in print	1.51	1.06–2.15	0.014
		birth + years in print	1.44	1.01–2.06	0.036

Notes: Time to the break instead of its occurrence: an author who died early could not break. Sample: the Catholics and the converts with a documented break year among the pre-1517 sample (168 authors, 38 breaks by 1530; the 8 converts without a documented year are left out). The clock starts in 1517; an author leaves the risk set in the year of his break or his death, and 1530 censors. Upper block: a discrete-time hazard model, a logit on author-years linear in time, with the tournament score (Equation (2)) standardised over the row's authors; the hazard ratio is the change in the yearly odds of a break per standard deviation, with a Wald 95 percent interval. Lower block: a Cox proportional hazards model (Efron method for ties) on the same data. Controls: year of birth and the number of years the author had been in print by 1517, each with a flag and the mean filled in where missing. p: a permutation test that reshuffles the score across the row's authors 1,000 times, two-sided.

C.2 The Tournament

Table C4 gives the reliability of the two tournaments and the model's language tilt. Figure C1 compares the author scores of four models with the tournament, the embeddings and the word n-grams; the four models agree among themselves at 0.81 to 0.88. Table C5 lists the eight works at each end of the scale; Bomhouver's indulgence instructions stand at the Roman end. Figure C2 drops one author at a time.

Table C4. Tournament reliability

Measure	Before 1517	1517–30 window
Order consistency (the two reading orders agree)	0.925	0.956
Split-half reliability of the author score	0.89	0.85
German passages' win share in cross-language meetings	43.4 % (40.9–46.0)	49.1 % (47.0–51.3)
Correlation of the author score with his German share	−0.07	−0.02

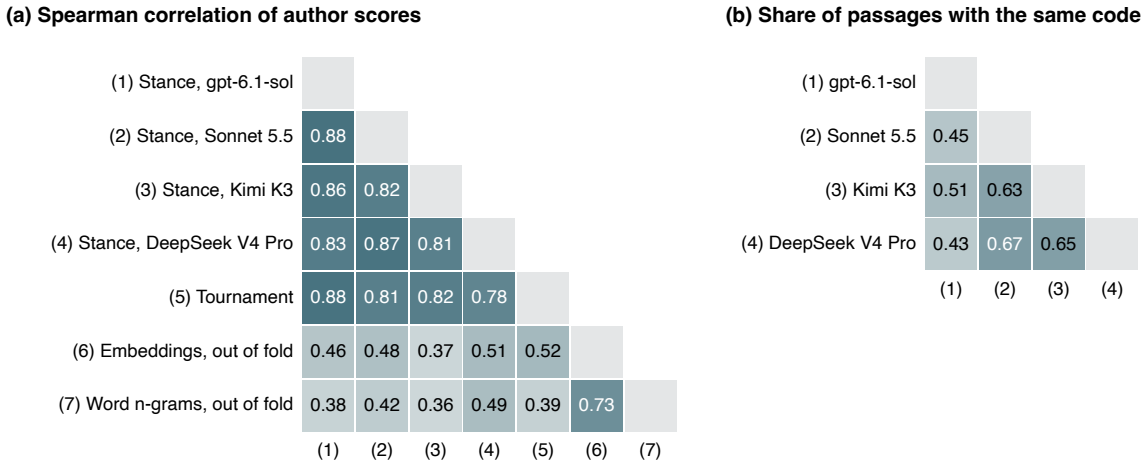
Notes: Reliability of the two tournaments, whose strengths and scores are those of Equations (1) and (2): before 1517 (176 authors, 2,348 passages, 9,392 meetings) and 1517–30 (224 authors, 1,434 passages, 5,736 meetings). Order consistency: the share of meetings in which the two reading orders pick the same passage; a meeting whose orders disagree counts as a tie in the fit. Split-half reliability: each author's passages are split at random into two halves, the two half-mean strengths are correlated across authors with two or more passages, and the correlation is averaged over 200 random splits and stepped up by the Spearman–Brown formula to the full set of passages; corrected for this attenuation the AUC before 1517 would be 0.641. German win share: among decided meetings between a German and a Latin passage (1,465 before 1517), the share won by the German one, with a normal-approximation 95 percent interval; the language is that of the passage's work in the catalogue. Last row: the Spearman correlation across authors between the author score and the share of his passages from German works.

Table C5. Works at the two ends of the scale

Reform end	θ	Passages
Manuel (P), Ein schon bewerts lied von der reynen unbefleckten entpfengnüß Marie, 1510	+7.6	10
Cordus (P), Quattuor hereticoru[m] ex Predicatoru[m] ordine Berne combustoru[m] ap, 1509	+7.3	2
Murmellius (C), JOānis Mur mellij Ruremũdensis Didasca lici libri duo ... Jn quorũ alt, 1510	+7.2	2
Jacobus (C), Correctoriu[m] Biblie Sammlung, 1508	+6.7	2
Busche (P), Spicilegium Philosophorum pene omnium, 1513	+6.7	2
Badius Ascensius (C), Jodoci Badij asce[n]sij Stultifer[a]e nauicul[a]e seu scaph[a]e, Fatua, 1502	+6.6	2
Dottanius (C), Carmen Lysitelilogon de poetices co[m]moditatibus co[n]tra sacrilogos, 1502	+6.4	3
Gengenbach (P), Practica zu teutsch uff das xvc. und new Jar, 1515	+6.2	2
Roman end	θ	Passages
Cyclops (P), Elegion de immaculata conceptione D. Virginis ..., 1511	−7.5	7
Koebel (C), Dialogvs Libertatis Ecclesiastice Defensorius cum Imperatorum sanction, 1516	−7.6	2
Bomhouver (C), (JNstructio et Ordinatio Venerabi lis #[et] Egregij viri dñi Chřiani, 1508	−7.9	6
Scheurl (C), Epistola Doctoris Scheurli ad Charitatem Abbatissam Sanctae Clarae de, 1515	−8.1	2
Sacranus (C), Errores atrocissimorum Ruthenorum, 1507	−8.3	2
Bomhouver (C), Summaria instructio et ordinatio Christiani Bomhouver ... pro executio, 1508	−8.6	4
Zangenried (C), Sermo de Imaginibus et Picturis Ecclisiarum vulgaris Heydelberge habit, 1502	−8.7	4
Bomhouver (C), Jnstructio et Ordinatio p[ro] negotio sacratissimar[um] indulgentiar[u, 1507	−8.7	5

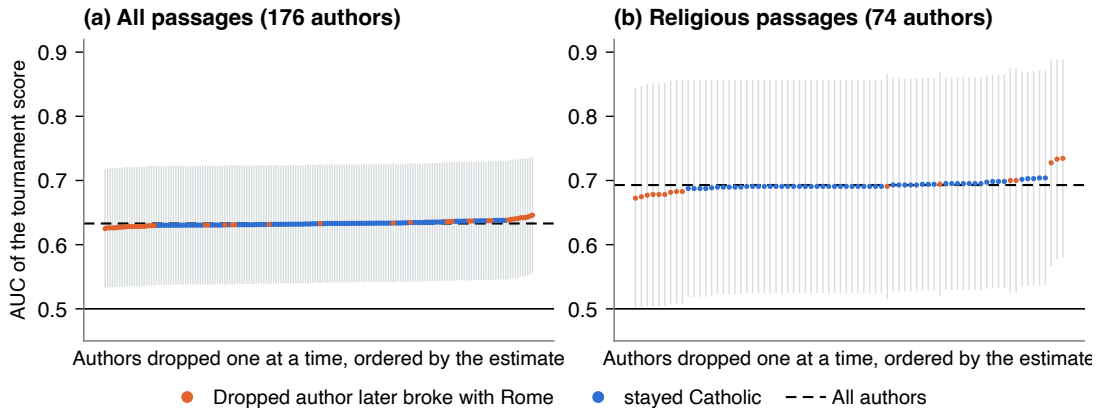
Notes: The works at the two ends of the pre-1517 scale: the eight works with the highest and the eight with the lowest mean passage strength θ among the works of the sample's authors with at least two passages in the fit. $\theta_i = \log \pi_i$ is the Bradley–Terry strength of passage i (Equation (1)), averaged over the passages of the work; the author's score is the mean over his passages (Equation (2)). P marks an author who later broke with Rome, C one who stayed Catholic. Titles are the catalogue titles cut to 70 characters, with the print year.

Figure C1. Agreement across models



Notes: Panel (a): the Spearman correlation across authors between author scores from different measures, on the authors of the pre-1517 sample who have both scores. Stance models: the author's mean overall lean over his chunks as coded by four models given the same chunks and the same prompt (gpt-6.1-sol, the main model of the paper; Claude Sonnet 5.5; Kimi K3; DeepSeek V4 Pro); the tournament score; and the out-of-fold scores of the embedding and word n-gram classifiers (Figure 1, Table C1). Panel (b): the share of the 1,285 chunks read by all four stance models to which two models give the same holistic code. Author-level AUCs of the four models' scores: 0.60, 0.67, 0.62 and 0.63.

Figure C2. Dropping one author at a time



Notes: Each point is the AUC of the tournament score with one author of the sample left out, the dropped author's side as colour, ordered by the estimate; the grey spike is the 95 percent confidence interval (DeLong), the dashed line the AUC with every author, the solid line 0.5. Panel (a): all passages, 176 authors, estimates between 0.625 and 0.646 around 0.633. Panel (b): the score from religious passages on the 74 authors who have one, between 0.672 and 0.734 around 0.693. Scores are not refitted when an author leaves; a refit without his passages moves the range in the third decimal. The coefficients of Table 3 behave the same way: column (1) stays between 0.080 and 0.099, and columns (1) and (4) of Table 3 stay clear of zero at the 95 percent level whichever author is dropped.

C.3 Alternative Explanations

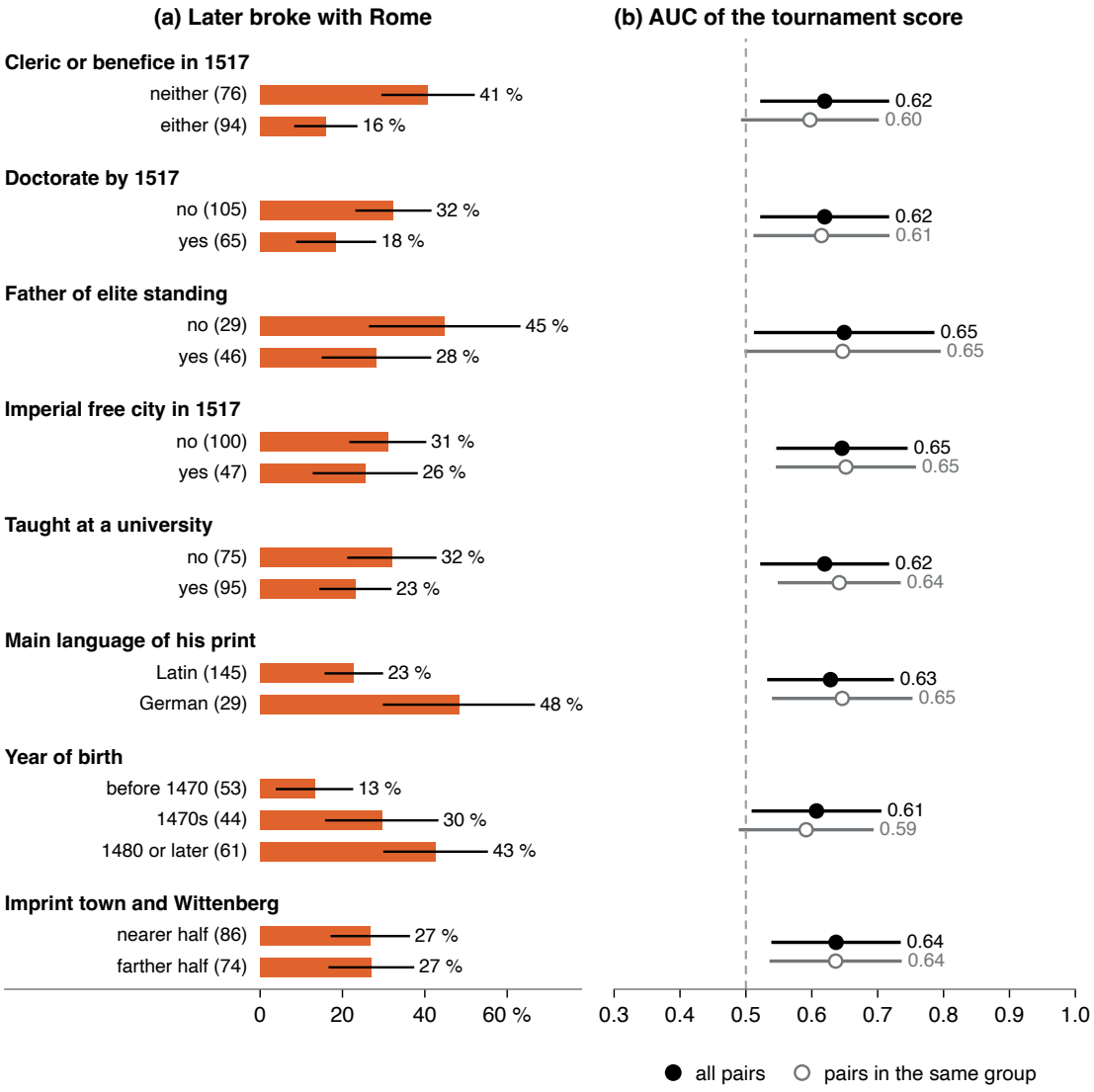
Two tables and a figure stand behind Section 5. Table C6 gives, for eight groupings of the authors, the share of each group that later broke with Rome and the AUC of the tournament and of the embeddings, over all pairs of a convert and a Catholic and over pairs within the same group; Figure C3 draws the tournament columns. Table C7 gives the matched work pairs before 1517 (Section 4.1 and Section 5.3) and in 1517 to 1530, before the convert's break (Section 4.3); the rule is in Appendix A.

Table C6. Breaks and scores within groups

			Tournament: AUC		Embeddings: AUC	
Authors compared within	Later broke with Rome, by group	Authors	all pairs	same group	all pairs	same group
Stakes: cleric or benefice in 1517	neither 41 % (31 of 76); either 16 % (15 of 94)	170	0.62	0.60 ± 0.10	0.65	0.60 ± 0.10
Doctorate by 1517	no 32 % (34 of 105); yes 18 % (12 of 65)	170	0.62	0.61 ± 0.10	0.65	0.65 ± 0.10
Father of elite standing	no 45 % (13 of 29); yes 28 % (13 of 46)	75	0.65	0.65 ± 0.15	0.63	0.60 ± 0.15
Imperial free city in 1517	no 31 % (31 of 100); yes 26 % (12 of 47)	147	0.65	0.65 ± 0.10	0.67	0.65 ± 0.11
Taught at a university before 1517	no 32 % (24 of 75); yes 23 % (22 of 95)	170	0.62	0.64 ± 0.09	0.65	0.66 ± 0.09
Main language of his print	Latin 23 % (33 of 145); German 48 % (14 of 29)	174	0.63	0.65 ± 0.10	0.65	0.64 ± 0.10
Year of birth	before 1470 13 % (7 of 53); 1470s 30 % (13 of 44); 1480 or later 43 % (26 of 61)	158	0.61	0.59 ± 0.10	0.66	0.66 ± 0.10
Imprint town and Wittenberg	nearer half 27 % (23 of 86); farther half 27 % (20 of 74)	160	0.64	0.64 ± 0.10	0.67	0.67 ± 0.10

Notes: Who broke with Rome, by group, and whether the score separates the sides among authors of the same group. Sample: the pre-1517 sample; each row uses the authors for whom the grouping is known (Authors). Stakes: clerical estate or a benefice in 1517, from the biography coding. Doctorate: a doctorate earned by 1517, from the same coding. Father of elite standing: urban elite, office-holder or noble, on the authors whose father's standing is coded. Imperial free city: the type of polity of the author's place of activity in 1517 (home polity where that is unknown), imperial or Swiss confederate city against the rest. Taught: university teaching before 1517. Main language: a Latin share of his pre-1517 works of at least one half. Year of birth in three groups; distance to Wittenberg of the imprint town of most of his works, split at the median. All pairs: the AUC of the score over every pair of a convert and a Catholic in the row's sample. Same group: the AUC over pairs of a convert and a Catholic in the same group, the groups' AUCs weighted by their number of such pairs, with 95 percent confidence intervals (DeLong within each group, combined with the same weights). Tournament: the pre-1517 score. Embeddings: the out-of-fold score of the classifier of Figure 1 fitted with the sample's authors in the folds. An AUC that stays when authors are compared within groups is a signal that the grouping does not explain.

Figure C3. Breaks and AUC within groups



Notes: The same groupings as Table C6, drawn. Panel (a): the share of each group that later broke with Rome, with a normal-approximation 95 percent interval. Panel (b): the AUC of the tournament score over all pairs of a convert and a Catholic in the row's sample, and over pairs within the same group, with 95 percent confidence intervals (DeLong; within groups, combined with the pairs as weights). Sample and definitions as in Table C6.

Table C7. Matched work pairs

Design	Pairs	Convert ahead	Wins	Losses	Ties	p
Work pairs: same language, print year, birth cohort	122	62 %	71	44	7	0.015
Work pairs, same genre added	112	58 %	58	42	12	0.261
Works of 1517–21, before the convert's break	57	67 %	35	17	5	0.039
Works of 1517–25	98	74 %	66	23	9	0.002
Works of 1517–30	128	77 %	92	27	9	<0.001

Notes: Matched work pairs: a convert's work against a Catholic's work of the same catalogue language, print year (within two years) and birth cohort (within five years, when both birth years are known), each Catholic work used once; a work without a partner within these limits is matched within three years and ten. Rows 1 and 2: works printed before 1517; row 1 is the comparison of Section 4.1, row 2 also matches the genre of the works (Section 5.3). Rows 3 to 5: works printed in the window, the convert's before his documented break, matched on the genre as well. The model compares the two works with the tournament prompt (Prompt B1), every passage of the one work against every passage of the other, in both orders, on up to four passages of each work for rows 1 and 2 and on whole chunks for rows 3 to 5; a comparison is a win when both orders put the convert's text nearer reform, a loss when both put the Catholic's, a tie otherwise, and a pair is a win when it has more won than lost comparisons, a loss when fewer, a tie when equal. Comparisons with a passage the model names at confidence 90 or more are left out, and both authors are in the respective sample (Luther's pairs are not). Convert ahead: wins over decided pairs, ties excluded. p: an exact two-sided sign-flip test over the convert authors, so that an author with several pairs counts once. Comparing instead the two authors' tournament scores (Equation (2)) within the pairs of row 1, the convert has the higher score in 75 percent of the 122 pairs (p 0.007).

C.4 After 1517

Four tables and two figures stand behind Figures 5, 6, 7 and 8 of Section 4.3. Table C8 gives the tournament and the embeddings on three windows, 1517 to 1521, 1517 to 1525 and 1517 to 1530, and the linear probability of the break on each window's score, alone and with the controls of Table 3 without the mean print year. Table C9 gives the coefficients of Figure 6, the same model on the 17 converts with text both before 1517 and after their break, and the placebo of 5,000 draws.

Table C10 repeats the measures of Figure 8 with the converts' text split at each convert's break, shown where ten or more converts have such text in a period; its last rows are a check without the model's rubric (Section 4.3). Table C11 asks whether Luther's profile is the direction along which the sides divide: the period's own first component, and the alignment of the gap between the sides with Luther's profile and with Erasmus's. Figure C4 shows the correlations behind each mean of panel (b) of Figure 8 pair by pair, for each side before 1517 and in 1522 to 1530. Figure C5 shows, author by author, the agreement with Luther's profile and the AUCs of Section 4.3.

Table C8. Shorter post-1517 window

	Tournament		Embeddings		Linear probability of the break		
Window	Authors (converts)	AUC ± 95 % CI	Authors (converts)	AUC ± 95 % CI	Score alone	With controls	Authors (converts)
1517–21	137 (51)	0.78 ± 0.08	138 (52)	0.73 ± 0.09	0.238 (0.031)	0.186 (0.037)	121 (45)
1517–25	186 (60)	0.81 ± 0.07	189 (62)	0.71 ± 0.08	0.242 (0.026)	0.206 (0.029)	161 (50)
1517–30	224 (64)	0.80 ± 0.07	227 (66)	0.70 ± 0.07	0.224 (0.026)	0.205 (0.026)	191 (54)

Notes: The text of 1517–30 in three windows, each author's text printed from 1517 to the end of the window and, for a convert, before his documented break. Tournament: the AUC of the author score of the 1517–30 tournament on the authors with text in the window (the shorter windows score the shorter text with the 1517–30 fit's strengths). Embeddings: the cross-validated classifier of Figure 1 fitted on the window's text, all authors with text in the window in the folds. 95 percent confidence intervals (DeLong). Linear probability of the break: Equation (3) estimated by OLS with the window score, standardised over the window's tournament authors, alone and with Latin share, log characters of the window text, year of birth and distance to Wittenberg (the mean print year is left out, since a convert's text stops at his break); both columns on the authors with every control; HC1 standard errors in parentheses. Luther is left out of the tournament but not of the embeddings.

Table C9. Event study coefficients

Years to the convert's own break	5–3 before	2–1 before	0–2 after	3 or more after	Chunks	Authors	Converts
All converts and Catholics	0.65	1.15	1.72	1.74	2,342	280	76
	(0.31)	(0.33)	(0.36)	(0.32)			
Converts with text before 1517	0.95	1.22	1.89	1.93	2,014	221	17
and after the break	(0.39)	(0.39)	(0.39)	(0.34)			
Placebo: Catholics given	0.00	0.02	0.01	0.10	—	—	—
the converts' break years	(−0.45 to 0.44)	(−0.60 to 0.63)	(−0.71 to 0.75)	(−0.76 to 1.03)			

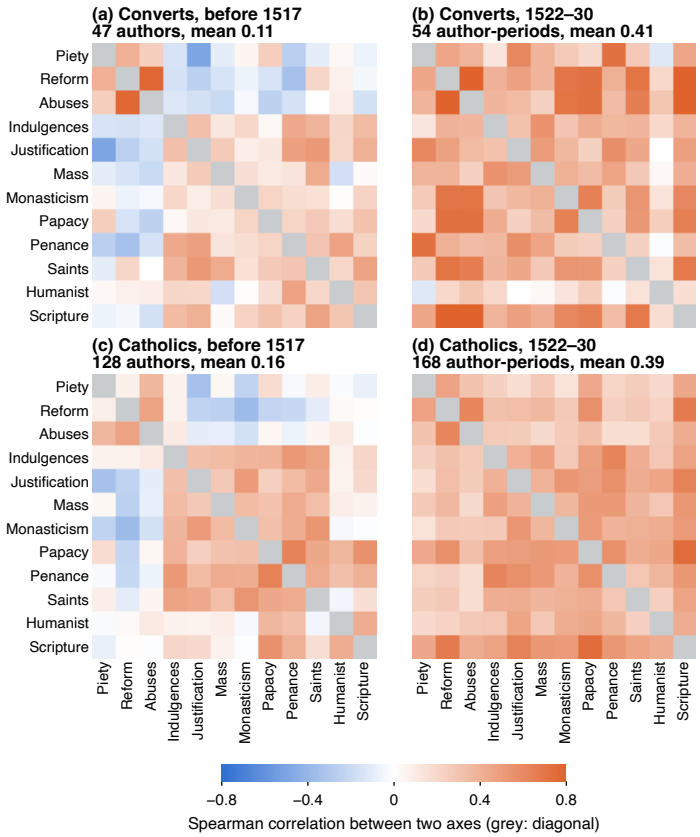
Notes: The event study of Equation (4) and Figure 6 (sample described there). Entries are the coefficients of the four event-time indicators with standard errors clustered by author in parentheses; 6 or more years before the documented break is the reference. Row 1: every chunk left after recognised passages are dropped, with a lean and a print year. Row 2: the Catholics and only the 17 converts with text both before 1517 and after their break, so that every convert is compared with his own earlier text on both sides of his break. Row 3: a placebo in which each Catholic is given a break year drawn with replacement from the converts' break years and the model is refitted on Catholics alone, 5,000 draws; the entry is the mean coefficient over the draws with the 2.5th to 97.5th percentile. Of the draws, 18, 0, 0 and 3 reach the row-1 coefficient in absolute value.

Table C10. Coordination without post-break texts

	Before 1517	1517–19	1520–21	1522–25	1526–30
Authors with text					
Catholics	128	55	51	75	93
Converts, all text	47	31	33	32	22
before their break	47	31	29	19	10
after their break	0	1	5	14	13
Mean lean					
Catholics	−0.40	−0.06	−0.04	−0.40	−0.78
Converts, all text	−0.07	0.61	1.07	1.43	1.39
before their break	−0.07	0.61	0.97	1.32	1.31
after their break	—	—	—	1.61	1.50
Correlation across the twelve axes					
Catholics	0.16	0.24	0.24	0.38	0.38
Converts, all text	0.11	0.09	0.29	0.40	0.39
before their break	0.11	0.09	0.27	0.39	0.23
after their break	—	—	—	0.39	0.40
Similarity to Luther's profile (cosine)					
Catholics	−0.03	−0.14	−0.31	−0.50	−0.44
Converts, all text	0.12	0.08	0.16	0.35	0.32
before their break	0.12	0.08	0.17	0.33	0.20
after their break	—	—	—	0.40	0.44
AUC: how often a convert stands nearer Luther than a Catholic					
Agreement with Luther's profile, all text	0.64	0.69	0.76	0.87	0.89
converts' texts before their break only	0.64	0.69	0.77	0.86	0.88
Line from Eck to Luther in embedding space, language held fixed, all text	0.59	0.58	0.51	0.70	0.73
converts' texts before their break only	0.59	0.57	0.48	0.71	0.79

Notes: The measures of Figure 8 with the converts' text split at each convert's own documented break. Authors with text: converts and Catholics with a stance-read chunk left after recognised passages are dropped in the period; the two indented rows count the converts whose chunks before (after) the break fall in the period. Mean lean, correlation across the twelve axes and similarity to Luther's profile are defined as in Figure 8; the centre of each period (the midpoint of the two sides' mean profiles) is taken from all text, so the indented rows are measured against the same centre as the rows above them. A split row is shown when ten or more converts have such text in the period. AUC rows: the share of pairs of a convert and a Catholic in which the convert stands nearer Luther. Agreement with Luther's profile is the projection of Figure C5. The line from Eck to Luther in embedding space is a check without the model's rubric: each author's embedded text of the period is projected on the direction from Eck's to Luther's 1517–21 texts after the language of the text has been partialled out; the second AUC of each pair uses only the converts' chunks before their break.

Figure C4. Correlation between axes



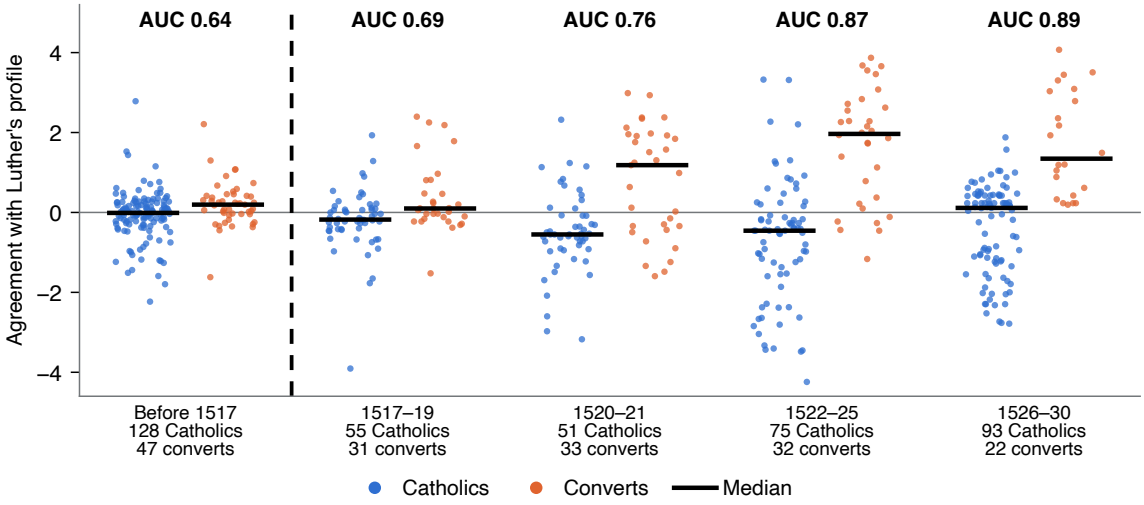
Notes: Panels (a) to (d): Spearman correlation between the positions on two axes across the authors of a side: one profile per author before 1517 (47 converts, 128 Catholics) and the author-period profiles of 1522–25 and 1526–30 pooled (an author who printed in both periods enters twice). An author's position on an axis is his mean code over his chunks of the period (−2 Roman to +2 reform, silence counted as 0). The mean of the 66 cells of a panel is panel (b) of Figure 8; a cell is grey where an axis does not vary in the panel. Colour from −1 to +1.

Table C11. Luther's profile and the main direction

Measure	Before 1517	1517–19	1520–21	1522–25	1526–30
Share of all variation on the period's own first component	0.32	0.40	0.49	0.60	0.56
AUC of the period's own first component	0.59	0.66	0.76	0.89	0.90
AUC of the projection on Luther's 1518–21 profile (Figure C5)	0.64	0.69	0.76	0.87	0.89
Alignment of the convert–Catholic gap with Luther's 1518–21 profile (18 chunks)	0.70	0.69	0.80	0.93	0.95
Same, with Erasmus's 1517–21 profile (5 chunks)	0.75	0.84	0.32	0.42	0.75
Authors (later converts)	175 (47)	86 (31)	84 (33)	107 (32)	115 (22)

Notes: Test of whether Luther's profile is the direction along which the sides divide. Authors with stance-read text in each period, as Figure 8. Row 1: the share of the variation of the authors' centred twelve-axis profiles that lies on their first principal component. Row 2: the AUC of the authors' scores on that component (it has no sign, so the larger of the AUC and its complement). Row 3: the AUC of the projection on Luther's 1518–21 profile, the measure of Figure C5. Rows 4 and 5: the cosine between the difference of the two sides' mean profiles and the reference profile, both measured from the period mean: Luther's profile from his writings of 1518–21 (18 chunks) and Erasmus's from his of 1517–21 (5 chunks). From 1520 to 1525 the gap points to Luther and no longer to Erasmus (it points to Erasmus again in 1526–30), and the period's own main direction separates the sides no better than Luther's does.

Figure C5. Agreement with Luther by author



Notes: One dot per author and period: the author's twelve-axis profile, centred at the mean profile of all authors of the period, projected on Luther's profile from his writings of 1518–21 (18 chunks): the sum of the author's centred codes, each weighted by Luther's code on that axis, divided by the length of Luther's profile. A dot is high when the author takes the reform side on the questions Luther pressed hardest. Same authors, chunks and periods as Figure 8; dots are spread horizontally within a period for legibility. The AUC above each period is the share of pairs of a convert and a Catholic in which the convert's dot is the higher one; the 95 percent confidence intervals (DeLong) are 0.55 to 0.73, 0.58 to 0.81, 0.65 to 0.88, 0.80 to 0.94 and 0.82 to 0.96. With the converts' texts before their break only, the AUCs are 0.64, 0.69, 0.77, 0.86 and 0.88 (Appendix Table C10).

C.5 Where the Signal Sits

Panel (b) of Figure 3 splits the gap in the mean score into a part that comes from which genres the authors wrote in and a part that comes from how they wrote within a genre. The split is the one Kitagawa (1955) gave for the difference between two rates, which Oaxaca (1973) and Blinder (1973) brought into economics for the difference between two wages; Fortin et al. (2011) survey the family. Here the rate is the mean score, and the groups are the genres.

Author i has a share s_{ik} of his tournament passages in genre k and a mean strength t_{ik} over them, so his score is $S_i = \sum_k s_{ik}\, t_{ik}$. For a side c with N_c authors (Catholics 0, later converts 1), let w_k^c be the side's average share of text in genre k and m_k^c its mean score inside the genre, each author's passages weighted by his share:

$$w_k^c = \frac{1}{N_c} \sum_{i \in c} s_{ik}, \qquad m_k^c = \frac{\sum_{i \in c} s_{ik}\, t_{ik}}{\sum_{i \in c} s_{ik}}.$$

The side's mean score is then $\bar{S}^c = \sum_k w_k^c\, m_k^c$ exactly, and the gap is $\Delta = \bar{S}^1 - \bar{S}^0 = \sum_k (w_k^1 m_k^1 - w_k^0 m_k^0)$. Each genre's term splits into a share part and a mean part. We evaluate each at the midpoint between the two sides, as Reimers (1983) does, so that neither side serves as the reference:

$$\Delta = \sum_k \underbrace{(w_k^1 - w_k^0)\left(\frac{m_k^1 + m_k^0}{2} - \bar{m}\right)}_{B_k,\ \text{between genres}} + \sum_k \underbrace{(m_k^1 - m_k^0)\,\frac{w_k^1 + w_k^0}{2}}_{W_k,\ \text{within genres}},$$

with $\bar{m} = \sum_k \frac{w_k^1 + w_k^0}{2} \cdot \frac{m_k^1 + m_k^0}{2}$ the mean over genres, each weighted by its average share of text. The identity holds without $\bar{m}$. Subtracting it changes no total, because the shares sum to one on each side; it only stops a genre's between part from depending on where the strength scale puts its zero. The figure shows B_k/Δ and W_k/Δ. For religion, $w^1 = 0.25$, $w^0 = 0.37$, $m^1 = 0.36$, $m^0 = -3.03$ and $\bar{m} = 0.23$, which gives $B = 0.20$ and $W = 1.05$ of $\Delta = 1.35$: the Catholics' extra 13 points of text in a genre that scores 1.56 below the mean pull their average down by 0.20, and the sides' difference of 3.39 inside religion, on 31 percent of the text, adds 1.05. Over 1,000 resamples of authors within each side, the 95 percent interval of Δ is 0.41 to 2.25, of W_{rel} 0.29 to 1.74 and of B_{rel} −0.01 to 0.48; the within part of religion is positive in 99.6 percent of the resamples, the between part in 97 percent. The shares W_k/Δ and B_k/Δ are ratios of two estimates, and the denominator is itself uncertain; their intervals are therefore wide and skewed (for the within part of religion, 31 to 149 percent: in some resamples it exceeds the whole gap, because the parts

of the other genres come out negative). The figure reports the shares as point estimates and the intervals on the parts. On the model's overall lean the within part of religion is 93 percent of the gap.

Gentzkow et al. (2019b) ask the same question of the speech of Congress, whether the parties differ in the topics they speak on or in the phrases they use within a topic, and answer it with a different tool: how much each source alone tells an observer about the speaker's party. We can ask that too. Score every passage twice, at the mean strength of its genre, μ_k, and as its distance from that mean. Averaged over an author's passages, the two scores split his own exactly: $S_i = C_i + W_i$, where $C_i = \sum_k s_{ik}\,\mu_k$ is his genre mix and $W_i = \sum_k s_{ik}(t_{ik} - \mu_k)$ is how he wrote within genres. Table C12 gives the AUC of each. The genre mix alone ranks a later convert above a Catholic with probability 0.585. The within part gives 0.61. The two are uncorrelated across authors (Spearman 0.06), which is why together they reach the score's 0.63. Within genres, the separation is religion's: 0.70, against 0.56 and 0.56 in the other two. Of the 44 authors who wrote nothing but religion, 36 stayed Catholic. The genre mix sorts the authors; it moves the average score little.

Table C12. Genre mix and writing within genres

	Tournament	Overall lean
All text	0.63	0.59
Genre mix alone (C_i)	0.58	0.57
Within genres, all (W_i)	0.61	0.57
within religion	0.70	0.73
within literature and education	0.56	0.52
within law, medicine, science	0.56	0.45

Notes: AUC of each author-level quantity over the 176 authors of the pre-1517 sample (48 later converts); the overall lean over the 171 authors with a coded chunk. Genre mix: each passage (chunk) scored at the mean of its genre, then averaged over the author's passages. Within genres: each passage scored as its distance from its genre's mean. The two sum to the author's score. The within rows by genre use the author's passages in that genre only, on the authors who have any (Appendix C.5).

One trace runs outside religion (Figure 4). In law, medicine and science, Catholic texts more often ground a practical matter in the church's authority, the date of Easter in the decrees of the Fathers, a point of law in the canons, the order of Europe's princes in the pope (AUC of Scripture against tradition 0.59, of papal authority 0.57); it rests on a handful of authors.

The model's own reasons show what it rewards. Every meeting of the tournament returns one sentence that states what decided it (Prompt B1). Table C13 counts these sentences for the 6,080 meetings in which both reading orders return the same verdict

and the reason has one clause on each passage, and sorts the clauses by theme. The winning passage is the one the model puts nearer reform. The label plays no part: the table describes the model's rule, not the converts.

Table C13. The model's reasons

The model's reason cites ...	in the passage that wins	in the passage that loses
Humanist learning, classical letters	27 %	5 %
Faith, grace, inward Christ-centred piety	13 %	4 %
The clergy: abuses, greed, conduct	9 %	7 %
Scripture	4 %	2 %
Saints, Mary, relics, pilgrimage	5 %	28 %
Scholastic method and its authorities	8 %	17 %
Monastic life and vows	7 %	10 %
Pope and councils	2 %	8 %
Indulgences, purgatory, merit	1 %	8 %
Mass, sacraments, confession	2 %	7 %
Meetings	6,080	

Notes: Every tournament meeting asks the model which of two masked passages leans further toward reform and returns one sentence that names what decided it (Prompt B1). Counted are the 6,080 of the 9,392 meetings of the pre-1517 fit in which both reading orders return the same verdict and the reason has one clause on each passage. A keyword rule assigns each clause to the themes listed; a clause can name several themes, and a clause that names none is not counted, so a column does not sum to 100. The keyword rule is a fixed list of terms for each theme; a clause names a theme when it contains one of them. The first column is the share of the winning passages, the one the model puts nearer reform, whose clause names the theme; the second column the same for the losing passage. The themes are grouped by the side on which the model names them more often: the first group characterises the passages it puts nearer reform, the second those it puts nearer Rome.

What wins is learning and inward piety. The model cites humanist learning in 27 percent of the winning passages and 5 percent of the losing ones, faith, grace and inward Christ-centred piety in 13 against 4. The clergy's conduct (9 against 7) is cited almost as often on both sides.

What loses is the defence of the church of 1500: the cult of saints, Mary and relics (28 percent of the losing passages against 5 of the winning ones), scholastic method (17 against 8), monastic life (10 against 7), pope and councils (8 against 2), indulgences (8 against 1). The passage the model puts nearer Rome upholds devotion, school, pope or indulgence; the passage it puts nearer reform shows learning and inward piety.