In the Name of God!

Religiosity and the Emergence of Modern Science and Growth

Lars Harhoff Andersen and Jeanet Sinding Bentzen*

Department of Economics, University of Copenhagen

June 2023

First draft: May 2021

Abstract

We examine the role played by religion in the emergence of modern science and economic growth. To do so, we develop novel measures of religiosity based on a literature in psychology and sociology, which shows that given names can reflect cultural identities of parents. To validate our measures, we find that among 487,000 individuals born in Europe between 1300 and 1940, those with religious names were more likely to hold religious occupations, to be born after earthquakes and in areas where surveys indicate higher levels of religiosity. Next, we document that individuals with religious names were less inclined to become scientists, engineers, or to engage in advanced studies. For identification, we exploit exogenous shocks to migrants' religiosity by earthquakes. Last, we find that declining religiosity is associated with rising economic growth. Results hold in an alternative dataset, accounting for a host of confounders, and comparing individuals born within the same 1x1 km^2 and year. The results corroborate a literature arguing that religion may hamper the development of modern science, which became increasingly important for economic growth after the technological revolution.

JEL codes: Z12, N33, I25.

Keywords: Science, economic growth, religiosity, knowledge, human capital, religion.

*Lars Harhoff Andersen, Department of Economics, University of Copenhagen. Jeanet Sinding Bentzen, Department of Economics, University of Copenhagen, CEPR, CAGE, SODAS, jeanet.bentzen@econ.ku.dk. We are thankful for helpful comments from Gustav Agneman, Sascha Becker, Roland Benabou, Karol Jan Borowiecki, Matías Cabello, Stuart Campbell, David de la Croix, Scott Cunningham, Carl-Johan Dalgaard, James Fenske, Roberto Galbiati, Oded Galor, Ed Glaeser, Claudia Goldin, Pauline Grosjean, Sergei Guriev, Mark Harrison, Latika Hartmann, Erik Hornung, Michael Jetter, Anne Sofie Beck Knudsen, Stelios Michalopoulos, Joel Mokyr, Kevin O'Rourke, Claudia Rei, Pablo Selaya, Moses Shayo, Uwe Sunde, Hans-Joachim Voth, Asger Mose Wingender, Edith Zink, Ekaterina Zhuravskaya, as well as comments from seminar participants at (when nothing else is indicated, the seminar was given at the Department of Economics) University of Sao Paolo 2023, the 2023 ASREC conferences at Harvard and Monash, Monash University 2022, Melbourne University 2022, the Natural Experiments Workshop at Deakin University 2022, the AusClio Conference at Canberra University 2022, the Quantitative Political Economy workshop at King's College London 2022, the Roots of Comparative Development conference at Brown University 2022, the Conflict and Development workshop at University of Southern Denmark 2022, Deakin University 2022, the PEPES combined Paris School of Economics and SciencesPo seminar 2021, Bergamo University 2021, Arne Ryde Workshop at Lund University 2021, the University of Aarhus Department of the Study of Religion 2021, University of Copenhagen Department of Sociology 2021, University of Warwick CAGE Economic History Workshop 2021, University of Copenhagen 2020, University of Aarhus DGPE Workshop 2020, University of Copenhagen DGPE Workshop 2019, University of Copenhagen MEHR Workshop 2019. We are grateful for excellent research assistance by Aisha Latif, Vladyslav Masenkov, and Abdelrhman Eldallal. Jeanet Bentzen is thankful for research funding from the Independent Research Fund Denmark.

1 Introduction

Modern economic growth began only a few centuries ago and was characterized by a rise in the rate of innovation in science and technology.[1] There is ample reason to believe that these factors could have been influenced by religion. On the one hand, many universities began as religious institutions and multiple medieval innovations were invented or spread through the vast networks of monasteries. On the other hand, a recent literature argues that religion might have been a stumbling block for the emergence of modern science. Thus esteemed economic historian, Joel Mokyr (2016a, 17) reasons that "*If the culture is heavily infused with respect and worship of ancient wisdom so that any intellectual innovation is considered deviant and blasphemous, technological creativity will be similarly constrained.*" History abounds with examples of clashes between religion and science and also examples of their peaceful co-existence. The role played by religion, therefore, remains an empirical matter, which we set out to address.

A substantial body of literature before us has examined the impact of types of religion for economic outcomes, identifying differences between Protestants and Catholics (Weber, 1905; Becker & Woessmann, 2009), between Muslims and Christians (Kuran, 2018; Rubin, 2017), or Jews and Christians (Becker & Pascali, 2019; Carvalho & Koyama, 2016), for instance. We move beyond this work by focusing on the impact of *religiosity*, ie. the intensity of peoples' beliefs or the significance of religion in their lives, rather than the type of God they worship. Early contributions include Iannaccone (1998), Barro & McCleary (2003), and Guiso *et al.* (2003) exploiting measures of religiosity from surveys to examine its' contemporary socioeconomic consequences.[2]

To obtain proxies for religiosity at a time before surveys, we rely on a literature within psychology, anthropology, and sociology arguing that given names may reflect the identities of parents (Lieberson, 2000; Mateos, 2014). Researchers before us have shown that given names reveal parental characteristics, such as their degree of individualism (Knudsen, 2019; Bazzi *et al.*, 2020), ethnicity (Fryer & Levitt, 2004; Andersen, 2021), or nationalism (Jurajda & Kovač, 2021; Assouad, 2020). We use individuals' first names to infer the significance of religion in the lives of their parents and potentially their upbringing. We propose the hypothesis that while name choice is influenced by multiple factors, people for whom religion is central are on average more likely to choose religious names for their children, as opposed to names that might reflect preferences for nationalism, for

[1] Jacob (1997); Mokyr & Voth (2009); Mokyr (2010); Squicciarini & Voigtländer (2015).

[2] Cf reviews of the literature on religiosity by Iannaccone (1998); Iyer (2016); Bentzen (2021b). Global variation in religiosity is not explained by types of religion. For instance, Bentzen (2019b) documents that the major religious denominations explain only 3.4% of the global variation in religiosity.

instance. We test this hypothesis using historical data. As our testing ground is historical Europe – the cradle of modern economic growth, – we focus on Christianity.

We propose two distinct definitions of a (Christian) religious name. Our main measure defines a name as religious if it is shared by a significant Christian figure. The other is based on a data-driven approach that defines a name as religious if it is over-represented among individuals holding religious occupations, such as theologians or priests. Given the inherent complexity of naming choices, our measures should not be interpreted as precise indicators of an individual's personal level of religiosity. In addition, religion is a complex construct with multiple dimensions, many of which naturally escape naming.[3] Our empirical analysis aims to assess the extent to which the measures provide valuable insights into the role played by religion in the lives of ones' parents *on average* across large samples of individuals. To do so, we construct two independent historical databases of 487,000 authors born in Europe between 1300 and 1940 and 47,000 university students in the Holy Roman Empire.[4] These databases hold information such as author names and occupations, their productivity, name origins, birth and death year, geocodes of birth- and deathplace, and geographic features such as earthquakes.

Our analysis relies on a two-way-fixed effects research design, where the variation comes from comparing individuals born in the same year and 1x1 km^2 grid cell. This design allows us to rule out various confounding factors. First, differences across 1x1 km^2 grid cells (e.g., naming practices may differ between rural and urban locations); second, differences across time that affect individuals within a grid cell in the same way (e.g., secularization trends or regional changes in schooling);[5] third, reverse causality (e.g., your occupation may influence your religiosity);[6] fourth, differences between authors or students and the rest (e.g., farmers may have different naming practices compared to engineers);[7] fifth, differences due to gender, ethnicity, religious denomination, migration

[3]No single measure can capture all dimensions of religion. Nevertheless, scholars before us have measured religiosity using surveys (Barro & McCleary, 2003; Guiso *et al.*, 2003), behavior such as church attendance, donations, or Google searches (Stark & Glock, 1970; Spilka *et al.*, 2003; Gibson, 1989; Gruber & Hungerman, 2008; Bentzen, 2021a), interviews (Hill & Pargament, 2003; Hodge & Holtrop, 2002), biological measures, such as brain stimuli during prayer (Newberg & Waldman, 2009; Miller *et al.*, 2019), or religious activity revealed by light during Ramadan (Livny, 2021). We follow this tradition with a novel approach, potentially useful across time and space.

[4]The individuals in the first database are not only authors; they have several other occupations, such as scientists and priests. They enter the database if they ever wrote a text that ended up in a library.

[5]These two are taken into account in a specification that includes grid cell by year fixed effects.

[6]Your occupation later in life does not influence your name. There are examples of individuals changing name to suit their occupation. While this is not a concern for the validity checks, we document that this cannot explain results in the second part of the analysis either in Section 4, where we further rule out explanations based on birth order effects.

[7]These factors are accounted for as the samples include rather similar types of individuals; either authors or university students.

status, as well as parental preferences for individualism, conformity, or nationalism (e.g., naming practices may differ across Protestants vs Catholics);[8] sixth, parental socioeconomic status revealed by their naming practices (e.g., educated parents may name their children differently and at the same time influence their children's occupation choices).[9] In addition, we complement the two-way-fixed effects model with specifications that exploit shocks to religiosity caused by earthquakes.[10] The latter makes unobserved characteristics, such as parental characteristics, an unlikely explanation for the results. Earthquakes influence society in multiple ways, thus introducing a bias of its own. We remove physical effects of earthquakes by focusing on earthquakes in migrants' birthplace.

Lending credibility to the usefulness of the names-based religiosity measures, we document that individuals who share name with a significant religious figure were more likely to i) partake in occupations or studies associated with religion, and to be born ii) after earthquakes, iii) in areas where respondents in contemporary surveys reveal higher religiosity levels, and iv) in areas of France where the clergy was more loyal to the church.[11] The size of the effects are economically meaningful. For instance, sharing name with a prominent religious figure raises the probability of having a religious occupation by 14% of the mean or studying theology in the Middle Ages by 15% of the mean.

The names-based measures of religiosity have certain advantages to alternatives. First, naming our offspring is a universal human endeavor and nearly everyone throughout the history of modern humans had a first name. Second, compared to measures based on surveys, names-based measures reveal the *actual* relative preferences of parents instead of their *stated* preferences and is not subject to non-response bias. Third, since first names reflect the religiosity of the previous generation, we can treat it as pre-determined relative to outcomes later in life. Fourth, as names do not have a direct economic cost, they are less influenced by differences in wealth than other candidates. Alternative proxies, such as church density reflect wealth and power as much as religiosity (Buringh *et al.*, 2020). Fifth, the names-based measures allow for individual-level analysis of the impact

[8] Examining differences across Protestants and Catholics is of particular interest, given the voluminous literature starting with Weber (1905) that has documented differences in socioeconomic status between Catholics and Protestants. A priori, we have reason to believe that our results are not caused by such differences. For instance, Bentzen (2021b) shows that the correlation between contemporary religiosity and GDP per capita is 10 (3) times larger than that between the share of Protestants (Muslims) and GDP per capita. These correlations prove nothing causal, but speak to the relevance of examining religiosity.

[9] Inspired by Fryer & Levitt (2004)'s "blackness of a name" measures, we calculate the "urbanness" of a name based on the relative frequency of a name in urban locations. Likewise for nobles, advanced university students, and areas with high education levels.

[10] We exploit previous research documenting rising religiosity in the aftermath of earthquakes, e.g., Sibley & Bulbulia (2012); Bentzen (2019a). The research is based on the religious coping hypothesis, which states that people tend to use their religion to cope with adversity (Pargament, 2001).

[11] The latter involves a measure of religiosity used by Tackett (1986) and Squicciarini (2020).

of religiosity, which greatly increases our chances of estimating causal effects. Sixth, the measures can be computed for large samples of similar individuals, which further facilitates their comparability. A similar method was used by Hofstede *et al.* (2005) who compared cultural differences across countries by focusing on IBM programmers.

We proceed to determine the impact of religiosity on individual-level outcomes. We document that authors with religious names are less likely to become scientists or engineers and that religiously named university students were less likely to pursue advanced studies. These results emerge when defining religious names based on significant religious figures or the data-driven approach. Again, the results are robust to including controls such as proxies for the socioeconomic status of parents and restricting comparison to authors or students born in the same year and square kilometer. The size of the effects are economically meaningful. Persons with religious names (who thus potentially had a more religious upbringing) are 6% of the mean less likely to become scientists or engineers and 3-8% of the mean less likely to take an advanced degree in the Middle Ages.

These samples of similar individuals raise concerns about whether the samples are representative for the general population and whether selection into the samples may bias results. We judge such concerns to be unlikely to compromise our analysis. First, the share of authors with religious names is higher in regions where survey responses and the loyalty of the clergy reveal higher religiosity. This increases our confidence that they are representative at the regional level. Second, our samples represent the knowledge elite who were the likely drivers of the innovation and technological progress conducive to modern growth (Mokyr, 2010; Squicciarini & Voigtländer, 2015). Thus, samples of authors or university students are preferred to many alternatives. Third, we can rule out selection bias as an explanation for the results. For instance, if scientists and persons with religious names are more likely to become authors, a concern is that this could explain some of the results. However, we can eliminate the majority of this potential bias by removing individuals with religious occupations.[12] Thus, if selection drives results, it is of a rather unusual type based on names that are unrelated to authors' religious occupation. Furthermore, if selection bias is responsible for our results, we would expect that results strengthen if we make selection worse. The data does not support this. Last, the analysis leveraging exogenous shocks to religiosity removes selection bias altogether.

Another potential concern is that our names-based measures capture other factors

[12]The logic is that if selection drives results, we would expect that the majority of the selection based on religious names would be associated with religious occupations. And thus, removing authors with religious occupations should greatly reduce estimates. This prediction is not borne out in the data. Find more details in Section 4.

that influence occupations later in life, independently of religion. For instance, parents who have a general preference for tradition or conformity, unrelated to religion, could discourage their children from pursuing careers in science. If the names that we define as religious simply capture a preference for tradition, this could explain the science results. If scientific and religious occupations are substitutes, it could also explain the religious occupation results. This type of explanation is unlikely to explain the results. First, the analysis leveraging exogenous shocks to religiosity is difficult to explain with a general preference for tradition. Second, the most common non-religious names tend to be the names of the royalty and thus are no less traditionalist than religious names. Third, individuals with royal names are not more likely to hold religious occupations which we would have expected if religious and royal names reflect the same dimension of identity. Fourth, results hold when adjusting the names-based measures for name commonness, accounting for name commonness independently, or excluding most or least common names. These exercises strengthen our confidence that we capture something specific to *religiosity* per se.

In an attempt to determine the overall impact of religiosity, our final analysis investigates the relationship between the proportion of authors with religious names and economic growth at the city level. Previous research has emphasized the crucial role of scientists and engineers in facilitating the transition towards modern economic growth. For instance, Mokyr & Voth (2009, 29) note that "*The technological changes of the 19th century created a demand for highly skilled mechanics and engineers in the upper tail of the distribution.*" If individuals with a more religious upbringing are on average less likely to become scientists and engineers, as our results suggest, we expect a negative impact of religiosity on economic growth when these occupations become increasingly important for economic growth. We find that European cities with a larger share of religiously named authors grew at a slower pace than other cities. The association increases in size and significance over time, especially as we pass the Enlightenment and the Technological Revolution. A rise in religiosity by one standard deviation is associated with a reduction in city growth by a third. Consistent with the mechanisms envisioned, we find that religious cities had fewer inventors.[13] While causality is better identified in the individual-level analyses, the city-level results should be interpreted with more caution. For instance, we cannot rule out that unobservable factors causing rising economic growth and declining religiosity explain results.

[13] As a consistency check of these data and specifications, we confirm results in the existing literature that cities with more innovators or more scientists experienced larger growth rates.

We contribute to research examining the consequences of differences in religiosity – reviewed in more detail in Section 5. Research most related to ours has shown that school curriculum was less technical in more religious districts in 19th century France (Squicciarini, 2020), that innovation rates are lower in more religious societies across the globe today (Bénabou *et al.*, 2022), and that science fell behind in medieval Islamic societies with stronger religious institutions (Chaney, 2016). We contribute by identifying the causal impact of religiosity for the emergence of modern science and growth across Europe. Also, we provide the first historical European-wide measure of religiosity at the yearly and individual level which could be foundational in future research and which could be used to compute religiosity measures for all religions.[14]

More broadly, we relate to a literature that has grouped the so-called deep determinants of economic outcomes into geographic factors (Diamond, 1997; Olsson & Hibbs Jr, 2005), cultural values (Giuliano & Nunn, 2021; Landes, 1998; Spolaore & Wacziarg, 2013), and institutions (Acemoglu *et al.*, 2001; North, 1991). The impact of religion could work through the cultural values of individuals or the institutions that surround them. As our measures vary at the individual level, we can disentangle the individual-level impact from that of institutions. For instance, the association between religious names and occupations is remarkably stable when adding more and more detailed spatial fixed effects. As institutions vary mostly at the country or regional level, this informs us that a non-negligible part of the impact of religiosity on occupation choices occurs at the individual level. Last, we relate to a literature by sociologists and historians that explores the reasons for the decline in religiosity and the rise of secularism. For instance, Stark & Finke (2000) and Bruce (2002) both emphasize the church's failure to adapt to new ideas and technologies as a main explanation. This is our motivation for the choice of econometric framework designed to identify the causality running from religiosity to science.

Our paper is structured as follows. In Section 2, we describe and present basic tests of the validity of our names-based religiosity measures. Section 3 examines the impact of religiosity for the development of science. Section 4 addresses additional concerns and identifies the causal effects of religiosity on occupation choice. In Section 5, we review the literature on science, religiosity, and economic growth and examine the impact of religiosity for the emergence of modern growth at the city level.

[14]As pendents to Catholic saints, the Jewish tzadik, the Islamic walī, the Hindu rishi or Sikh guru, the Shintoist kami, and the Buddhist arhat or bodhisattva are also referred to as saints. The pendent to Biblical figures would be central figures in the Quran, Torah, etc.

2 Measuring historical religiosity

We set out to measure the intensity of religious beliefs from the late Middle Ages to the early 20th century. That is, from a period in European history when religion was at the center of most aspects of life to one where large parts of the continent were secularized. Though belief in a higher Deity may have been close to universal in all but the very end of the period, the role of religion differed significantly across individuals and regions. At the time Luther (1484-1546) began the Protestant Reformation, Machiavelli (1469-1527) declared religion to be man-made, and while deism spread in the 18th century, the same period also saw the great awakening in Britain and its colonies.[15] While some areas saw religious awakenings, in others the churches stood largely empty. We propose a measure to capture how much personal importance individuals found in religious belief.

While multiple measures of religiosity exist for the contemporary world (e.g., reviewed in Bentzen (2021b)), the best other measure of historical religiosity that we know of is a measure of (Catholic) religiosity across French departments in 1791, measured by the reluctance of French clergy to swear an oath to the secular government (Squicciarini, 2020). We suggest a measure of the religiosity of individuals born throughout Europe during the period 1300-1940.

Anthropologists, sociologists, and psychologists have documented that given names reflect cultural identities of parents.[16] When naming their children, parents are inspired by tradition, people in popular culture, aesthetics, and by feelings of identity. This identity can relate to factors such as ethnicity, ideology or religion, for instance. Studies of colonial American child-naming patterns suggest that parents invested a great deal of thought into naming their children, and that their choices reflected attitudes about themselves and the underlying values of society.[17] Parents would have to choose between naming their child after grandparents, religious or political figures, for instance. While some of these sources of inspiration may overlap, the final choice reveals the relative importance of the particular source. Nationalistic parents may give their child a name of national significance, while parents for whom religion is more important, are potentially more likely to give their child a name of religious significance. We hypothesize that

[15] Historical record suggests that even though there is no evidence that atheism played any important role for the majority of the period, the level of piety and religious intensity varied extensively between regions and individuals (Strauss, 1975).

[16] Lieberson (2000); Mateos (2014); Edwards & Caballero (2008); vom Bruck & Bodenhorn (2006); Carsten *et al.* (2004); Smith (1994).

[17] Main (1996); Smith (1984, 1985); Tebbenhoff (1985); Zelinsky (1970). In many traditions, grandparents were the main inspiration for naming. In that case, ones' name instead potentially reflects the identity of grandparents, great grandparents, etc. This practice means that we can proxy religiosity even further back in time. Our datasets enable us to test the average lag of the naming process.

naming patterns can provide a signal of the average religiosity of parents. Furthermore, if religiosity is (partly) passed on from parent to child, names also give a signal of the religiosity of the child.[18]

Scholars before us have used first names to proxy worldviews that are otherwise difficult to measure. Knudsen (2019) used the relative scarcity of first names to approximate individualism of parents. Fryer & Levitt (2004) computed the "blackness" of first names to measure the strength of black ethnic identity. Andersen (2021) generalized the method to proxy ethnic identities across Europe. Assouad (2020) exploited first names to examine the impact of Atatürk's nation-building efforts. Other scholars have used names to reflect assimilation efforts of migrants (Abramitzky *et al.*, 2020; Fouka, 2020; Saavedra, 2021). The hypothesis that first names may signal the religiosity of parents has been previously elaborated by demographic historian Hacker (1999), who assumed a relation between children's biblical names and the religiosity of parents to estimate an impact of religiosity on fertility in 19th century America. Knudsen (2019) made the same assumption, using biblical first names as a control variable.

We use two independent datasets with first names for individuals born in Europe long before the onset of modern growth and beyond. Our main database contains nearly half a million authors registered in the Virtual International Authority File (VIAF) (www.viaf.org). The VIAF database holds information on all recorded name variants of authors of texts in libraries across the globe, including their birth and death year and a unique identifier that can be linked to other data sources. The database gathers information from more than 40 national libraries, cultural agencies, and other major institutions in more than 30 countries around the world.[19] To obtain geocodes for the birthplace of authors, we cross-reference with the German National Library (DNB) and Wikidata (see Appendix A.4 for details). As our definition of a religious figure is restricted to Christianity and since our aim is to investigate the transition to modern growth, we restrict the sample to Europe.[20] We restrict further to authors born after 1300 when Christian names had fully spread to Germanic Europe (Sargent, 1990) and before 1940 when most potential authors had been born at the time of download (April 2021).[21] We remove

[18]Parents naturally pass on their believed truth about the world and religion may have benefits that parents would like to pass on to their children, such as improved mental health (Miller *et al.*, 2014; Park *et al.*, 1990), higher life satisfaction (Ellison *et al.*, 1989; Campante & Yanagizawa-Drott, 2015), abilities to cope with adversity (Clark & Lelkes, 2005), and refraining from deviant behavior (Lehrer, 2004).

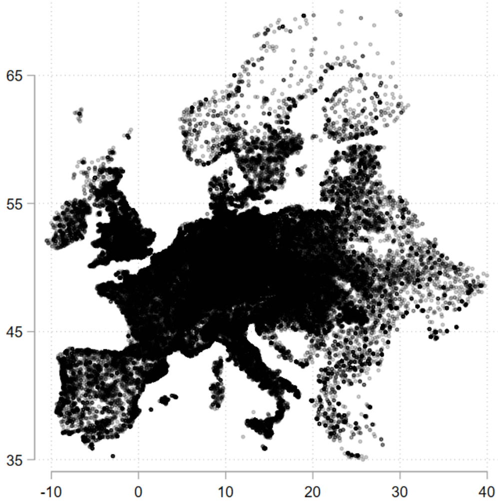
[19]The database is also used by Chaney (2020), who uses birthplaces of authors to improve on population size measures.

[20]Extending to Christian societies beyond Europe would be inferior, as Christian names outside Europe also indicate European descent.

[21]The youngest authors in our sample are therefore 81 years old, making it probable that all potential

Muslim majority countries and end up with 487,156 authors born in Europe between 1300 and 1940 for which we have first names and birthplace geocodes.[22] The locations of the authors' birthplaces are depicted in Figure 1. Our second database consists of 47,125 university students in the Holy Roman Empire and is described in Appendix I.

Figure 1: Locations of author birthplaces



For our baseline measure, we define a name as signifying higher religiosity of parents if it was shared by a religious figure that was considered sufficiently central.[23] Which figures were considered central varied between Protestants and Catholics. While saints are notably important to Catholics, Protestants tend to be more inspired by biblical figures (Dues, 1992; Hacker, 1999). There is a large overlap between the two, as many biblical figures became saints.

To identify which saints are significant, we exploit the Catholic tradition of dedicating churches to important saints, so-called patron saints. Whether or not a saint had a church dedicated to him or her is a signal of their importance as a religious figure. Accordingly, we define a saint as central if a large European medieval church was dedicated to him or her.[24] We exploit data collected by Buringh *et al.* (2020) on patron saint names of major

authors had entered the database at the time of download. This avoids problems of adverse selection by genre, prominence, etc.

[22]Based on contemporary numbers, the three Muslim majority countries in Europe are Albania, Bosnia and Herzegovina, and Kosovo. With 209, 498, and 89 authors born in these countries, respectively, this restriction of the data is insignificant for the results. We will further account for ethnicity by controlling for names that are more prevalent among different linguistics groups, of which Semitic speakers are of particular interest because of the sizable Jewish minority in Europe at the time.

[23]Alternatively, one could argue that parents inspired by minor religious figures were particularly religious. However, as there are multiple minor religious figures they would span close to the full set of names, rendering our measure useless. We exclude the rarest names for robustness in Tables B2 and G2. Another concern is that our measure is inaccurate for societies with a substantial Jewish population, which is one reason for constructing alternative names-based measures, described below.

[24]We refrain from conducting the simpler exercise of using all saints names, as their names would span close to all names. Also, choosing all churches throughout Europe, instead of the largest, would also not

churches built in the Middle Ages in present-day Italy, France, Switzerland, Germany, Belgium, the Netherlands, and Great Britain. Buringh *et al* focus on this region, as it was active in the commercial revolution in the 12th and 13th centuries and thus has the best and most accessible data on church building in the Middle Ages. These countries were also central in terms of religious influence with the Pope being seated in the Vatican City and the Protestant Reformation starting in Germany, for instance. Even if more countries were included in the sample, we do not expect to gain much, as important saints are highly similar across countries.[25] [26] To keep the task manageable and to focus on churches of greatest economic significance, Buringh *et al* further restrict their database to urban churches and to the top-24% of present-day churches built in the Middle Ages with a completed floor area of at least 1,000 square meter.[27] Buringh *et al* provide the identity of saints in the case of 1,778 different dedications for a total of 1,510 churches.[28]

Names may differ across countries and languages, even if they refer to the same saint. To eliminate differences caused by language and spelling, we reduce the first names of patron saints to their etymological branch, meaning their original form. We do so using the online etymology database www.behindthename.com, containing a total of 23,938 names rooted in 7,606 name branches. As an example, Jean, Johannes, Hans, and Giovanni are all part of the same name branch as the English John.

We are able to identify the name branch of 1,601 of the 1,778 saint names.[29] The first names represent 176 distinct name branches, of which Table 1 shows the top-10. Around one out of four churches were dedicated to a central Catholic figure, the Virgin Mary. After Mary comes Peter, arguably the most prominent of the apostles of Jesus and according to tradition the first leader of the Catholic church. Next comes John, referring

generate much variation in our religiosity measure, as the vast majority of names would be represented. Last, focusing on medieval rather than current churches, for instance, has the great advantage of being pre-determined relative to the vast majority of our sample.

[25]For instance, Mary is the most common patron saint among the churches in all seven countries. Also, the frequencies of the saints names in each of the countries are highly correlated, the average of the pairwise correlations being 0.86.

[26]The alternative idea of using information on the church location would not improve our measure, as we use the churches to define which saints were considered central in *general*. Using the European average rather than local saints ensures that we capture major patron saints and that our definition of important figures is not affected by local economic factors that we will later have as our dependent variable. Also, some saints were likely chosen in response to local competition from other religions (Barro & McCleary, 2016), which further reduces the usefulness of local saints' names. Last, names inspired by local saints may proxy for localism rather than religiosity.

[27]Even if rural churches were available to us, adding them would not improve our measure as the clear majority of large churches were built in urban areas and thus the sample most likely spans close to the full set of major patron saints.

[28]Some churches are dedicated to more than one saint.

[29]The remainder are rare names absent from the behindthename database and thus excluded from the analysis. This is unlikely to impact results, as few authors in our sample would have such rare names.

to John the Apostle or John the Baptist.[30]

Table 1: Top-10 name branches of saints to which large medieval churches were dedicated

Saint	Share
Mary	0.27
Peter	0.09
John	0.06
Martin	0.04
Paul	0.04
Nicholas	0.04
Stephen	0.03
Jacob	0.02
Michael	0.02
Francis	0.02
Remaining 126 saints	0.35

We define major biblical figures as Jesus, his relatives, and the apostles, following Hacker (1999). Their name branches include Andrew, Barnabas, Cleopatra, Jacob, Jesus, John, Joseph, Judah, Mary, Paul, Peter, Philip, Simon, Talmai, Thomas (described in Section "Religious name" in Appendix A.4). To the definitions of central saints or biblical figures, we add the name branch Christos, which is an obvious reference to Jesus, even though it is not formally his name, but rather his title. We also add name branches that refer directly to the Christian god, such as Amadeus and Godfrey.

According to our preferred definition, a religious name is one that is shared by a major saint or biblical figure. The top-10 most frequent religious and non-religious names among the authors are shown in Table 2. The list of religious names contains name branches of well-known religious figures, such as John, Peter, Paul, and Francis.[31]

Table 2: Top-10 religious and non-religious name branches among authors

Religious		Non-religious	
Name	Share	Name	Share
John	0.11	Henry	0.03
Charles	0.05	Frederick	0.03
Joseph	0.03	William	0.02
Francis	0.03	Ernest	0.01
George	0.03	Rudolf	0.01
Ludwig	0.02	Herman	0.01
Peter	0.02	Otto	0.01
Paul	0.02	Walter	0.01
Anthony	0.02	Conrad	0.01
Christos	0.01	Gerard	0.01
The rest	0.68	The rest	0.85

Top-10 most common name branches among authors who share name with patron saints or biblical figures in the first columns and top-10 most common name branches of authors who do not share name with these figures in the last columns.

Central kings in the history of Europe are among the most frequent name branches that were not shared by religious figures. The name branch Henry tops the list, a common name for rulers of Germanic kingdoms in the Middle Ages, especially following the reign of Henry the Fowler (876 – 936). Nearly as popular, the name branch Frederick was borne by rulers of the Holy Roman Empire, Germany, Austria, Scandinavia, and Prussia.

[30] For descriptions of the remaining names in Table 1, see Section "Religious name" in Appendix A.4.

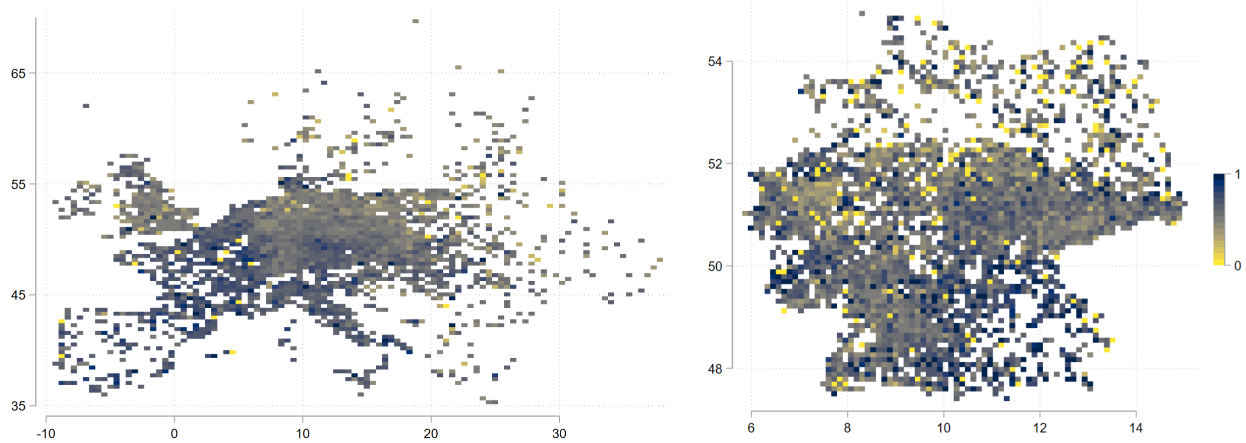
[31] The latter refers to Francis of Assisi (1181-1226), who is among the most venerated figures in Christianity.

Next comes William, which exploded in popularity across Europe following William the Conqueror (1028-1087) and the Norman conquest of England. The main results hold if we exclude each name on either list individually or all together, signifying that results are not sensitive to any of these names (Tables B1 and G1).

Nearly 60% of the authors share name with a major patron saint or biblical figure. This average covers a general decline from around 75% between 1300 and 1750, to around 50% in 1900 and beyond (Figure B1). The share fluctuates intensively until around 1450, reflecting that few authors in the sample were born at this time (Figure B2). Furthermore, the share covers large spatial heterogeneity (Figure 2).

Figure 2: Share of names shared by significant religious figures

(a) Full sample (b) Example: Germany



The share of authors in a grid cell who share name with a major saint or biblical figure for the full sample in panel a and zoomed in on the country with most authors, Germany, in panel b to help visualize the detail of the data. A dark blue grid-cell indicates that all authors born in this cell had a religious name, while no authors in clear yellow grids had a religious name. Any tones of blue or yellow between these extremes indicate that a mix of authors with or without religious names were born here. The sample in the figure is restricted to grids with at least 10 authors.

Table 2 also reveals that the top religious names are more common than the non-religious. To ensure that our religiosity measures do not simply capture normal names, we account for whether or not a name is common in all specifications. Further, we construct a "normality adjusted" measure of religiosity, which we term the Religious Name Index (RNI). The measure is inspired by the seminal work of Fryer & Levitt (2004) and captures the frequency of a name among religious figures, relative to the full population:

$$RNI_j = \frac{Pr(name_j|RelFig)}{Pr(name_j|RelFig) + Pr(name_j|Population)} \tag{1}$$

for each name branch j and for each definition of a significant religious figure $RelFig$. $Pr(name_j|Population)$ is the percentage of the population with a name belonging to name branch j. When defining significant religious figures based on saints, $Pr(name_j|RelFig)$ is the percentage of churches that bear name branch j. The highest possible value for the index is one, which means that only churches bear the name and no persons. The lowest possible value is zero obtained if no churches had the name (but at least one person did). The index equals 0.5 if the name is equally common among people and churches. When defining central religious figures based on the Bible, we calculate $Pr(name_j|RelFig)$ as the frequency with which the name is used in the Bible.[32] These RNI measures face limitations. For instance, it is not obvious that a person was more important in Christianity if he or she is mentioned more frequently in the Bible. Furthermore, the divide between religious and non-religious names may be closer to a binary choice. The main benefit of the RNI measures is that they account for the general normality of a name and we regard them as important robustness checks (Tables B3 and G3).

One may question our definition of which figures are religiously significant. For instance, our measures may be inaccurate for societies with a large Jewish population. We therefore supplement with a measure that does not rely on religious figures, but instead defines a name as more religious if it is more frequent among persons with religious occupations, compared to the general population (cf. Appendix A.4 and Table G3).[33]

2.1 Validity tests

We proceed to validate the names-based religiosity measures vis-à-vis alternative behaviors associated with religiosity. First, we exploit information on the authors' occupations, available from the German National Library (DNB) for 362,666 of the authors. The authors in our sample hold between 1 and 24 occupations. The average (or median) author holds two occupations. The most frequent occupations are writers, doctors, painters, and lawyers, held by 10%, 8%, 7%, and 7% of the authors, respectively. As a simple illustration of differences between authors with religious names and those without, the semantic analysis in Figure 3 shows the excess frequencies of occupations among authors who share name with a religious figure in panel a and for those who do not in panel b. For instance, 3% of the occupations of authors with religious names is "pastor", compared to 1.7% for the authors without a religious name. Pastors are therefore 1.3 percentage points more frequent among authors with religious names, symbolized by the size of the word "pastor" in panel a. For this exercise, we keep the original disaggregation level of occupations, which means that some occupations overlap, such as "teacher" and "university teacher". Some terms that enter occupations are not formally occupations, such as "catholic", but

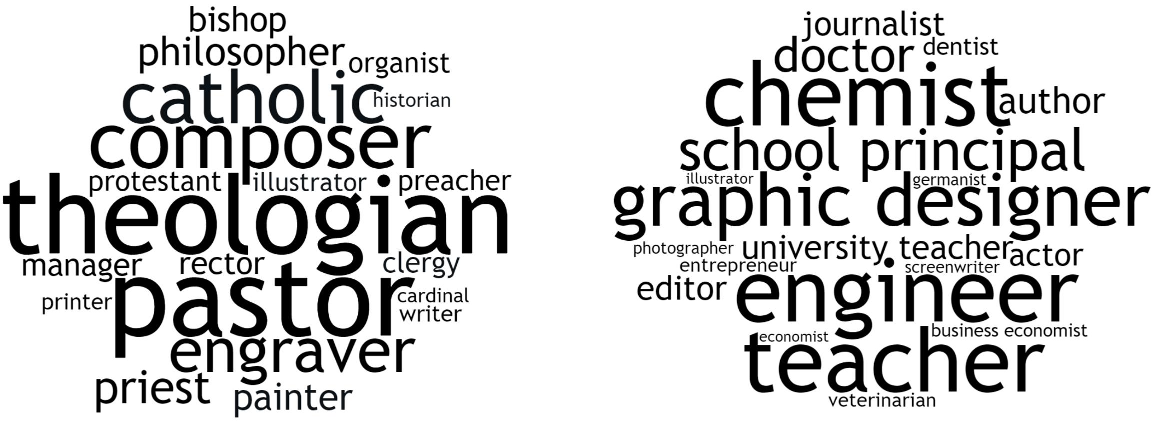
[32] We do this for the Old and New Testament. For more information, see Appendix A.4. Results are robust to using any of these alternative measures independently, cf Tables B3 and G3.

[33] In addition, the results hold to using Old Testament names, which contains a great number of Jewish names (Tables B3 and G3).

since they do elicit useful information about the persons, we keep them as occupations.[34]

Figure 3: Occupations split by author names

(a) Religious names (b) Non-religious names



The size of the words reflects the excess frequency of the top-20 words describing occupations of authors with a religious name, relative to those without in panel a and for authors without religious names, relative to those with religious names in panel b.
Result: Authors who share name with a religious figure are more likely to have a religiously associated occupation.

The figure reveals clear differences in occupations for the two author types. Ten of the twenty relatively most frequent occupations among authors with a religious name are associated with religion (panel a). None of the relatively more frequent occupations for the remaining authors are directly associated with religion (panel b). We examine the latter more formally in Section 5.

A concern is that the pattern in Figure 3 is explained by temporal or other characteristics with a simultaneous bearing on religiosity and science. For instance, this period saw a rise in secularization and a simultaneous rise in modern science (Stark & Finke, 2000; Bruce, 2002). To test econometrically whether the association between religious names and occupations is real, we estimate regressions of the type:

$$occupation_{igt} = \beta relname_{igt} + \gamma_g + \gamma_t + \kappa t_g + \gamma_{gt} + \omega X_{igt} + \varepsilon_{igt} \tag{2}$$

where $relname_{igt}$ measures whether or not author i born in grid cell g shared name with a significant religious figure. $occupation_{igt}$ is a dummy equal to one if the author held at least one religious occupation, defined using the European Roots Genealogy list of old German occupations and Wikipedia's list of Christian Religious Occupations.[35] The main occupations of authors with at least one religious occupation are pastor (18%), theologian (17%), Catholic (4%), and Protestant (4%), while the main occupations of authors who have no religious occupations are writer (5%), painter (4%), lawyer (4%), and doctor (4%)

[34]Only 1% of the authors have a mention of Catholic or Protestant, which indicates that such a mention is not a proxy for whether or not the person was Catholic or Protestant. Instead, we find it more plausible that this mention indicates higher religiosity.

[35]We put more focus on German occupations, as the authors' occupations are written in German.

(Table A2). γ_g are grid cell fixed effects of size 1x1 degree amounting to approximately 100x100 km^2 at the Equator in the baseline specification. We document robustness to including grid cell fixed effects as detailed as 1x1 km^2. γ_t are birth year fixed effects, and t_g are grid cell specific trends. In the most flexible model, we include instead grid cell by birth year fixed effects (γ_{gt}) which means that we compare authors born in the same grid cell in the same year. X_{igt} captures individual-level factors that potentially correlate simultaneously with a person's name and occupation choice. These include the number of occupations held by a person, their gender, different indicators of the commonness of a name, birth year, a Protestant area indicator, ethnicity, birthplace latitude and longitude, distance to Wittenberg, migration distance, an urban area indicator, and a list of variables meant to capture the part of a name that is influenced by parental socioeconomic characteristics.

As opposed to many other samples, our dataset includes a rather narrow group of individuals –authors–, which limits the potential influence of many of the usual potential confounders, such as social status, income, and education. Nevertheless, we compute different proxies meant to account for bias caused by socioeconomic factors. The rationale is that to remove potential bias by parental socioeconomic characteristics, we can simply identify the part of the name that is influenced by these characteristics.[36] Therefore, we compute the commonness of a name among urban residents, nobles, advanced university students, and high-education areas. In addition, we follow Knudsen (2019) and include last name by country fixed effects as another catch-all for parental socioeconomic status.[37]

The error term, ε_{igt}, captures all other factors that determine occupation choice. To address the concern that some of these factors also correlate with what is left of the religious name variation, we proceed with exploiting shocks to religiosity in Section 4. A significant and positive estimate of β indicates that an individual with a religious name was more likely to end up in a religious occupation, compared to others. The direction of causality is implicit in the model, as choice of occupation later in life is unlikely to influence one's first name.[38] We cluster standard errors at the 1x1 degree grid cell level.[39]

[36]Omitted confounders that is correlated with the outcome variable enters ε_{igt}. If these confounders do not correlate with the measure of a religious name, they do not bias β. They only pose a cause for concern if they are simultaneously correlated with the measure of a religious name.

[37]For controls in addition to those in Table 3, see Table 4 and Tables B1-D2.

[38]There are examples of name-changes in response to choice of occupation, but these cannot explain results, cf. Section 4.

[39]The results are robust to using various different levels of clustering (Tables B4 and G4), which is unsurprising given the lack of spatial clustering of the data (Voth, 2021). See Figures A1, A2, and 2 for the spatial distribution of the data. The main result stays remarkably stable when adding spatial fixed effects consecutively from no fixed effects down to 1x1 km grid cell fixed effects, which further indicates that the results are driven by individual-level effects, rather than spatial effects (Tables 4 and 6).

The main results for this first validity check are presented in panel A of Table 3: individuals who share name with a significant religious figure are more likely to hold religious occupations later in life. The raw correlation is shown in column (1). We add a control for the number of occupations held by the author to remove mechanical effects (column 2). This has no impact on the relation between religious names and occupations. Column (3) adds a control for gender. Women are less likely to hold religious occupations, consistent with the focus on traditional gender roles in most major religions (e.g., Sharma & Young (1987)). We next add an indicator of the commonness of a name in column (4) as a proxy for the collectivism-individualism dimension of culture, which potentially influences various economic factors (Gorodnichenko & Roland, 2011).[40] Individuals with common names are more likely to hold religious occupations, although this is exclusively driven by a trend in both variables (column 6).

Table 3: Religious names and occupations related to religion

Dep var: Religiously associated occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full sample								
Religious name	0.055***	0.055***	0.055***	0.050***	0.047***	0.020***	0.017***	0.014***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.002)	(0.001)	(0.001)
Number professions		0.018***	0.016***	0.016***	0.016***	0.012***	0.014***	0.012***
		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Female			-0.083***	-0.071***	-0.056***	-0.035***	-0.038***	-0.037***
			(0.005)	(0.004)	(0.004)	(0.003)	(0.002)	(0.002)
Common name				0.032***	0.033***	-0.0074***	-0.0085***	-0.0057***
				(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
Birth year						-0.00093***		
						(0.000)		
R-squared	0.0076	0.011	0.018	0.020	0.049	0.12	0.14	0.29
Observations	362666	362666	362666	362666	362609	362609	362592	325462
Mean dep var	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.10
Number grids	917	917	917	917	860	860	860	652
Panel B: Excluding scientific occupations								
Religious name	0.060***	0.060***	0.060***	0.054***	0.051***	0.022***	0.019***	0.015***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.002)	(0.002)	(0.002)
R-squared	0.0077	0.0097	0.019	0.021	0.058	0.13	0.15	0.31
Observations	297143	297143	297143	297143	297077	297077	297060	261582
Mean dep var	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.12
Number grids	907	907	907	907	841	841	841	639
Year FE	N	N	N	N	N	N	Y	Y
1x1 grid FE	N	N	N	N	Y	Y	Y	Y
Grids x year FE	N	N	N	N	N	N	N	Y

OLS across authors. The dependent variable is a dummy equal to one if one of the occupations of the author relates to religion. Religious name is a dummy equal to one if the author shares name with a major saint or biblical figure. Controls include the dummy for top-10 most common names, gender, number of occupations, birth year fixed effects, 1x1 degree grid cell fixed effects, and 1x1 degree grid cell by birth year fixed effects. Panel A includes the full sample of authors, while panel B excludes authors with at least one scientific occupation. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors who share name with a significant religious figure are more likely to hold occupations related to religion. This is not due to an impact on the likelihood of holding scientific occupations.

To remove a large set of factors that may influence both name giving and occupation choice, we add grid cell fixed effects in column (5) with no change to the conclusion. Column (6) adds a control for birth year, halving the estimate on the religious name dummy, consistent with the simultaneous sharp decline in religious names and occupa-

[40] The common name indicator measures whether or not the author's name is among the top-10 most common first names in a country within 10 years prior to their birth, a standard measure used by Knudsen (2019). Knudsen calculates the measure for men and women separately. As our samples include few women, our preferred measure does not take gender into account. The results are robust to including the gender-based measure or other variants (Tables B7 and G7).

tions during the sample period, coinciding with secularization of Europe.[41] Column (7) removes variation caused by time-differences more flexibly by adding birth year fixed effects, while column (8) increases flexibility further by adding grid cell by birth year fixed effects instead.[42] The estimate on the religious name dummy remains highly significant, even in this specification comparing persons born in the same year and grid cell, of the same gender, name normality, and holding the same number of occupations.

Taking the estimate in column (8) at face value, sharing name with a religious figure increases the likelihood of holding a religious occupation later in life by 14% of the mean. This amounts to nearly two thirds of the well-known gender gap in religious occupations.[43]

The estimated association between religious names and occupations is not specific to any groups of observations or countries and is nearly identical in Protestant and Catholic areas (Figure B3, B4, and Table B5). Also, results are unchanged when including a Protestant area dummy or a measure of Protestant names, cf Table B6.[44] Gauging heterogeneity across time (panel c of Figure B4), we find that the relation between religious names and occupations is positive throughout the period, but only significantly different from zero before 1600 and after 1800. The estimates between 1600 and 1750 are indistinguishable from zero, which may reflect the rapid secularization of Europe at this time. This is a period of declining popularity of religious occupations and religious naming, the latter potentially somewhat more persistent than the former. Notably, our measure seems to work well even for people born in the most recent period of our sample (1921-40).

One may still worry that omitted confounders explain results. For instance, rural families may face different naming conventions than urban families or scientists may choose certain names for their offspring and encourage them to pursue a career in science, thus potentially reducing their likelihood of ending up in religious occupations. To address, we remove individuals with scientific occupations in panel B of Table 3. The β estimate on religious names is unaltered. If anything, it rises slightly, indicating that any explanations involving shifts in scientific occupations cannot explain these validity checks.

To further eliminate bias by social status, we add proxies for the part of names that

[41]The estimate on common names turns negative, reflecting rising names diversity.

[42]The latter specification drops 35,000 observations, since only one author is born in these grid cell-years. The result is unaltered if instead including grid cell specific trends, which spares the singletons.

[43]This calculation is based on the standardized beta coefficients, which are 0.023 for the religious name dummy and -0.036 for the gender dummy. For further comparison, the standardized betas for the common name dummy is -0.009 and number professions is 0.044.

[44]To measure Protestant names, we measure the frequency of a name among Protestant regions, compared to the rest. The latter checks relate to a large literature documenting various socioeconomic impacts or lack thereof of the Protestant Reformation, e.g., Becker & Woessmann (2009); Cantoni (2015); Cantoni *et al.* (2018); De la Croix *et al.* (2020b). See a review by Becker *et al.* (2016).

captures parental socioeconomic status in Panel A of Table 4. If rural families face different naming conventions than urban families, this would be reflected in the names they choose for their children. Our datasets allow us to compute four "Social status of a name" indices, SNI, to capture whether a name was more frequent among the urban population, nobles, advanced university students, and high-education areas (described in detail in Section "SNI indices" in Section A.4). These indices mimick our RNI indices and the indices by Fryer & Levitt (2004) capturing the "blackness" of a name. To remove the variation in names caused by urban-rural differences, we calculate the frequency of a name in urban areas compared to the full population.[45]

Next, we compute the frequency of a name among nobles, using information on whether or not students belonged to the nobility in the RAG database. Similar to the RNI measures of the excess-frequency of names among churches relative to the population, we compute the frequency of names among nobles in the RAG database and divide this by the frequency of the same name among authors.[46] The RAG database also contains information on whether or not students proceeded with advanced studies. For each name branch, we compute the frequency of the name among students with advanced degrees and divide this by the frequency among authors. Last, we elicit information on educational attainment from the European Values Study (EVS), a survey of the European respondents about various values, demographics, and socioeconomics (for more details, see Appendix F). We do not know the name of the EVS respondents, but we know the NUTS2 region in which they were interviewed. From this, we calculate the average educational attainment in each NUTS2 region. Similar to the urban-rural measure, we then compute the frequency of author names in NUTS2 regions with above median education, relative to the rest.[47]

Authors with names that are more common in urban areas, more common among the nobility and in areas with above-median educational attainment are less likely to end up in religious occupations (columns 2-6). If these names reflect lower socioeconomic status of the parents, this is consistent with the existing literature.[48] Accounting for this reduces

[45]We also add a control for whether or not the author was born within 5 km of a city and city by year fixed effects in Table B9. The results are unchanged.

[46]We restrict to authors with name branches that exist in both the RAG and author database. The median person in the RAG database was born in 1488, while the median author was born in 1884. The assumption is that the part of author names that stems from differences between the nobility and the rest is correlated across the two datasets. While the nobility often adhered to a strict hereditary rule, this correlation may be quite small. This would tend to reduce significance of the nobility name index.

[47]While the difference in timing between the RAG database and the authors is quite large, the median birth year among the EVS respondents is 1937. The assumption is that a sufficiently large share of NUTS2 regions with above median educational attainment around 1884 are also above the median in 1937.

[48]See, e.g., Norris & Inglehart (2011) documenting higher religiosity in more insecure areas and

the relation between religious names and occupations by up to around a quarter, which is mainly due to the higher likelihood of religious names in rural areas. The religious name dummy remains highly significant statistically (at the 1% level) and economically (it equals 10% of the mean). As last crude proxy for socioeconomic status, we add last name fixed effects in column (7). The estimate on religious names is unchanged.[49]

Table 4: Religious names and religious occupations - socioeconomic confounders

Dependent variable: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Socioeconomic confounders							
Religious name	0.014***	0.010***	0.014***	0.013***	0.013***	0.0098***	0.0085***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Urban name index		-0.11***				-0.11***	-0.20***
		(0.017)				(0.018)	(0.021)
Noble name index			-0.012***			-0.015***	-0.019***
			(0.003)			(0.003)	(0.004)
Advanced degree name index				-0.0069		-0.016***	-0.017*
				(0.006)		(0.006)	(0.010)
High education name index					-0.040***	-0.0070	0.021
					(0.011)	(0.012)	(0.015)
R-squared	0.30	0.30	0.30	0.30	0.30	0.30	0.48
Observations	273721	273721	273721	273721	273721	273721	169544
Grid x year FE	Y	Y	Y	Y	Y	Y	Y
Baseline controls	Y	Y	Y	Y	Y	Y	Y
Last name x Country FE	N	N	N	N	N	N	Y
Panel B: Within-neighborhood comparisons							
Fixed effect:	None	Country	2x2 grid	1x1 grid	0.5x0.5 grid	0.1x0.1 grid	0.01x0.01 grid
Religious name	0.013***	0.016***	0.015***	0.016***	0.015***	0.013***	0.012***
	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.13	0.13	0.14	0.14	0.15	0.19	0.23
Observations	339666	339666	339666	339666	339666	339666	339666
Number spatial FE	0	43	256	792	2225	10981	18522
Number year FE	603	603	603	603	603	603	603
Panel C: Within neighborhood and socioeconomic status comparisons							
Religious name	0.0097***	0.0097***	0.0099***	0.010***	0.010***	0.0093***	0.0091***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.35	0.35	0.35	0.36	0.36	0.40	0.43
Observations	182576	182576	182576	182576	182576	182576	182576

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the author shares name with a significant saint or biblical figure. Urban Name Index measures the frequency of the name in urban areas compared to the rest of the population. Noble name index measures the frequency of the name among nobles compared to the rest of the population. Advanced degree name index measures the frequency of the name among students with advanced degrees compared to the rest of the population. High education name index measures the frequency of the name among regions with above median level education, relative to other regions. All regressions include controls for gender, top-10 most common names, number of occupations, and birth year fixed effects. In addition, panel A includes 1x1 degree grid cell by year fixed effects. In addition, we add 32,150 last name by country fixed effects in column (7) of Panel A. In addition to the baseline controls, panel B includes spatial fixed effects at the country level in column (2), at the 2x2 degree grid cell level in column (3), 1x1 degree grid cell level in column (4), 0.5x0.5 degree grid cell level in column (5), 0.1x0.1 degree grid cell level in column (6), and at the 0.01x0.01 degree grid cell level in column (7). Panel C replicates panel B, but adds the Urban, Noble, Advanced Uni, and High Education name indices and last name by country fixed effects throughout. The sample in panel A is restricted to the sample in column (6). Panel B and C are restricted to the sample in column (7) in these panels, respectively. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and occupations holds after eliminating parts of the name which is influenced by parental socioeconomic status and when restricting comparison to increasingly narrow neighborhoods down to 1x1 km^2.

To account for a large set of unobservable characteristics, we add increasingly fine spatial fixed effects in panel B from no spatial fixed effects in column (1) down to 1x1 km^2 grid cell fixed effects in column (7). In the most flexible model, we thus compare authors born within the same square kilometer, reducing urban-rural and general socioeconomic differences substantially, as individuals of similar socioeconomic status tend to cluster within neighborhoods. The parameter estimate on religious names is largely unaltered as we move from no spatial fixed effects to 18,522 1x1 km^2 grid cell fixed effects. At the same time, the explanatory power of the model increases by three quarters from a R^2 of

occupations. See also surveys by Iannaccone (1998); Iyer (2016); Bentzen (2021b)

[49]Estimating the specification in column (6) using the sample from column (7), produces an estimate on the religious name dummy at 0.0088 and R^2 remains at 0.30.

0.13 to 0.23, meaning that a substantial share of the variation in religious occupations stems from variation across neighborhoods. The Oster (2019) methodology reveals that selection on unobservable characteristics should be nearly four times larger than selection based on neighborhood differences to explain away our results.[50] In panel C, we add all four social status name indices from panel A and the last name by country fixed effects to the specifications of panel B. Even in this highly restricted model, having a religious name increases the likelihood of taking up a religious career by 7% of the mean.

Despite these various controls and spatial fixed effects, one may still be concerned that we have omitted important variation in unobserved parental characteristics that influence naming in a way correlated with religious naming. To obtain exogenous variation in religious naming, we exploit a phenomenon termed religious coping, which is a tendency for individuals to use their religion to cope emotionally with adversity (Pargament, 2001). Previous research has shown that survey-based measures of religiosity rise after earthquakes and this increased religiosity is transmitted across generations (Bentzen, 2019a). The impact persists for areas that are not physically hit, indicating that individuals use their religion to cope *emotionally* with the experience. If religious names capture the religiosity of parents, we expect that more individuals are given religious names when a major earthquake hit before they were named.

To examine, we combine a database of major earthquakes with the authors' place and year of birth. We find that authors born after one or more earthquakes are more likely to have a religious name (Figure 4, see Appendix C for details). Reassuringly, earthquakes that hit after birth did not influence naming and neither did earthquakes that hit more than 30 years before birth.[51] To prevent physical damage of earthquakes to contaminate results, we exclude authors born within 20 km of an earthquake. To increase homogeneity of the control group, we exclude authors born more than 500 km from an earthquake. The results are not sensitive to these restrictions of the data, cf. Table C1. We limit the analysis to authors born after 1700 to prevent improvement in earthquake detection technology to contaminate results (cf. panel b of Figure B4). Relaxing this restriction

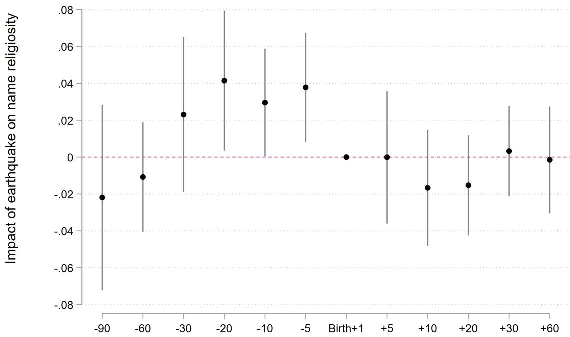
[50]The rationale behind Oster's approach is that stability of coefficients when adding controls is informative of the potential bias by unobservables if the controls add sufficiently to the explanatory power of the model. Following Oster (2019), we assume a maximal R^2 that is 30% larger than the R^2 from the controlled regression. Comparing the baseline in column (4) of panel B of Table 4 to the specification with the full set of detailed neighborhood fixed effects in column (7), we calculate a δ of 3.8.

[51]We do not have priors about the degree of persistence. For instance, the practice to name after grandparents or great grandparents would increase persistence. Transmission of religiosity across generations would work similarly. Previous research documents that religiosity remains higher up to 9-12 years after a large earthquake has hit in the modern era (Bentzen, 2019a). Thus finding persistence up to 20-30 years seems reasonable, given the historical setting.

increases standard errors, but otherwise does not change results. We do not include as fine-grained a set of grid cell fixed effects as in the analysis of occupations, since we define birthplaces as hit if an earthquake hit within a radius of 100 km and thus the earthquake indicator does not vary much at the 1x1 degree grid cell level. We include 5x5 degree grid cell fixed effects, year of birth fixed effects, and 2x2 degree grid cell specific trends.

These findings support the hypothesis that earthquakes reinforced the role of religion in the lives of parents beyond other considerations, which strengthens our confidence that names reflect the importance of religion in parents' lives. We will exploit this exogenous variation further to examine causal effects of religiosity on occupations in Section 4.

Figure 4: Religious names and earthquakes



Estimates of Equation (10) across 265,646 authors. Each dot reflects the impact of earthquakes in the particular time-period on the likelihood that the person was given a religious name. The vertical lines reflect the 95% confidence intervals, clustered at the 2x2 grid cell level. Control variables include the top-10 most common names dummy, gender, 5x5 degree birthplace fixed effects, birth year fixed effects, and 2x2 degree birthplace specific time-trends. The sample includes authors born within 20-500 km of an earthquake.
Result: Parents are more likely to give their child a religious name when naming after an earthquake.

We proceed to validate our names-based measure vis-à-vis two groups of measures of regional religiosity. First, the European Values Study asks eleven questions regarding respondents' religiosity, such as "How important is religion in your life?" and "How often do you attend religious services?", for instance (described in detail in Appendix F). We restrict analysis to respondents born before 1940 and to authors born after 1900. We match the survey respondents to the authors using information on the NUTS2 regions in which the respondents were interviewed. Whether we regress author names on the average NUTS2-level survey-based religiosity levels or regress answers across survey respondents on the NUTS2-level share of religiously named authors, we find a positive association between religious names and religious survey responses across all specifications (panels A and B of Table F1). Eighteen of the 22 associations are significantly different from zero. The surveys further allow us to account for respondents' income and education or their father's education (panels C and D of Table F1). Again, higher socioeconomic status is associated with lower religiosity, but the association between religious names and religious survey responses remain mostly unchanged.

Second, we probe validity of our names-based measure relative to the best other measure of historical religiosity that we know of; a measure of (Catholic) religiosity in French departments in 1791, used by Squicciarini (2020). During the French Revolution, the government asked all clergymen to swear an oath of allegiance to the Civil Constitution (Tackett, 1986). We match the share of clergy that remained loyal to the Church to the share of authors born within 30 years of 1791 in the particular French department (see Appendix E for details). We find that the share of religiously named authors is higher in departments where the clergy were more likely to reveal loyalty towards the church, consistent with the idea that both measures capture local religiosity (Figure E1).

In our last set of validity checks, we use the alternative dataset of 47,125 individuals who studied at a university in the Holy Roman Empire between 1300 and 1550. Since the clear majority of the students are observed before the Protestant Reformation, we define a name as religious if it was shared by a patron saint. We document that students who share name with a patron saint were more likely (15% of the mean) to study theology rather than medicine or law (cf Appendix I). This increases our confidence that we can proxy religiosity by peoples' names.

3 Religiosity, science, and studies

Equipped with a measure of religiosity that has passed a range of validity checks, we proceed to test the impact of religiosity on science. We commence with a formal econometric scrutiny of the wordcloud in panel b of Figure 3. We define occupations as pertaining to the fields of science or engineering if they are included in the "List of scientific occupations" in Wikipedia in the categories natural science, life science, earth science, formal science, statistics, physical science, or engineering. The main occupations among authors with at least one science or engineering occupation include doctor (21%), chemist (5%), university teacher (4%), engineer (3%), mathematician (3%), and physicist (2%) (Table A3).[52] The main occupations among authors who are not scientists or engineers include writer (6%), painter (4%), lawyer (4%), pastor (3%), and politician (3%).[53]

We estimate Equation (2) with a dummy variable equal to one if the author held at least one science or engineering occupation as dependent variable in Table 5.[54] The results confirm that authors with religious names are less likely to become scientists or

[52]Since each author may have several occupations, some occupations, such as university teacher, may enter the top-10 list without being themselves defined as belonging to science or engineering per se.

[53]Writer refers to a occupation ("schriftsteller" in German) as opposed to author ("autor" in German), which refers to the actual writer of a text.

[54]For completeness and to enable placebo checks, we use the full sample including the period 1600-1750, even though our measure may not capture religiosity well in this period (cf panel c of Figure B4). The results in this section strengthen if we restrict to authors born after 1750.

engineers. Taking the estimate in column (8) of panel B at face value, sharing name with a religious figure lowers the likelihood of having a scientific occupation by 6% of the mean. This amounts to 15% of the gender gap in science choice.[55] These findings are not simply the reverse of the results in Table 3: The estimates drop very little when excluding authors with at least one religious occupation (panel B). This rules out that the results are caused by crowding out of science occupations by religiously named authors who chose religious occupations.

Table 5: Religious names and scientific occupations

Dependent variable: Science occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Full sample								
Religious name	-0.034***	-0.034***	-0.034***	-0.033***	-0.026***	-0.015***	-0.013***	-0.0099***
	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.001)	(0.001)	(0.001)
Number professions		-0.026***	-0.028***	-0.028***	-0.028***	-0.027***	-0.027***	-0.032***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Female			-0.097***	-0.100***	-0.095***	-0.10***	-0.10***	-0.11***
			(0.005)	(0.005)	(0.006)	(0.005)	(0.005)	(0.005)
Common name				-0.0066***	-0.0019	0.014***	0.015***	0.011***
				(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Birth year						0.00036***		
						(0.000)		
R-squared	0.0020	0.0076	0.013	0.013	0.028	0.035	0.039	0.16
Observations	362666	362666	362666	362666	362609	362609	362592	325462
Mean dep var	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18
Grids x year FE	917	917	917	917	860	860	860	652
Panel B: Excluding individuals in religious occupations								
Religious name	-0.028***	-0.028***	-0.027***	-0.027***	-0.018***	-0.012***	-0.011***	-0.0084***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
R-squared	0.0012	0.0079	0.016	0.016	0.035	0.038	0.041	0.17
Observations	322489	322489	322489	322489	322429	322429	322404	288138
Mean dep var	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
Number grids	913	913	913	913	853	853	853	647
Year FE	N	N	N	N	N	N	Y	Y
1x1 grid FE	N	N	N	N	Y	Y	Y	Y
Grids x year FE	N	N	N	N	N	N	N	Y

OLS across authors. The dependent variable is a dummy equal to one if one of the occupations of the author is associated with science or engineering, zero otherwise. Religious name is a dummy equal to one if the author shares name with a major saint or biblical figure. Controls include the dummy for top-10 most common names, gender, number of occupations, birth year fixed effects, 1x1 degree grid cell fixed effects, and 1x1 degree grid cell by birth year fixed effects. Panel A includes the full sample of authors, while panel B excludes authors with at least one religious occupation. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with religious names are less likely to become scientists or engineers. This is not due to crowding out caused by the rise in religious occupations.

We add the proxies meant to remove variation in names caused by the socioeconomic status of parents in panel A of Table 6. As expected, having a name associated with higher socioeconomic status increases the likelihood of becoming a scientist or engineer. The association between religious names and science occupations remains highly significant, but falls by a third when accounting for different naming conventions in urban vs rural areas. The latter is the reason for exploiting shocks to religious names in Section 4. Comparing only individuals with the same last name in column (7) more than doubles the explanatory power of the model, but leaves the estimate on religious names unchanged.

Including the increasingly detailed spatial fixed effects in panel B, the estimate on religious naming remains significant throughout, even after removing cross-neighborhood

[55] The latter comparisons are based on the standardized beta coefficients calculated from column (8). They are -0.013 and -0.083 for religious name and the female dummy respectively. For further comparison, the standardized beta coefficients for the common name dummy and number of professions is 0.013 and -0.092, respectively.

differences down to the 1x1 km^2 level. This signals that at least part of the relation between religiosity and science occurs at the individual level (or other level that varies within one square kilometer), as opposed to the societal level.[56] Comparing the estimate in column (7) of Panel B to the baseline in column (4), we compute Oster (2019)'s δ to 6.14, meaning that omitted confounders would have to reduce the estimate by more than a factor six compared to the addition of 1x1 km^2 fixed effects to explain away results.

Table 6: Religious names and scientific occupations - socioeconomic confounders

Dependent variable: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Socioeconomic confounders							
Religious name	-0.011***	-0.0077***	-0.011***	-0.011***	-0.0096***	-0.0073***	-0.0088***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Urban name index		0.12***				0.10***	0.10***
		(0.015)				(0.017)	(0.031)
Noble name index			0.0022			0.0043	-0.0010
			(0.003)			(0.004)	(0.005)
Advanced degree name index				0.0043		0.0043	0.0051
				(0.007)		(0.007)	(0.011)
High education name index					0.070***	0.040***	0.039*
					(0.010)	(0.011)	(0.021)
R-squared	0.17	0.17	0.17	0.17	0.17	0.17	0.36
Observations	273721	273721	273721	273721	273721	273721	169544
Grid x year FE	Y	Y	Y	Y	Y	Y	Y
Baseline controls	Y	Y	Y	Y	Y	Y	Y
Last name x Country FE	N	N	N	N	N	N	Y
Panel B: Within-neighborhood comparisons							
Fixed effect:	None	Country	2x2 grid	1x1 grid	0.5x0.5 grid	0.1x0.1 grid	0.01x0.01 grid
Religious name	-0.020***	-0.015***	-0.013***	-0.013***	-0.013***	-0.011***	-0.010***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.025	0.033	0.037	0.040	0.046	0.073	0.095
Observations	339666	339666	339666	339666	339666	339666	339666
Number spatial FE	0	43	256	792	2225	10981	18522
Panel C: Within neighborhood and socioeconomic status comparisons							
Religious name	-0.010***	-0.010***	-0.0099***	-0.0095***	-0.0094***	-0.0083***	-0.0074***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.25	0.25	0.26	0.26	0.26	0.29	0.31
Observations	182576	182576	182576	182576	182576	182576	182576

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the author shares name with a significant saint or biblical figure. Urban Name Index measures the frequency of the name in urban areas compared to the rest of the population. Noble name index measures the frequency of the name among nobles compared to the rest of the population. Advanced degree name index measures the frequency of the name among students with advanced degrees compared to the rest of the population. High education name index measures the frequency of the name among regions with above median level education, relative to other regions. All regressions include controls for gender, top-10 most common names, number of occupations, and birth year fixed effects. In addition, panel A includes 1x1 degree grid cell by year fixed effects. In addition, we add 32,150 last name by country fixed effects in column (7) of Panel A. In addition to the baseline controls, panel B includes spatial fixed effects at the country level in column (2), at the 2x2 degree grid cell level in column (3), 1x1 degree grid cell level in column (4), 0.5x0.5 degree grid cell level in column (5), 0.1x0.1 degree grid cell level in column (6), and at the 0.01x0.01 degree grid cell level in column (7). Panel C replicates panel B, but adds the Urban, Noble, Advanced Uni, and High Education name indices and last name by country fixed effects throughout. The sample in panel A is restricted to the sample in column (6). Panel B and C are restricted to the sample in column (7) in these panels, respectively. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and scientific occupation holds after accounting for various proxies for parental socioeconomic status and when restricting comparison to increasingly narrow neighborhoods down to 1x1 km neighborhoods.

We conduct robustness checks identical to those of the results for religious occupations, most of which are relegated to the appendix for brevity (Appendix G, Tables G1-G11 and Figures G1-G2). Regarding heterogeneity of the results, we find that the religion gap in science i) is not particular to any group of observations (Figure G1), ii) nor to any one country (panel a of Figure G2), iii) exists within eight of the ten most represented countries in the sample (panel b of Figure G2), iv) is nearly identical within Protestant or Catholic areas (panel b of Figure G2), and v) emerges for individuals born after

[56]The estimate drops by 25% when removing cross-country variation and another 25 pct points when removing 1x1 km^2 cross-neighborhood variation. Whether this reflects societal-level effects or whether specifications in column (1)-(3) suffer from omitted variables bias is hard to say. To be conservative, we regard column (4) specifications as our baseline as done throughout.

1820, coinciding with the onset of the Technological Revolution later in the lives of these individuals (panel c of Figure G2). Insignificance between 1600 and 1750 is expected from the poor performance of our names-based measure in this period (cf. Figure B4).[57]

In the RAG database of medieval university students, we find that students who shared name with a major patron saint were 3-8% of the mean less likely to continue with advanced studies (Table I3). This amounts to between 55 and 135% of the gender gap in advanced studies.[58] This means that the students who have presumably had a more religious upbringing were more likely to end their university studies after their master, compared to their fellow students with a less religious upbringing who were more likely to continue with advanced studies. The latter were particularly more likely to continue with law studies (Table I4). This is consistent with a historical narrative suggesting that the pious were much less attracted by the study of law due to its' association with money and wealth (Brundage (2008), cf. Appendix I).

4 Further identification

We interpret the results so far as indicating that a person's religiosity can be proxied by their first name and that having had a more religious upbringing reduces the average likelihood of careers in science or advanced studies. A priori, several other factors could explain results, such as preferences for tradition, name changing, selection, or birth order effects. To accommodate, we leverage exogenous shocks to religiosity, but first we present evidence that results are rather unlikely to be caused by these alternative explanations.

One alternative explanation, based on the idea that religious names are more common names, is that choosing a religious name for their child reflects parents' preference for tradition or conformity, independently of their religiosity. A concern is that this focus on tradition explains why religious names are more frequent for persons who end up in religious occupations and less frequent for those in scientific occupations for reasons unrelated to religion. We find that this cannot explain results. First, this alternative argument relies on the assumption that names that are not shared by religious figures are names chosen by non-traditionalist parents. This is not necessarily the case. In fact, eight

[57]Insignificance during 1750-1820 begets explaining. If religion was exclusively bad for science throughout history, this insignificance would be puzzling, as 1750-1820 is a period for which our measure performs well and where the Enlightenment had been underway for more than a century. We speculate that potential positive effects of religion may have canceled out any negative effects in this period. After 1920, the estimate converges toward zero again, which could reflect a shift in occupation choices during the world wars or that the gap between the more and less religious is closing.

[58]The reported estimates are based on beta estimates calculated in specifications corresponding to columns (8) and (9) of Table I3, which include the baseline controls plus 1x1 grid cell by year fixed effects and 2x2 grid cell by year fixed effects, respectively. The binned added variables plot in Figure I4 reveals that results are not driven by individual groups of observations.

of the top-10 non-religious names owe their prominence to famous European monarchs, which are potentially similarly traditional as religious figures and hard to classify as non-conformist (Table 2). Thus, if the results cover tradition in general, one would have to explain why parents with a preference for tradition chose religious rather than royal names for their children, if not for religion.

Second, if general tradition explains the relation between religious names and religious occupations, we would expect individuals with royal names to be more likely to hold religious occupations as well. We test this more formally in Table 7, where we define name branch j of authors born in country c as royal if a king or queen who reigned country c before the time of birth of the author had a name within this name branch (more details in Appendix A.4). We find that having been given a royal name does not predict religious occupation choice (columns 1 and 2) and controlling for it does not change the impact of religious names (column 2). This is not because royal names are irrelevant for occupational choice; individuals with royal names are less likely to choose scientific occupations (columns 4 and 5).[59] This indicates that while tradition or nationalism are ideological dimensions that potentially influence occupation choice, the effect of religion on occupational choice cannot be reduced to its association with traditionalism.

Table 7: Religiosity and occupations - exploring general tradition as explanation

Dependent variable:	Religious occupation			Scientific occupation		
	(1)	(2)	(3)	(4)	(5)	(6)
Royal name	0.00099	-0.00029		-0.012***	-0.011***	
	(0.002)	(0.002)		(0.003)	(0.003)	
Religious name		0.014***	0.014***		-0.0098***	-0.010***
		(0.001)	(0.001)		(0.001)	(0.001)
R-squared	0.29	0.29	0.29	0.16	0.16	0.16
Observations	302596	302596	302596	302596	302596	302596

OLS across authors. The dependent variable is a dummy equal to one if one of the occupations of the author relates to religion in columns (1)-(3) and a dummy equal to one if one of the occupations of the author relates to science or engineering in columns (4)-(6). Religious name is a dummy equal to one if the author shares name with a major saint or biblical figure. Controls include the dummy for top-10 most common names, gender, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: A dominant group of traditional names –royal names– do not influence religious occupation choice and do not explain the impact of religious names on occupation choice.

Third, another measure that could reflect traditionalism is the choice of common names. If traditionalism explains results, we would expect this alternative measure of traditionalism to influence occupation choice in a way similar to religious names. That is, individuals with common names should be more likely to hold religious occupations and less likely to be scientists. We find the opposite. Individuals with common names are *less* likely to hold religious occupations and *more* likely to be scientists (column 8 of Tables 3 and 5). Thus, if not for religion, one would have to explain why having a common name associated with religion leads to a behavior that is opposite of what having a common name in more general predicts.

[59]For size comparison of the estimates on religious and royal names in column (5), the standardized beta estimates are -0.013 and -0.009, indicating that the importance of royal names is 70% the importance of religious names.

Fourth, in addition to controlling for name commonness throughout, we reduce variation caused by name commonness by using the RNI measure of religious naming that implicitly accounts for the commonness of names (Tables B3 and G3). The result is also insensitive to excluding top-religious or top non-religious names and to excluding the most common or least common names (Tables B1, B2, G1, and G2). These results enhance our confidence that religious names capture religiosity per se, rather than a more general preference for tradition or conformity.

Another potential explanation is that the relation between religious names and occupations covers a tendency for priests to change their name to a religious name to increase their credibility or as part of the tradition in certain religious orders. In this case, the relation between religious names and occupations still serves as a validity check of names reflecting religiosity, even though the religiosity would be that of the child instead of the parent. A concern, though, is that the relation between names and scientific occupations reflects this name-changing. Since we rule out name-changing based on religious occupations when excluding religious occupations in panel B of Table 5, the name-changing would have to relate to the science occupation. Thus, the concern is that scientists changed their name away from religious names. We do not find support for this. First, even when restricting the sample to those without religious occupations, we find that we would have to replace more than 35% of the sample randomly with religious names in order to remove significance (Table G11). We know of no literature suggesting name-changing among non-religious occupations of this scale. Second, we exploit that changing ones primary first name is arguably easier for those with more than one first name. If name changing explains part of our results, we would expect results to strengthen for those with more than one first name. We do not find support for this (Table D1). Based on this evidence, we deem it unlikely that the observed association between names and occupations is caused by name changes.

A third potential explanation is that our samples contain a selected group of individuals, which may bias our results. For instance, if the likelihood of becoming an author is higher for individuals with religious names and occupations, this could create a spurious correlation between religious naming and religious occupations. This does not seem to explain results. First, the correlations between names and regional-level measures of religiosity and the rise in religious naming after earthquakes make it hard to argue that the relation between religious names and religious occupations is spurious. Second, the mentioned selection bias is likely to go against finding the documented results. For instance, if the true relation between religious occupations and names is random, the mentioned selection would create a *negative* correlation between religious names and religious occupations, which is the opposite of what we find.[60] Thus, the selection would have to be of a type where the likelihood of becoming an author is higher for those with religious occupations, but lower for those with religious names. Third, a related concern is that selection into authorship is higher for individuals with religious names and for scientists, which could generate a spurious correlation between religious names and science occupations. Again, this selection has to be of a rather peculiar type in order to explain results. Since results hold after excluding authors with religious occupations, the selection based on religious names should be unrelated to their religious occupations (cf panel B of Table 5). Thus, something about the name, which is unrelated to other religious dimensions should be causing the selection, which seems unlikely to explain the estimate sizes documented. Fourth, if selection explains the results, increasing the selection bias should strengthen results. Potential selection is arguably stronger for more successful authors, as they face a higher likelihood of ending up in a library. We measure the "succes" of authors by being registered in another online register, the Worldcat and by having written texts that are more cited. The results do not strengthen with either definition (Table D2).

A fourth alternative explanation is the practice of "giving" younger sons to the priesthood, while the eldest son was typically expected to inherit the family's property. This could explain results in Table 3 if parents ere more likely to give religious names to younger sons than to the eldest son. We judge that this cannot explain results. First, again the correlations between names and regional religiosity and the rise in religious names after earthquakes would not arise if religious names do not reflect a general family-level religiosity. Second, the practice was most prevalent among Catholics and declined in many Protestant regions after the Reformation in the 16th century (Le Goff, 1986). Thus, the fact that the results hold for Catholics and Protestants and continue to hold well into the 20th century make them hard to explain with this practice (Figures B4 and G2 and Tables B6 and G6). Third, due to the exclusion of women from the priesthood in many religious traditions, including the Catholic Church, the practice of dedicating daughters to religious service was limited. Results hold when restricting to the 25,000 women in the sample, indicating that the explanation is not male-specific (Table D3).

Last, the spike in religious names following earthquakes is difficult to explain with any of the concerns raised. We can leverage this shock to obtain exogenous variation in

[60]To realize this, think of Religious occupation as "Random noise 1" and Religious name as "Random noise 2" in Figure D1. The red circle shows a sample selection based on religious names and occupations.

religiosity in order to further probe its causal impact on occupation choice. However, as earthquakes impact society in multiple ways, incurring potential direct shifts in occupation choice, we would not want to use them directly. One concern is that earthquakes lead to an increased demand for builders, thus crowding out other occupations, such as those in the scientific field. Before proceeding, we note that it is not obvious that any potential direct effects of earthquakes should reduce the demand for scientists. Disasters are equally likely to *increase* the demand for scientists, as disasters may spark a desire for scientific understanding of the phenomenon.[61] Furthermore, if earthquakes lead to crowding out of other occupations, we would expect them to impact occupations in religion and science in the same direction. This is not what we find below.

Nevertheless, we conduct an analysis focusing on the emotional effects of earthquakes. We rely on previous research documenting that i) religion is used for emotional coping, even by individuals who are not influenced materially by the disaster[62] and ii) parents hit by earthquakes may experience a strengthened role of religion, which they transfer to their children. The mere knowledge about a recent earthquake can instigate religious coping, either through thoughts of adversity[63] or because friends or family were affected, both of which may instigate religious coping, independently of the physical damage. The rationale behind our analysis is to remove variation due to physical damage as much as possible. First, we exclude authors born within 20 km of an earthquake in all samples.[64] Informed by the dynamic analysis in Figure 4, we add dummies equal to one if an earthquake hit within 20 years before or after birth. Earthquakes raise the likelihood of having a religious name or a religious occupation, and reduce the likelihood of becoming a scientist (religious occupations not significantly, though, cf columns 1, 5, and 9 of Table 8).

Second, we leverage disparities in expected effects dependent on the timing of earthquakes. Consistent with the idea that earthquakes raise the religiosity of parents influencing their naming decision, only earthquakes before birth matter for naming (column 1 and Figure 4). For occupations, earthquakes before *and* after birth matter, consistent

[61] Innovation efforts towards environmentally friendly technologies rose in U.S. counties exposed to the Dust Bowl in the 1930s (Moscona, 2021), natural disasters lead to more risk-mitigating innovations (Miao & Popp, 2014), solar eclipses raised curiosity and the need for explaining, thus increasing social complexity and technological progress in pre-historic societies (Litina & Fernández, 2020), and the Great Influenza pandemic led to increased innovation and scientific occupation choice for individuals from less religious backgrounds, but instead strengthened religiosity for devout individuals (Berkes *et al.*, 2023).

[62] Bentzen (2019a) documents that earthquakes raise religiosity, even after removing districts that were hit. The paper also includes a case study of an earthquake sized 6.0 on the Richter scale that hit Napa Valley in 2014. Bentzen shows that Google searches for the earthquake rose in Napa Valley, but also in surrounding Metropolitan areas and even in surrounding states.

[63] Norenzayan & Hansen (2006) show that experimentally induced thoughts of death raise religiosity.

[64] Many earthquakes result in destruction further than 20 km. Results hold if we remove authors hit by the larger earthquakes. Available upon request.

with the idea that earthquakes can strengthen religious upbringing and the person's own religiosity later in life. For science occupation choice, the estimate before birth is nearly twice the size of the estimate after birth.[65] This larger impact before birth for choices 20-40 years later is hard to explain if not for inter-generational transmission of religiosity.

Third, we exploit that individuals are arguably more likely to be concerned with adversity in their own country, rather than in neighboring countries. Indeed, the effects on naming and occupation choice are solely attributed to earthquakes that occur in the authors' country of birth (columns 2, 6, and 10). An obvious reason may be the proximity of earthquakes that occur in one's own country. To accommodate, we control for earthquake distance in columns (3), (7), and (11). Results are driven by earthquakes that strike within the person's own country, even for earthquakes of similar distance.

Table 8: Earthquakes as shocks to religiosity

Dependent variable:		Religious name				Religious occupation				Scientific occupation		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Earthquake before birth	0.031**	-0.032	-0.12**	-0.14**	0.0027	-0.016**	-0.022	-0.037	-0.027***	0.019	0.068	0.11**
	(0.013)	(0.022)	(0.055)	(0.068)	(0.007)	(0.008)	(0.028)	(0.038)	(0.010)	(0.027)	(0.043)	(0.049)
Earthquake after birth	-0.022	-0.021	0.050	-0.018	0.0036	-0.014*	-0.0077	0.010	-0.016**	0.016	-0.030	-0.026
	(0.014)	(0.030)	(0.068)	(0.070)	(0.005)	(0.008)	(0.028)	(0.037)	(0.008)	(0.014)	(0.037)	(0.044)
Earthquake before birth same country		0.077***	0.076***	0.061**		0.024**	0.024**	0.028**		-0.057**	-0.056**	-0.054**
		(0.027)	(0.025)	(0.026)		(0.010)	(0.010)	(0.011)		(0.027)	(0.027)	(0.024)
Earthquake after birth same country		0.00018	0.0039	-0.022		0.022**	0.022**	0.013		-0.040***	-0.043***	-0.032**
		(0.034)	(0.033)	(0.045)		(0.011)	(0.011)	(0.013)		(0.015)	(0.015)	(0.016)
Avg earthquake distance before birth			-0.49*	-0.57*			-0.030	-0.092			0.27	0.41
			(0.258)	(0.330)			(0.155)	(0.203)			(0.182)	(0.250)
Avg earthquake distance after birth			0.39	-0.075			0.034	0.080			-0.26	-0.18
			(0.349)	(0.326)			(0.153)	(0.190)			(0.189)	(0.226)
R squared	0.072	0.072	0.072	0.072	0.052	0.052	0.052	0.053	0.042	0.042	0.042	0.021
Observations	196000	196000	196000	88383	196000	196000	196000	88383	196000	196000	196000	88383
Mean dep var	0.54	0.54	0.54	0.57	0.076	0.076	0.076	0.092	0.18	0.18	0.18	0.15
5x5 grid and Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
2x2 grid trends	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Sample	Full	Full	Full	Migrant	Full	Full	Full	Migrant	Full	Full	Full	Migrant

OLS across authors. The dependent variable is a dummy equal to one if the author shares name with a significant religious figure in columns (1)-(4), has a religious occupation in columns (5)-(8), or has a scientific occupation in columns (9)-(12). "Earthquake before birth" is a dummy equal to one if an earthquake hit within 100 km of the author's birthplace during the 20 years before birth. "Earthquake after birth" is a dummy equal to one if an earthquake hit within 100 km of the author's birthplace during the 20 years after birth. Baseline controls included throughout: Top-10 most common names, gender, number occupations, 5x5 grid cell fixed effects, year of birth fixed effects, and 2x2 grid cell trends. The sample excludes authors born within 20 km of an earthquake 20 years before or after their birth and excludes authors further away than 500 km from earthquake that hit within 60 years before or after their birth. Robust standard errors clustered at the 2x2 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Earthquakes raise the likelihood of having a religious name or religious occupation and lower the likelihood of becoming scientists. For naming, only earthquakes before birth matter, while occupation choice is influenced by earthquakes before and after birth. These effects are particularly pronounced for earthquakes that hit in the author's country of birth and persist for authors that have migrated since their birth.

Fourth, while these results are consistent with religious coping, they are also consistent with occupation choice shifts that could be stronger in ones' own country, independently of religion. To address, we restrict the sample to individuals who die in a different location from where they were born (columns 4, 8, and 12).[66] As we measure earthquakes in the birth location, this specification is not biased due to direct crowding out effects from an elevated demand for, say, builders. The impact of earthquakes before birth is unchanged, consistent with an impact on religious upbringing. However, earthquakes after birth now matter less for occupation choice (25-40% less), consistent with the idea that these earthquakes most likely do not influence the person as much in their new location.

Fifth, results carry through when using earthquakes as instruments for religious nam-

[65] Calculated based on the standardized beta coefficients of the estimates in column (9), which are -0.0112 and -0.0057, respectively.

[66] The information on the location of death is available for 113,607 of the 196,000 authors in Table 8. Of these, 88,383 died in a place that was different from their birth place.

ing in regressions on occupation choice (Table D4). Religious upbringing (proxied by having been given a religious name) raises the likelihood of ending up in religious occupations and lowers the likelihood of becoming a scientist. In the strictest specification, the exogenous variation in religious upbringing comes exclusively from earthquakes hitting before birth in ones home country. We compare individuals who no longer live where they were born, are hit by the same number of earthquakes after birth, within the same distance to earthquakes, are born in the same year, with the same gender, name commonality, and who are all authors.

We conclude that increased religiosity of parents seems to have a causal impact on children's occupation choices later in life. We next attempt to understand this through the lens of existing literature and examine how this matters for aggregate economic development.

5 Religiosity, science, and modern growth

When examining the causes of economic activity throughout European modern history, scholars have documented the importance of law and legal institutions for the Commercial Revolution starting in the 11th century (Van Zanden, 2009; Cantoni & Yuchtman, 2014), of upper-tail knowledge and a rising appreciation of reason for the Enlightenment (Mokyr, 2010; Squicciarini & Voigtländer, 2015; De la Croix *et al.*, 2020a), of coal, cultural values, property rights institutions and the fertility transition for the Industrial Revolution (Acemoglu *et al.*, 2005; Clark, 2008; Allen, 2011; Galor, 2011), and the development of modern science and engineering for the Technological Revolution (Mokyr & Voth, 2009; Squicciarini & Voigtländer, 2015). Religion –and religiosity in particular– may have given rise to different trajectories for societies in each of these revolutions.

The examination of the role played by religion in these processes is complicated by the fact that religion is a multidimensional concept which can manifest itself in multiple ways. Hence, it is not obvious from the outset whether its implications for the development of modern science are positive or negative. For instance, the historical record provides countless examples of technological breakthroughs made by people of strong religious faith, a famous example being Isaac Newton. Also, many of Europe's leading universities were established by monasteries. In addition, the medieval Church in Europe may have kept the power of secular rulers in check (Hoffman, 2015; Rubin, 2017), supported primogeniture, potentially strengthening political stability (Acharya & Lee, 2019), and may have been better organized than its royal counterparts (Stark & Finke, 2000;

Ullmann, 1970).[67] Extending beyond examples, research has documented benefits of religion, such as stress-relief (Pargament, 2001), happiness (Campante & Yanagizawa-Drott, 2015), pro-sociality (Purzycki *et al.*, 2016; Henrich *et al.*, 2010), and the emergence of early states (Norenzayan, 2013). These factors may well have spurred the development of science and eventually modern growth.

There are also cases to make for negative effects of religiosity on science and modern growth. For instance, expert on the Enlightenment and its importance for modern growth, Joel Mokyr, reasons that (Mokyr, 2012, 39): "*For an economy to create the technical advances that enabled it to make the huge leap of modern growth, it needed a culture of innovation, one in which new and sometimes radical ideas were respected and encouraged [...]*" According to Mokyr, this cultural appreciation of knowledge and reason emerged during the 17th century and took over from the traditional "*worship of ancient wisdom*" (Mokyr, 2016b). This potential historical link between science and religion was formalized in a theoretical model by Cantoni & Yuchtman (2013), who argue that the emergence of modern science depends on whether the government or an elite see the new knowledge as threatening to their position of power. Cantoni and Yuchtman interpret two historical cases through this lens: The introduction of Roman law in medieval Europe and Western science in late imperial China. In Europe, the states and Church found individuals trained in Roman law valuable, whereas elites in late 19th-century China felt threatened by the introduction of Western science and engineering and instead continued to base their selection criteria of civil servants on their knowledge of Confucian classics. Cantoni and Yuchtman extend their arguments to the Muslim world, noting that Islamic elites who controlled educational institutions in the Middle Ages initially promoted the study of logic and science, because the gains from spreading these skills outweighed the costs (potential criticism of the religious elites). This period of elite support for scientific study saw the flourishing of Islamic society. However, as Islam took over, the gains to elites from the study of philosophy and science declined. Elites started to oppress the study of science, and the Muslim world fell behind.

Bénabou *et al.* (2022) model the potential challenge between the worldviews of science and religion. In their model, the recurrent arrival of scientific discoveries generates productivity gains and at the same time threatens to erode religious beliefs by contradicting aspects of religious doctrines. As a result, religious elites have incentives to curb the development of science and new ideas. Examples abound of such conflicts, from the trial of Galilei in 1633 to President George W. Bush's restrictions of federal funding for

[67]See Grzymala-Busse (2020) for a review.

embryonic stem-cell research. In support of the model, Bénabou *et al.* document fewer patents and other measures of innovation per capita in contemporary countries or U.S. states that are more religious in surveys.

Chaney (2016) provides econometric support for the idea that religion may have played a crucial role for the decline of science in the Muslim world. Chaney links the fall of scientific and technological production in the Muslim world in the late Medieval period to the extensive spread of madrasas, educational centers where Islamic law was taught. Using data on book production during the period 800-1800, Chaney documents a rising trend in the proportion of religious books written in the Middle East, accompanied by a drop in original and scientific books.

Squicciarini (2020) documented that French areas with higher religiosity were less likely to introduce new technical curriculum in schools, thus forgoing skills essential for experiencing the Technological Revolution. In France, the Catholic Church promoted a conservative, anti-scientific program that hindered the introduction of technical curricula, while pushing for religious education.

It thus remains an empirical question whether religion and science contested or complemented one another in European history and whether the potential impact of religiosity was sufficiently large to matter for economic growth. Furthermore, religiosity in the reviewed literature is mostly modeled as working through institutions. However, there are reasons to believe that religiosity influences behavior directly, including choice of occupation (Lehrer, 2008; Audretsch *et al.*, 2013). Occupational choice can be modeled as determined by external and internal factors, the latter including factors such as ability, preferences, and parental value orientations. In their seminal work, Blau *et al.* (1956, 537) note that "*people's value orientations determine the relative significance of different kinds of rewards and thus the attractive force exerted by them.*"[68] In the same manner, religious values could make some occupations and fields of study more (or less) intrinsically meaningful and valuable to individuals or their parents.

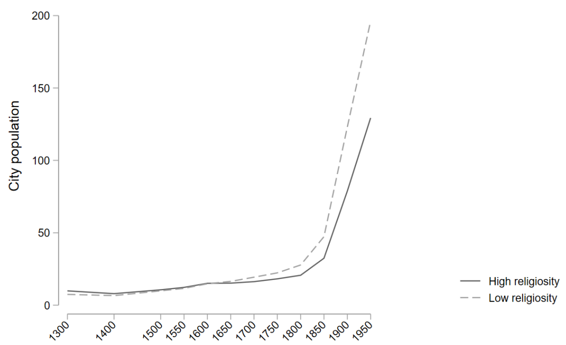
5.1 Addendum: Economic growth

We interpret the results so far as indicating that individuals characterized by a more religious upbringing were less likely to proceed with advanced studies or occupy occupations in science and engineering. If these types of knowledge production influenced economic growth as argued by the reviewed literature, we expect religiosity to have slowed growth as we pass the Enlightenment and particularly the Technological Revolution. To examine,

[68] Another seminal paper is Sewell *et al.* (1970). For more recent summaries of occupational choice research, see Albert & Weeden (2011) and Jonsson *et al.* (2009).

we employ a widely used measure of past economic prosperity; urbanization rates.[69] We use data on the population of 2,262 European cities from year 700 to today constructed by Buringh (2021).[70] As a proxy for the religiosity of a city, we compute the share of authors with religious names born within 100 km of a city center in the particular period. We restrict analysis to cities with at least 10 authors born in each period.[71]

Figure 5: Simple relation between religiosity and economic growth



Average city size over time. The cities were split in two by the median level of religiosity during the period 1300-1600.

As a first illustration of the results, we split cities into two equally sized groups based on the share of religiously named authors during the period 1300-1600, the earliest period for which names provide a reliable measure of religiosity (cf panel c of Figure B4). The sample for this exercise includes 417 cities. For each available year with data on city population, we calculate average city size for the two groups and depict these averages in Figure 5. The two groups were similar in terms of average city size until around 1600, where-after the less religious cities grew faster, particularly from 1800 on-wards. This difference coincides with the Enlightenment and later the increased importance of technological knowledge in production during the Technological Revolution (Mokyr, 2010; Squicciarini & Voigtländer, 2015). The intensified gap also coincides with the point in time when religiosity started to matter for scientific occupation choice in the previous analysis.[72] The figure also reveals a slight tendency for the more religious cities to be larger than the less religious cities before 1600, consistent with the historical examples of early universities and inventions being associated with monasteries.[73]

[69] Acemoglu *et al.* (2005); Bosker *et al.* (2013); Dittmar (2011); Nunn & Qian (2011).

[70] This is an updated version of Bairoch *et al.* (1988), which includes 2,200 European cities with a minimum of 5,000 inhabitants, measured in centuries between 800 and 1700 and half centuries 1750 and 1800. Buringh (2021) extended the dataset with an additional 62 cities that had reached 100,000 inhabitants by year 2000 and included additional years 700, 1100, 1550, 1650, 1900, 1950, and 2000.

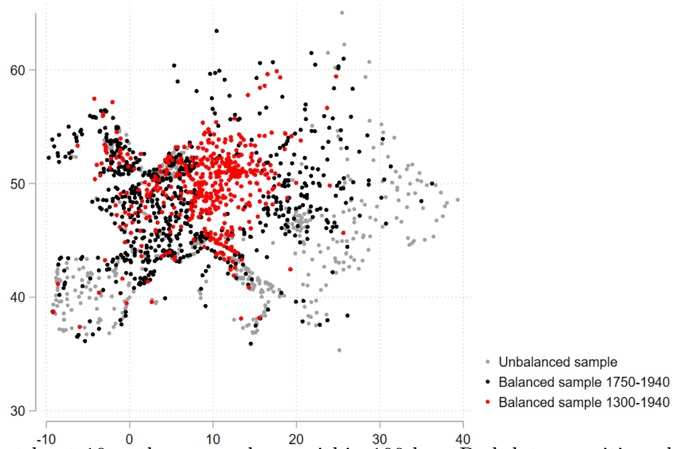
[71] We match an author to the city located closest to their birth place. Thus, for authors who were born within 100 km of more than one city, they will contribute to the average for the closest city only.

[72] The formal econometric analysis below reveals that the slightly earlier timing in Figure 5 is not significant. Instead, the difference based on religiosity emerges only after 1750, consistent with the individual-level analysis.

[73] The pattern of similar growth and slightly larger religious cities extends back in time to year 700

Since religiosity is measured before most population measures in Figure 5, we are not concerned about reverse causality, but the pattern may be caused by other factors that correlate with religiosity and city size simultaneously. To get a sense of the extent to which omitted factors may contaminate results, we conduct a simple balancing test of whether more or less religious cities differed systematically in terms of potentially relevant observables before 1300 (Table H1). As we do not have a clear pre and post period, the balancing test should be interpreted with caution.[74] We find no differences before 1300 between cities that were more or less religious in 1300-1600 in terms of growth rates, share of women or people with common names, vicinity to the Atlantic coast, or longitudes. The more religious cities, though, were more populated, more likely to be Catholic, and were more concentrated inland and in Southern Europe. These differences disappear when adding 5x5 degree grid cell fixed effects, which amount to 500x500 km2 grids at the Equator. This exercise thus enhances our confidence that the econometric analysis that follows is not greatly biased by unobservables.

Figure 6: Cities for analysis



The dots are cities where at least 10 authors were born within 100 km. Red dots are cities where at least 10 authors were born in each of the four time-periods between 1300 and 1940. Black dots are additional cities with at least 10 authors born between 1750 and 1940. Gray dots are additional cities with at least 10 authors born at any point in time.

To conduct a more formal analysis, we calculate time-varying religious name shares. Acknowledging that average religiosity does not change much over time, we calculate the shares in intervals of around 100 years and match them to the city population numbers. The lengthy intervals also ensure that enough authors were born within 100 km of the city center during the period. We match city population in 1950 to the share of authors

(panel a of Figure H1). The differential rise in the growth of less religious cities may have started already around year 1400 (panel b Figure H1).

[74]We conduct analysis similar to Hornbeck & Naidu (2014), where we restrict the sample to the period 700-1300 and regress each dependent variable on a dummy equal to one if a city had above-median share of authors with religious names during the period 1300-1600. We add all controls from Table 9 consecutively, except that instead of city fixed effects, we add 28 5x5 grid cell fixed effects, as the religiosity variable in this analysis does not vary over time.

with religious names born between 1850 and 1940, city population in 1850 to the share of religiously named authors born between 1750 and 1849, and city population in 1750 to the share of religiously named authors born between 1650 and 1749.[75] Due to few authors as we go further back in time, we will not attempt to calculate the religious name shares in intervals before 1650, but instead keep city population in 1650 as our earliest year in this analysis and match this to authors born between 1300 and 1649.

The cities in our samples are shown in Figure 6. Red indicates that at least 10 authors were born in all of the four intervals. This balanced set of 435 cities forms our main sample. As the dataset includes more authors born in the later periods, we also conduct analyses for cities with at least 10 authors born in each of the latter two periods (indicated by either black or red dots). Gray indicates cities with at least 10 authors born in at least one of the intervals, but that are not part of either balanced sample. Our balanced samples consist of cities spread across all of Western Europe, but somewhat more concentrated in the central part. We estimate equations of the form:

$$growth_{ct} = \alpha_0 + \alpha_1 size_{ct-1} + \sum_{t=1650}^{1950} \beta_t relname_{ct} + \gamma_c + \gamma_t + \lambda t_c + \omega X_{ct} + \varepsilon_{ct} \tag{3}$$

where $growth_{ct}$ is the growth rate of the population size of city c in century t, calculated as (log 1+) the population size of city c at time t minus (log 1+) the population size of city c at time $t-1$.[76] $size_{ct-1}$ is the (log 1+) city population size at time $t-1$.[77] $relname_{ct}$ is the share of authors with a religious name born within 100 km of city c in the period leading up to time t. γ_c are city fixed effects, γ_t are century fixed effects, and t_c are city-specific time-trends. X_{ct} are sets of control variables for city c at time t, including the share of female authors or authors with common names born within 100km of a city.[78] We cluster standard errors at the city level. We allow the impact of religious name shares to vary over time, indicated by β_t.

Table 9 shows results for the unbalanced sample of cities (column 1), the balanced

[75]The most recent period, 1850-1940, ends 10 years before the city population measure, since our main author data ends in 1940. We judge that the bias arising from not having authors born throughout the full period 1850-1949 is small. The measure of city sizes, however, is exactly 100 years apart throughout this analysis, which is crucial for the comparison of the effects over time.

[76]To ease readability of the estimates, the growth rates capture the total growth during each century, instead of the yearly growth rates.

[77]By including the initial level, a concern is that the estimates are biased by Nickell bias (Nickell, 1981). Our results are nearly identical when not controlling for initial population, when controlling for population in the first period interacted with time dummies, or when having the level of the population size as the dependent variable, indicating that Nickell bias cannot explain our results (Tables H2 and H3). Also, the working paper version of this paper had all city-level regressions with (log) city population size as dependent variable. We find that growth rates make for a more intuitive interpretation.

[78]Both of these are calculated over the same time-intervals as the religious name shares.

sample of cities for each of the periods between 1300 and 1940 (columns 2-8), and in the sample of cities in the periods between 1750 and 1940 (column 9). Columns (1) and (2) show the simple correlations, revealing that cities with larger shares of religiously named authors experienced lower growth rates. By adding city fixed effects in column (3), we find that city growth follows falling religiosity. The estimate drops some when adding time fixed effects in column (4) and city-specific trends in column (5) disclosing that part of the raw correlation is due to the fact that cities are growing and religiosity is falling over time. These trends may or may not be causally related and we remove them to be conservative. The estimate on religious name shares remains significant.

Table 9: Religiosity and economic growth

Dependent variable: City population growth	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(log) Initial population size	0.0058	0.055***	0.074**	-0.39***	-1.16***	-1.19***	-1.19***	-0.44***	-0.88***
	(0.013)	(0.012)	(0.034)	(0.036)	(0.076)	(0.075)	(0.075)	(0.038)	(0.046)
Religious name share	-2.07***	-2.12***	-2.51***	-0.65***	-0.32**				
	(0.077)	(0.095)	(0.117)	(0.127)	(0.151)				
Religious name 1300-1649 x 1650						0.14	0.17	-0.092	
						(0.345)	(0.354)	(0.202)	
Religious name 1650-1749 x 1750						0.21	0.26	0.14	
						(0.158)	(0.168)	(0.179)	
Religious name 1750-1849 x 1850						-0.87***	-0.84***	-1.31***	-0.42**
						(0.251)	(0.254)	(0.197)	(0.205)
Religious name 1850-1940 x 1950						-1.77***	-1.59***	-1.98***	-2.04***
						(0.388)	(0.412)	(0.308)	(0.218)
Common name share							-0.18	0.046	-0.35***
							(0.133)	(0.110)	(0.135)
Female share							0.16	0.24	1.09***
							(0.341)	(0.286)	(0.372)
R-squared	0.19	0.30	0.34	0.54	0.75	0.76	0.76	0.56	0.47
Observations	3809	1740	1740	1740	1740	1740	1740	1740	2208
Mean dep var	0.69	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.86
Balanced sample	N	Y	Y	Y	Y	Y	Y	Y	Y
City FE	N	N	Y	Y	Y	Y	Y	Y	Y
Year FE	N	N	N	Y	Y	Y	Y	Y	Y
City-trends	N	N	N	N	Y	Y	Y	N	N
Number cities	1603	435	435	435	435	435	435	435	1104

OLS estimates across cities and time. The dependent variable is population growth. Religious name share is the share of authors born within 100 km of the city center who share name with a religious figure. The sample is the full unbalanced sample of cities with at least 10 authors born within 100 km of the city center in column (1), the balanced sample of 435 cities for which at least 10 authors were born within 100 km of the city center in each of the four periods in columns (2)-(8), and the balanced sample of 1104 cities for which at least 10 authors were born within 100 km of the city center in the two last periods in column (9). Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities with an increasing share of religious authors experienced slower growth rates from 1750 onward.

Columns (6)-(9) allow the correlation between religiosity and city growth to vary over time. Consistent with what Figure 5 indicated, the negative association strengthens over time. Column (7) includes the controls for shares of women and authors with common names. These do not exert an independent impact on city growth in this sample, neither linearly or when the impact is allowed to vary over time (Table H4). Restricting to the last two periods in column (9) nearly triples the number of included cities and increases the number of countries represented in the sample from 24 to 37. The negative association between religiosity and city growth persists in this arguably more representative sample of cities.[79] As the last set of regressions in column (9) include only two periods, the city-

[79] Also, cities with a larger share of women and a lower share of individuals with common names experience higher growth rates in this sample, consistent with a literature on the economic benefits of

specific trends are caught by the time fixed effects. To ease comparison, we show the long balanced sample without city-specific trends in column (8), revealing that the inclusion of city-specific trends does not influence the estimate on the religious name shares much.

Taking the estimate on religiosity in the late 1800s and early 1900s in column (7), -1.6, at face value, moving from a city with no authors with religious names to a city with a full population of religiously named authors cuts city growth by three times the average growth rate. This is an implausibly large change in religiosity, though.[80] Instead, if we take a city in the 1900s and increase its' religiosity level from being among the 10% least religious cities (less than 38% religious names) to being among the 10% most religious cities (above 76% religious names), the model predicts that the growth rate falls by 0.38 units, amounting to two thirds of the average growth rate (0.55). The effects are large, but not implausibly so.

5.1.1 Identification and mechanisms at the city-level

The estimates in Table 9 are not driven by few observations, countries or religious denominations (cf. Figures H2 and H3). Reverse causality is again not an issue as we proxy the religiosity of parents or the degree of religion in a person's upbringing during the century leading up to the end-year of city growth. Since we add city fixed effects throughout, we mute confounders that are constant over time. We remove all general century-specific events with the time fixed effects and linear trends in religiosity or other confounders by adding city-specific trends. However, religiosity and city size are both rather persistent phenomena and may correlate with confounders not accounted for. To accommodate, we proceed by adding an extended list of controls. We will not leverage earthquakes as exogenous variation in religiosity, as we do not want to make the claim that religiosity was the only mechanism through which earthquakes mattered for growth. At the same time, we do not have the same detail of the data to account for these direct effects as convincingly as in the individual-level analysis.

First, the results are unchanged when allowing the impact of the control variables –share with common names, women, and the initial population size– to vary over time (Table H4). Additionally, the results hold up to adding measures of ethnicity (Table H5), migration (Table H6),[81] dummies interacted with time for cities located within 100 km of

individualism (Gorodnichenko & Roland, 2011).

[80]When interpreting results at the individual level, one unit change meant a change in whether or not one person had a religious name. In the city-level analysis, one unit amounts to changing from a city where none of the authors had religious names to a city where all authors were named religiously.

[81]We measure the share of authors who did not die in the same location as they were born. We attach this share either to their birth place or their death place. Thus, this analysis also shows robustness towards attaching the religious name shares to the authors' location of birth or death. This analysis is

the Atlantic ocean (Table H7), the latter to accommodate a concern that Atlantic trade drives our results, inspired by Acemoglu *et al.* (2005). Cities located on the Atlantic coast or any other coast do not experience higher growth in our samples, the reason being that many large cities in recent times, such as Berlin and Paris, are not situated on a coast.

We interpret the results as indicating that more religious cities experienced slower economic growth. Based on the individual-level analysis and the existing literature, we envision one mechanism that individuals for whom religion is more important are less likely to become scientists, an occupation that became increasingly important for modern growth. In support, we find that city growth slowed in cities with increasing shares of authors in religious occupations, while cities with increasing shares of scientists grew more (Table H8). However, while the sign on the estimates are as predicted by the literature, the estimates are only significant at conventional levels in specifications without city-specific trends. This means that we cannot distinguish the association between city growth and occupations from city-specific trends. Part of the reason may be that the number of cities thus halves in this analysis. Last a note of cation; since this analysis is based on actual occupations, causality may well run both ways.

As a last test of the mechanism, we exploit data on the number of innovators and patents in Europe collected by Bergeaud & Cyril (2022). To match with the data on authors and cities, we restrict to innovators who registered a patent in Europe at some point before 1950. Further, we note that the data include only 379 innovators who registered a patent before 1850. We exclude these, leaving us with data on 2,527,983 innovators who registered a patent between 1850 and 1950. For each city, we calculate the number of innovators who registered a patent within 100 km of the city center. We find that cities with larger shares of religious authors (as measured by naming patterns) have fewer innovators, even after controlling for the mentioned controls (Table H9). At the same time, cities inhabited by a larger share of scientist have more innovators and last, cities with more innovators manage to experience higher economic growth rates.

To conclude, we interpret these latter results as consistent with existing theoretical models and literature: Differences in religiosity across European societies may have been influential in terms of affecting their adoption of modern science, engagement in innovative activity, and thus for their production of modern economic growth. Again, we emphasize that these city-level results should be interpreted with caution. Omitted confounders may have raised economic growth and at the same time reduced religiosity in a similar

exclusively conducted for the late-panel 1750-1950, as we do not have information on their place of death for enough authors in the earlier periods.

increasing manner as that documented.

6 Concluding remarks

The intensity of religious beliefs of our ancestors can be proxied by the names they gave their children. We find empirical support for this across two large independent databases of authors and university students. On average, individuals with religious names were more likely to later become priests or theologians. In addition, parents experiencing a major earthquake before having a child were more likely to give that child a religious name. Last, the share of religious names is higher in regions that are otherwise regarded as more religious.

While philosophers have long emphasized the potential socioeconomic impact of religion, the empirical study of the phenomenon is still in its nascent stage. It is our hope that this novel measure of religiosity can propel a significant advancement in this research, testing whether hypotheses based on historical narrative hold across time and space, for instance.

In this paper, we use the names-based religiosity measures to address one central question: How did religiosity influence the development of science and was the impact large enough to influence the trajectory to modern economic growth? We find that religiously raised individuals have been less likely to become scientists or engineers. To establish causality, we leverage shocks to religious upbringing from earthquakes in migrants' birth places.

For identification purposes, we have removed regional variation by including grid cell fixed effects, meaning that our results reflect individual-level effects. However, in keeping with models by e.g., Cantoni & Yuchtman (2013), Chaney (2016), and Bénabou *et al.* (2022), religion may have caused institutional barriers to the development of science in addition to the effects we identify. We thus interpret our estimates as lower bounds of the true impact of religion on modern science production.

Consistent with a literature emphasizing the significance of science and technology for the transition to modern growth, we last document that economic growth was lower in cities with higher religiosity levels. The effect emerges after 1750, coinciding with the point in time where science and technology became increasingly important for economic growth. Religion may well have come with simultaneous emotional benefits, such as happiness and improved ability to cope with stress. However, our findings suggest that over time, religion became a stumbling block for the emergence and spread of science and knowledge useful for the launch of modern growth.

References

Abramitzky, Ran, Boustan, Leah, & Eriksson, Katherine. 2020. Do immigrants assimilate more slowly today than in the past? *American Economic Review: Insights*, **2**(1), 125–41.

Acemoglu, Daron, Johnson, Simon, & Robinson, James A. 2001. The colonial origins of comparative development: An empirical investigation. *American economic review*, **91**(5), 1369–1401.

Acemoglu, Daron, Johnson, Simon, & Robinson, James. 2005. The rise of Europe: Atlantic trade, institutional change, and economic growth. *American economic review*, **95**(3), 546–579.

Acharya, Avidit, & Lee, Alexander. 2019. Path dependence in European development: medieval politics, conflict, and state building. *Comparative Political Studies*, **52**(13-14), 2171–2206.

Albert, Kyle, & Weeden, Kim. 2011. *Occupations and professions.* Oxford University Press.

Allen, Robert C. 2011. Why the industrial revolution was British: commerce, induced invention, and the scientific revolution 1. *The Economic History Review*, **64**(2), 357–384.

Andersen, Lars Harhoff. 2021. The Germanic Origins of Comparative Democracy. *Unpublished Manuscript.*

Andersen, Lars Harhoff, & Bentzen, Jeanet. 2022. In the Name of God! Religiosity and the Transition to Modern Growth.

Assouad, Lydia. 2020. Charismatic Leaders and Nation Building.

Audretsch, David B, Boente, Werner, & Tamvada, Jagannadha Pawan. 2013. Religion, social class, and entrepreneurial choice. *Journal of Business Venturing*, **28**(6), 774–789.

Bairoch, Paul, Batou, Jean, & Pierre, Chevre. 1988. *Population des villes européennes de 800 à 1850: banque de données et analyse sommaire des résultats (la).* Librairie Droz.

Barro, Robert J, & McCleary, Rachel M. 2003. Religion and economic growth across countries. *American Sociological Review*, **68**(5), 760.

Barro, Robert J, & McCleary, Rachel M. 2016. Saints marching in, 1590–2012. *Economica*, **83**(331), 385–415.

Bazzi, Samuel, Fiszbein, Martin, & Gebresilasse, Mesay. 2020. Frontier culture: The roots and persistence of "rugged individualism" in the United States. *Econometrica*, **88**(6), 2329–2368.

Becker, Sascha O, & Pascali, Luigi. 2019. Religion, division of labor, and conflict: Anti-Semitism in Germany over 600 years. *American Economic Review*, **109**(5), 1764–1804.

Becker, Sascha O, & Woessmann, Ludger. 2009. Was Weber wrong? A human capital theory of Protestant economic history. *The quarterly journal of economics*, **124**(2), 531–596.

Becker, Sascha O, Pfaff, Steven, & Rubin, Jared. 2016. Causes and consequences of the Protestant Reformation. *Explorations in Economic History*, **62**, 1–25.

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

Bentzen, Jeanet Sinding. 2019a. Acts of God? Religiosity and natural disasters across subnational world districts. *The Economic Journal*, **129**(622), 2295–2321.

Bentzen, Jeanet Sinding. 2019b. Why are some societies more religious than others? *Pages 265–281 of: Advances in the Economics of Religion*. Springer.

Bentzen, Jeanet Sinding. 2021a. In crisis, we pray: Religiosity and the COVID-19 pandemic. *Journal of economic behavior & organization*, **192**, 541–583.

Bentzen, Jeanet Sinding. 2021b. Religiosity and Development. *Oxford Research Encyclopedia*.

Bergeaud, Antonin, & Cyril, Verluise. 2022. *PatentCity: a dataset to study the location of patents since the 19th century.*

Berkes, Enrico, Coluccia, Davide Maria, Dossi, Gaia, & Squicciarini, Mara P. 2023. *Dealing With Adversity: Religiosity or Science? Evidence From the Great Influenza Pandemic*. Tech. rept. CEPR Discussion Papers.

Blau, Peter M, Gustad, John W, Jessor, Richard, Parnes, Herbert S, & Wilcock, Richard C. 1956. Occupational choice: A conceptual framework. *ILR Review*, **9**(4), 531–543.

Bosker, Maarten, Buringh, Eltjo, & Van Zanden, Jan Luiten. 2013. From baghdad to london: Unraveling urban development in europe, the middle east, and north africa, 800–1800. *Review of Economics and Statistics*, **95**(4), 1418–1437.

Bruce, Steve. 2002. *God is dead: Secularization in the West*. Vol. 3. Blackwell Oxford.

Brundage, James A. 2008. The Medieval Origins of the Legal Profession. *In: The Medieval Origins of the Legal Profession*. University of Chicago Press.

Buringh, Eltjo. 2021. The Population of European Cities from 700 to 2000: Social and Economic History. *Research Data Journal for the Humanities and Social Sciences*, **1**(aop), 1–18.

Buringh, Eltjo, Campbell, Bruce MS, Rijpma, Auke, & van Zanden, Jan Luiten. 2020. Church building and the economy during Europe's 'Age of the Cathedrals', 700–1500 CE. *Explorations in Economic History*, **76**, 101316.

Callaway, Brantly, & Sant'Anna, Pedro HC. 2021. Difference-in-differences with multiple time periods. *Journal of Econometrics*, **225**(2), 200–230.

Campante, Filipe, & Yanagizawa-Drott, David. 2015. Does religion affect economic growth and happiness? Evidence from Ramadan. *The Quarterly Journal of Economics*, **130**(2), 615–658.

Cantoni, Davide. 2015. The economic effects of the Protestant Reformation: testing the Weber hypothesis in the German lands. *Journal of the European Economic Association*, **13**(4), 561–598.

Cantoni, Davide, & Yuchtman, Noam. 2013. The political economy of educational content and development: Lessons from history. *Journal of Development Economics*, **104**, 233–244.

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

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

Carsten, Janet, *et al.* 2004. *After kinship*. Vol. 2. Cambridge University Press.

Carvalho, Jean-Paul, & Koyama, Mark. 2016. Jewish emancipation and schism: Economic development and religious change. *Journal of Comparative Economics*, **44**(3), 562–584.

Chaney, Eric. 2016. Religion and the rise and fall of Islamic science. *Work. Pap., Dep. Econ., Harvard Univ., Cambridge, MA*.

Chaney, Eric. 2020. Modern Library Holdings and Historic City Growth.

Clark, Andrew, & Lelkes, Orsolya. 2005. Deliver us from evil: religion as insurance. *Papers on Economics of Religion*, **603**, 1–36.

Clark, Gregory. 2008. *A farewell to alms: a brief economic history of the world*. Vol. 25. Princeton University Press.

De Chaisemartin, Clément, & d'Haultfoeuille, Xavier. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, **110**(9), 2964–96.

De la Croix, David, Docquier, Frédéric, Fabre, Alice, & Stelter, Robert. 2020a. The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800).

De la Croix, David, Morault, Pauline, *et al.* 2020b. *Winners and Losers from the Protestant Reformation: An Analysis of the Network of European Universities*. Tech. rept. Institut de recherche économiques et sociales, UC Louvain.

Diamond, Jared. 1997. *Guns, Germs, and Steel: The Fates of Human Societies*. WW Norton and Company.

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

Dues, Greg. 1992. *Catholic customs & traditions: A popular guide*. Twenty-Third Publications.

Edwards, Rosalind, & Caballero, Chamion. 2008. What's in a name? An exploration of the significance of personal naming of 'mixed'children for parents from different racial, ethnic and faith backgrounds. *The sociological review*, **56**(1), 39–60.

Ellison, Christopher G, Gay, David A, & Glass, Thomas A. 1989. Does religious commitment contribute to individual life satisfaction? *Social Forces*, **68**(1), 100–123.

EVS. 2022. *EVS Trend File 1981-2017*. GESIS, Cologne. ZA7503 Data file Version 3.0.0, https://doi.org/10.4232/1.14021.

Fouka, Vasiliki. 2020. Backlash: The unintended effects of language prohibition in US schools after World War I. *The Review of Economic Studies*, **87**(1), 204–239.

Friedman, Benjamin M. 2021. *Religion and the Rise of Capitalism*. Knopf.

Fryer, Roland G Jr, & Levitt, Steven D. 2004. The causes and consequences of distinctively black names. *The Quarterly Journal of Economics*, **119**(3), 767–805.

Galor, Oded. 2011. *Unified growth theory*. Princeton University Press.

Gibson, Henry M. 1989. Measuring attitudes towards Christianity among 11-to 16-year-old pupils in non-denominational schools in Scotland. *Educational Research*, **31**(3), 221–227.

Giuliano, Paola, & Nunn, Nathan. 2021. Understanding cultural persistence and change. *The Review of Economic Studies*, **88**(4), 1541–1581.

Gorodnichenko, Yuriy, & Roland, Gerard. 2011. Individualism, innovation, and long-run growth. *Proceedings of the National Academy of Sciences*, **108**(Supplement 4), 21316–21319.

Gruber, Jonathan, & Hungerman, Daniel M. 2008. The church versus the mall: What happens when religion faces increased secular competition? *The Quarterly journal of economics*, **123**(2), 831–862.

Grzymala-Busse, Anna. 2020. Beyond war and contracts: The medieval and religious roots of the European state. *Annual Review of Political Science*, **23**, 19–36.

Guiso, Luigi, Sapienza, Paola, & Zingales, Luigi. 2003. People's opium? Religion and economic attitudes. *Journal of monetary economics*, **50**(1), 225–282.

Hacker, J David. 1999. Child naming, religion, and the decline of marital fertility in nineteenth-century America. *The History of the Family*, **4**(3), 339–365.

Henrich, Joseph, Ensminger, Jean, McElreath, Richard, Barr, Abigail, Barrett, Clark, Bolyanatz, Alexander, Cardenas, Juan Camilo, Gurven, Michael, Gwako, Edwins, Henrich, Natalie, *et al.* 2010. Markets, religion, community size, and the evolution of fairness and punishment. *science*, **327**(5972), 1480–1484.

Hill, Peter C, & Pargament, Kenneth I. 2003. Advances in the conceptualization and measurement of religion and spirituality: Implications for physical and mental health research. *American psychologist*, **58**(1), 64.

Hodge, David R, & Holtrop, CR. 2002. Spiritual assessment: A review of complementary assessment models. *Social work and Christianity: Readings on the integration of Christian faith and social work practice*, **2**, 167–192.

Hoffman, Philip T. 2015. Why did Europe conquer the world? *In: Why Did Europe Conquer the World?* Princeton University Press.

Hofstede, Geert, Hofstede, Gert Jan, & Minkov, Michael. 2005. *Cultures and organizations: Software of the mind*. Vol. 2. Mcgraw-hill New York.

Hornbeck, Richard, & Naidu, Suresh. 2014. When the levee breaks: black migration and economic development in the American South. *American Economic Review*, **104**(3), 963–90.

Iannaccone, Laurence R. 1998. Introduction to the Economics of Religion. *Journal of Economic Literature*, **36**(3), 1465–1495.

Iyer, Sriya. 2016. The new economics of religion. *Journal of Economic Literature*, **54**(2), 395–441.

Jacob, Margaret C. 1997. *Scientific culture and the making of the industrial West*. Oxford University Press.

Jonsson, Jan O, Grusky, David B, Di Carlo, Matthew, Pollak, Reinhard, & Brinton, Mary C. 2009. Microclass mobility: Social reproduction in four countries. *American Journal of Sociology*, **114**(4), 977–1036.

Jurajda, Štěpán, & Kovač, Dejan. 2021. Names and behavior in a war. *Journal of Population Economics*, **34**(1), 1–33.

Knudsen, Anne Sofie Beck. 2019. *Those Who Stayed: Selection and Cultural Change during the Age of Mass Migration*. Tech. rept. Working Paper.

Kuran, Timur. 2018. Islam and economic performance: Historical and contemporary links. *Journal of Economic Literature*, **56**(4), 1292–1359.

Landes, David S. 1998. The Wealth and Poverty of Nations: Why Some Are So Rich and Some So Poor.

Le Goff, Jacques. 1986. *The birth of purgatory*. University of Chicago Press.

Le Goff, Jacques. 1993. *Intellectuals in the middle ages*. Blackwell Oxford.

Lehrer, Evelyn L. 2004. Religion as a determinant of economic and demographic behavior in the United States. *Population and development review*, **30**(4), 707–726.

Lehrer, Evelyn L. 2008. The role of religion in economic and demographic behavior in the United States: A review of the recent literature. *Available at SSRN 1145920*.

Lieberson, Stanley. 2000. *A matter of taste: How names, fashions, and culture change*. Yale University Press.

Litina, Anastasia, & Fernández, Èric Roca. 2020. Celestial enlightenment: eclipses, curiosity and economic development among pre-modern ethnic groups.

Livny, Avital. 2021. Can Religiosity be Sensed with Satellite Data? An Assessment of Luminosity during Ramadan in Turkey. *Public Opinion Quarterly*, **85**(S1), 371–398.

Main, Gloria L. 1996. Naming children in early New England. *The Journal of interdisciplinary history*, **27**(1), 1–27.

Mallet, Robert, & Mallet, John William. 1858. *The Earthquake Catalogue of the British Association: With the Discussion, Curves and Maps, Etc.* Taylor & Francis.

Mateos, Pablo. 2014. Names, ethnicity and Populations. *Advances in Spatial Science*.

Miao, Qing, & Popp, David. 2014. Necessity as the mother of invention: Innovative responses to natural disasters. *Journal of Environmental Economics and Management*, **68**(2), 280–295.

Miller, Lisa, Bansal, Ravi, Wickramaratne, Priya, Hao, Xuejun, Tenke, Craig E, Weissman, Myrna M, & Peterson, Bradley S. 2014. Neuroanatomical correlates of religiosity and spirituality: A study in adults at high and low familial risk for depression. *JAMA psychiatry*, **71**(2), 128–135.

Miller, Lisa, Balodis, Iris M, McClintock, Clayton H, Xu, Jiansong, Lacadie, Cheryl M, Sinha, Rajita, & Potenza, Marc N. 2019. Neural correlates of personalized spiritual experiences. *Cerebral Cortex*, **29**(6), 2331–2338.

Mokyr, Joel. 2010. *The Enlightened economy an economic history of Britain 1700-1850*. Yale University Press.

Mokyr, Joel. 2012. Culture, institutions, and modern growth. *Pages 151–191 of: Institutions, Property Rights, and Economic Growth: The Legacy of Douglass North*. Cambridge University Press.

Mokyr, Joel. 2016a. *A culture of growth*. Princeton University Press.

Mokyr, Joel. 2016b. A culture of growth. *In: A Culture of Growth*. Princeton University Press.

Mokyr, Joel, & Voth, Hans-Joachim. 2009. Understanding growth in early modern Europe. *Cambridge Economic History of Europe*, **1**, 7–42.

Moscona, Jacob. 2021. Environmental catastrophe and the direction of invention: Evidence from the american dust bowl. *Available at SSRN 3924408*.

Newberg, Andrew, & Waldman, Mark Robert. 2009. *How God changes your brain: Breakthrough findings from a leading neuroscientist.* Ballantine Books.

Nickell, Stephen. 1981. Biases in dynamic models with fixed effects. *Econometrica: Journal of the econometric society*, 1417–1426.

Norenzayan, Ara. 2013. *Big gods: How religion transformed cooperation and conflict.* Princeton University Press.

Norenzayan, Ara, & Hansen, Ian G. 2006. Belief in supernatural agents in the face of death. *Personality and social psychology bulletin*, **32**(2), 174–187.

Norris, Pippa, & Inglehart, Ronald. 2011. *Sacred and secular: Religion and politics worldwide.* Cambridge University Press.

North, Douglass C. 1991. Institutions. *Journal of economic perspectives*, **5**(1), 97–112.

Nørtoft, Mikkel, & Schroeder, Hannes. 2020. The Ancient Genome Atlas.

Nunn, Nathan, & Qian, Nancy. 2011. The potato's contribution to population and urbanization: evidence from a historical experiment. *The quarterly journal of economics*, **126**(2), 593–650.

Olsson, Ola, & Hibbs Jr, Douglas A. 2005. Biogeography and long-run economic development. *European Economic Review*, **49**(4), 909–938.

Oster, Emily. 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, **37**(2), 187–204.

Pargament, Kenneth I. 2001. *The psychology of religion and coping: Theory, research, practice.* Guilford press.

Park, C., Cohen, L.H., & Herb, L. 1990. Intrinsic religiousness and religious coping as life stress moderators for Catholics versus Protestants. *Journal of personality and social psychology*, **59**(3), 562.

Purzycki, Benjamin Grant, Apicella, Coren, Atkinson, Quentin D, Cohen, Emma, McNamara, Rita Anne, Willard, Aiyana K, Xygalatas, Dimitris, Norenzayan, Ara, & Henrich, Joseph. 2016. Moralistic gods, supernatural punishment and the expansion of human sociality. *Nature*, **530**(7590), 327–330.

Rubin, Jared. 2014. Printing and Protestants: an empirical test of the role of printing in the Reformation. *Review of Economics and Statistics*, **96**(2), 270–286.

Rubin, Jared. 2017. *Rulers, Religion, and Riches: Why the West got rich and the Middle East did not.* Cambridge University Press.

Saavedra, Martin. 2021. Kenji or kenneth? pearl harbor and japanese-american assimilation. *Journal of Economic Behavior & Organization*, **185**, 602–624.

Sargent, Steven D. 1990. Saints' Cults and Naming Patterns in Bavaria, 1400-1600. *The Catholic historical review*, **76**(4), 673–696.

Sewell, William H, Haller, Archibald O, & Ohlendorf, George W. 1970. The educational and early occupational status attainment process: Replication and revision. *American sociological review*, 1014–1027.

Sharma, Arvind, & Young, Katherine K. 1987. *Women in world religions.* State University of New York Press Albany.

Sibley, Chris G, & Bulbulia, Joseph. 2012. Faith after an earthquake: A longitudinal study of religion and perceived health before and after the 2011 Christchurch New Zealand earthquake. *PloS one*, **7**(12), e49648.

Smith, Daniel Scott. 1984. Child-naming practices as cultural and familial indicators. *Local Population Studies*, **32**, 17–27.

Smith, Daniel Scott. 1985. Child-naming practices, kinship ties, and change in family attitudes in Hingham, Massachusetts, 1641 to 1880. *Journal of Social History*, **18**(4), 541–566.

Smith, Daniel Scott. 1994. Continuity and discontinuity in Puritan naming: Massachusetts, 1771. *The William and Mary Quarterly*, **51**(1), 67–91.

Spilka, Bernard, Hood, Ralph W, Hunsberger, Bruce, & Gorsuch, Richard. 2003. *The psychology of religion: An empirical approach.* Guilford Press.

Spolaore, Enrico, & Wacziarg, Romain. 2013. How deep are the roots of economic development? *Journal of economic literature*, **51**(2), 325–69.

Squicciarini, Mara P. 2020. Devotion and Development: Religiosity, Education, and Economic Progress in Nineteenth-Century France. *American Economic Review*, **110**(11), 3454–91.

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

Stark, Rodney, & Finke, Roger. 2000. *Acts of faith: Explaining the human side of religion.* University of California Press.

Stark, Rodney, & Glock, Charles Y. 1970. *American piety: The nature of religious commitment.* Vol. 1. Univ of California Press.

Strauss, Gerald. 1975. Success and failure in the German Reformation. *Past & Present*, 30–63.

Tackett, Timothy. 1986. *Religion, revolution, and regional culture in eighteenth-century France: the ecclesiastical oath of 1791.* Princeton University Press.

Tebbenhoff, Edward H. 1985. Tacit rules and hidden family structures: naming practices and godparentage in Schenectady, New York 1680-1800. *Journal of Social History*, **18**(4), 567–585.

Ullmann, Walter. 1970. *The growth of papal government in the Middle Ages.* Vol. 35. Routledge.

Van Zanden, Jan Luiten. 2009. Chapter Two. Why The European Economy Expanded Rapidly In A Period Of Political Fragmentation. *Pages 32–68 of: The Long Road to the Industrial Revolution.* Brill.

vom Bruck, Gabriele, & Bodenhorn, Barbara. 2006. *Entangled Histories: An Introduction to the Anthropology of Names and Naming, in "An Anthropology of Names and Naming" (Eds: vom Bruck, Gabriele and Bodenhorn, Barbara).* Cambridge: Cambridge University Press.

Voth, Hans-Joachim. 2021. Persistence–myth and mystery. *Pages 243–267 of: The handbook of historical economics.* Elsevier.

Weber, M. 1905. The Protestant Ethic and the Spirit of Capitalism. *trans. Talcott Parsons 1930.*

Yalçıner, Ahmet Cevdet. 2009. Tsunami risk and strategies for the European region.

Zelinsky, Wilbur. 1970. Cultural variation in personal name patterns in the Eastern United States. *Annals of the Association of American Geographers*, **60**(4), 743–769.

Appendix

(Potentially for online publication)

A Construction of the author database

This appendix describes the data on authors used in the main analysis and all the different datasets that were linked to calculate the variables. The data on medieval university students is described in Appendix I.

A.1 VIAF identifier

The main database of authors consists of authors with a Virtual International Authority File (VIAF) identifier from either the VIAF database, the Deutsche Nationalbibliothek (DNB, the German National Library), or Wikidata, all of which we scraped in April 2020. We commence by identifying the geocodes of the birthplaces of the authors. This information is available from the DNB and Wikidata databases.[82] Of the 4.1 million authors with a VIAF identifier, we were able to identify the location and date of birth for 798,930 of them in the DNB database and for 462,405 in Wikidata. When information from several sources is available, we prefer data from DNB which seems the most reliable source of information. Next comes Wikidata and last the VIAF database itself. We restricted the sample to authors born in Europe, before 1940 to ensure that the sample includes all authors that could have potentially been born within the sample window. We further restrict the sample to authors born after 1300, at which point in time it is reasonable to assume that Christian names were widespread in Germanic Europe (Sargent, 1990). In this reduced dataset we have 487,156 authors.

A.2 Treatment of names

We identify the first name of the authors from the DNB and Wikipedia databases; "forename" in DNB and "given name" in Wikipedia. Because some of the people in the database have multiple first names listed, our main measure of first names is based on the first name listed in each of the databases, under the assumption that this is their primary first name in most cases. We again use DNB as our primary source. If DNB has no information on the first name, we use the Wikipedia database. We standardize the names to the Latin alphabet using the R-function "stringi".

We identify the etymological origin of names – what we term name branches – using the website www.behindthename.com, which contains 23,938 names certified by the website. In addition, the website includes a number of user submitted names, which are not certified by the website and which we do not use. The database contains information on

[82]These two sources are the two main geo-referencing sources used by Chaney (2020) and the only two where exact geocodes for birthplace are available. We get a rather detailed representation of Europe by using these two geo-referencing sources (Figure 1).

the etymology of all names included which allows us to identify the name branch of each name. The 23,938 names can be reduced to a total of 7,606 name branches.

To understand what a name branch is and how it might work, consider the following. The name Jeanet is the Danish version of the French name Jeannette, which is itself a diminutive of the name Jeanne. Jeanne is the modern French form of the old French name Jehanne, which in turn was the feminine form of the name Iohannes. Iohannes has many versions and is by far the most common name in all our databases on personal names. Iohannes also comes in the forms John, Jon, Johan, Hans, Jean, Johnny etc. Iohannes itself is the latin version of the greek Ioannes, ultimately deriving from the Hebrew name Yôhānān, which means "Yahweh is gracious" which was the name of John the Baptist and later John the Apostle. This means that Jeanet, Jeannette, Jeanne, Iohannes, John, Jon, Johan, Hans, Jean, Johnny, and Yochanan are all in the same name branch, which we name after its modern English form "John".

To the etymological connections already coded in the behindthename dataset we make a few additions. Firstly, we combine name branches deriving from the word "Christ" with those deriving from the name branch "Jesus" since they both refer to the same person. Secondly, we merge name branches that etymologically refer to God with the behindthename name branch "Yahweh". This means that names that include the strings "gott", "theo", or "teo" are merged with "Yahweh" as well as the names "Godfrey", "Godabert", "Godehard", "Godica", "Godric", "Godwine", "Goteleib", "Gottschalk". Thirdly, we keep the name branches "Matthew" and "John" separate from the name branch "Yahweh" (even though the behindthenames database defines them as part of the "Yahweh" name branch) as these names refer to the apostles rather than directly to the Christian God. The first two amendments mean that some names that would otherwise not have been defined as religious will be defined as such. For instance, Godfrey, which is an obvious reference to the Christian God.

Table A1 shows an example of how first names are attached to their name branches. The sample includes 100 authors born within the same 1x1 km grid-cell in Germany.

Table A1: Example of the link between first names and name branches

Firstname	Lastname	Namebranch	Birthyear
adolf	klaus	adolf	1914
albert	koch	albert	1841
andreas	sievers	andrew	1931
anton	praetorius	anthony	1560
anton	freilinghausen	anthony	1597
anton	moller	anthony	1762
arnold	meshov	arnold	1591
arnold	monnich	arnold	1703
august	houckius	augustus	1677
bernhard	brinkmann	bernard	1627
bernhard	lippsmeier	bernard	1885
bernhard	eichholtz	bernard	1900
bernhard	lippsmeier	bernard	1914
carl	wibbe	charles	1896
christian	meien	jesus	1781
conrad	hansen	conrad	1906
eduard	vogt	edward	1812
eduard	gaffron	edward	1861
eduard	overhoff	edward	1905
emilie	schniewind	emil	1892
erika	stork	eric	1939
ernst	lahme	ernest	1855
ernst	hohenhaus	ernest	1879
ernst	giesecke	ernest	1900
eva	dahlkotter	eve	1924
ewald	kleine	ewald	1873
florian	meyer-langenfeld	florus	1936
franz	liebrecht	francis	1882
franz	bussen	francis	1906
franz	huneke	francis	1929
friedrich	houck	frederick	1689
friedrich	duncker	frederick	1797
friedrich	rose	frederick	1839
friedrich	ferrari	frederick	1884
friedrich	hoings	frederick	1908
fritz	saint paul	frederick	1843
fritz	eichholtz	frederick	1889
fritz	kassmann	frederick	1908
gerard	bodin	gerard	1620
gustav	klumpp	gustav	1888
hagen	galatis	hagen	1928
hanna	dahlkotter	john	1899
hans	lummerzheim	john	1919
heinrich	vockeradt	henry	1844
heinrich	trembur	henry	1871
heinrich	pagenkemper	henry	1878
heinrich	everz	henry	1882
heinrich	oestreich	henry	1904
heinrich	palmer	henry	1906
hermann	rothert	herman	1875
hermann	vornefeld	herman	1924
hildegard	kaup	hildegard	1924
horst	kayling	horst	1940
hugo	sauer	hugh	1885
jakob	koenen	jacob	1907
johann	kellerhaus	john	1632
johann	schwiggen	john	1641
johann	fabricius	john	1646
johann	schwackenberg	john	1671
johann	nottebohm	john	1671
johann	zwicke	john	1721
johann	drekmann	john	1724
johann	castorff	john	1738
johann	stuve	john	1752
johann	nonne	john	1785
johanna	post	john	1856
johanna	schrader	john	1888
johannes	cincinnius	john	1485
johannes	kayser	john	1654
johannes	devivere	john	1798
johannes	effelsberger	john	1889
josef	link	joseph	1913
julius	post	julius	1844
julius	petri	julius	1868
karl	boddeker	charles	1846
karl	rickelt	charles	1857
kaspar	ulenberg	jasper	1548
lambert	vent	lambert	1661
louis	buddeberg	ludwig	1836
magda	selter	magdalene	1878
margret	dietrich	margaret	1920
maria	brandt	mary	1878
marie	steinbecker	mary	1879
martin	niemoller	martin	1892
max	grunebaum	maximus	1851
oskar	fuchs	oscar	1866
paul	mehlitz	paul	1912
pauline	wurttemberg	paul	1854
peter	rosskampff	peter	1647
simon	gakel	simon	1616
sophie	modersohn	sophia	1799
stephan	klotz	stephen	1606
theodor	grewe	yahweh	1916
theresia	thranberend	theresa	1940
ursula	schlieffen	ursula	1930
walter	nies	walter	1918
wilhelm	gallenkamp	william	1820
wilhelm	wetekamp	william	1859
wilhelm	niemoller	william	1898
wolfgang	jansen	wolfgang	1924

One hundred authors born within the same 1x1 km grid cell in Germany. The grid-cell was chosen as one out of four grid cells with exactly 100 authors. The last name is not used for calculating the name branch, but is depicted here for completeness.

A.3 Worldcat database

To obtain information on authors' popularity and productivity we link the authors to their data from the Worldcat database in which we are able to trace 186,199 of the authors.

A.4 Variable descriptions

Birth Year

The year of birth of each author. We use data on birth year from the DNB-database if this is available (414,393 cases). For those where DNB does not have data we use Wikidata (72,480 cases) and for those that have neither we rely on the VIAF database itself (283 cases). For the 150,318 authors for whom we have birth years from both the DNB and Wikidata datasets, the birth years correlate with a ρ=0.9997.

Common name

In the baseline specification, we measure name commonness as a dummy equal to one if the first name is among the top-10 most common names within 10 years prior to birth among persons born in the same country, following Knudsen (2019) and Bazzi *et al.* (2020). For author i born in country c, we first calculate the share of authors born in country c within 10 years prior to the author's birth year, who share first name with author i. We next calculate a dummy equal to one if the first name is among the top-10 most frequent names in the country at the given point in time. We base the main measure on actual first names and not the aggregated name branches. We do this following Knudsen (2019). We depart partly from Knudsen in that Knudsen groups the name further into gender. For the main results, we refrain from this as the share of women is rather low. The results are unchanged if we also group the names in terms of gender (Tables B7 and G7). The results are also robust to other alternative measures of name commonness, based on name branches instead of first names or calculated throughout Europe instead of within countries.

Death Year

The year of death of each author. We use the death year from the DNB database when available. Otherwise, we use the death year from Wikipedia. This information is available for 374,992 authors. For the 131,860 authors for whom we have death years from both the DNB and Wikidata datasets, the death years correlate with a ρ=0.9658.

Ethnicity

Ethnicity measures the commonness of a name among persons of the particular ethnicity, compared to the rest of the population. We first determine the languages spoken at the place of birth of each author. We do so using shapefiles for the main languages spoken across time in Europe, available through the novel Ancient Genome Atlas constructed by Nørtoft & Schroeder (2020). The main language groups include Greek, Celtic, Italic, Baltic, Slavic, Germanic, Semitic, Uralic, Turkic, and Basque. We focus on the language groups before 1600, as late as possible to get detailed enough data, but early enough to be before modern times. We use these language areas to compute measures akin to the RNI for the ethnicity signaled by the name of each author. In particular, we measure how normal a name is among a specific ethnicity compared to the rest of the sample, following Andersen (2021). We use the following formula:

$$ENI_j^{ethnicity} = \frac{Pr(name_j|Population_{ethnicity})}{Pr(name_j|Population_{ethnicity} + Pr(name_j|Population_{not-ethnicity})},$$

where j is a particular name branch, ethnicity is a particular linguistic group, $Pr(name_j|Population_{ethnic}$ is the probability of having name j in the ethnic group, and $Pr(name_i|Population_{not-ethnicity})$ is the probability of having the name if one is not in the ethnic group. This approach is identical to the approach used in the original paper by Fryer & Levitt (2004), except that they have access to the self-reported race of each person in their sample, whereas we do not. This means that in order construct the ENI values we need some indicator of ethnicity from elsewhere; the language groups. Thus, for each name branch, we know how frequent the name is within each of the language groups.

Female

Dummy variable equal to one if the first name is defined as a female name in the behindthename database, zero otherwise.

Latitude and longitude

Latitude or longitude of the author's birthplace. We use data from the DNB database for the most cases (327,608 cases) and for the rest we use information from Wikidata (159,548 cases). For some authors, we have latitudes and longitudes from both sources. For the 56,505 authors for whom we have latitude (longitude) from both the DNB and Wikidata datasets, the latitudes (longitudes) correlate with a ρ=0.9354 (0.9058).

Number first names

We calculate the number of first names from the variable forename in DNB and the given name from Wikipedia. Based on the raw first names, the average (median) author in our data has 1.24 (1) first names. 81% of the sample has one first name, 16% has two, 2% has three names, and 2% has more than three names. The person with the largest number of first names in our data is Louise Élisabeth Félicité Françoise Armande Anne Marie Jeanne Joséphine de Croÿ de Tourzel, a French noblewoman (11 June 1749 – 15 May

1832). We use the number of first names to test whether the results could be caused by name changing in Table D1. Note that not the entire part of the above name is useful for name changing. We have therefore checked robustness of results using another measure of first names that comprises only the part of names before "de", "d'", "von", etc. As the share of authors with such a name is very low (0.3%), it is not surprising that results are unchanged (results available upon request).

Number first names dummy
A dummy variable equal to one if the person has more than one first name.

Occupation
A dummy variable equal to one if the particular occupation is mentioned in the list of occupations, available from the DNB database for a total of 362,666 authors. Each of these authors have between 1 and 24 occupations listed, but 99% of the authors with at least one occupation have five occupations or less. The average (median) number of occupations is 1.9 (2). We remove all special characters from the occupations. The occupations are stated in German, where it is customary that a occupation held by a female has an "in" added in the end of the male occupation. For instance, the German word for a painter is "malerin" if the painter is female and "maler" if the painter is male. We remove the "in" in the end of all occupations that have a male counterpart. In addition, sometimes the male version of the occupation has an "e" in the end. We remove these e's, so that male and female occupations are called the same. Last, we translate the occupations into English to produce Figure 3.

Religious occupation: A dummy variable equal to one if the author in question had a (Christian) religious occupation, zero otherwise. We define a religious occupation as all occupations categorized as "Christian Religious Occupations" in Wikipedia and the religious occupations listed in the European Roots Genealogy list of old German occupations in `www.european-roots.com/german_prof.htm`. In particular, we define an occupation as religious if it contains one of the following word parts: abt, abtiss, apostel, august, archimandrit, benediktin, bischof, christ, diakon, domherr, eremit, evangel, franziskan, geistlich, jesuit, johannit, inquisitor, kanon, kapitular, kaplan, kardinal, karmelit, katholi, kleriker, kirch, laienbruder, martyr, missionar, monch, mystiker, nonne, ordensbruder, ordensschwester, papst, pastor, pater pfarrer, pilgervater, pradikant, pralat, prediger, priest, prior, probst, propst, protestant, religionslehrer, religionspedagog, religionspadagog, religios, scholast, seelsorg, theol, theosoph, and vikar. To not obtain false-positives, we only include occupations based on the two short word parts, abt and pater, if they are followed by or surrounded by space, i.e. "abt " and " pater ". This is important, as we will otherwise capture words, such as abteilungsleider (department leader) or spater (later).

Table A2 shows the top-10 most frequent occupations among authors with at least one religious occupation in the sample in the first columns and among authors who do not have a religious occupation in the last columns. The total number of observations exceeds the number of authors in the sample, as several authors have more than one occupation. This also means that some occupations can enter both lists, which is the case with university teacher.

Table A2: Top-10 religious and non-religious occupations

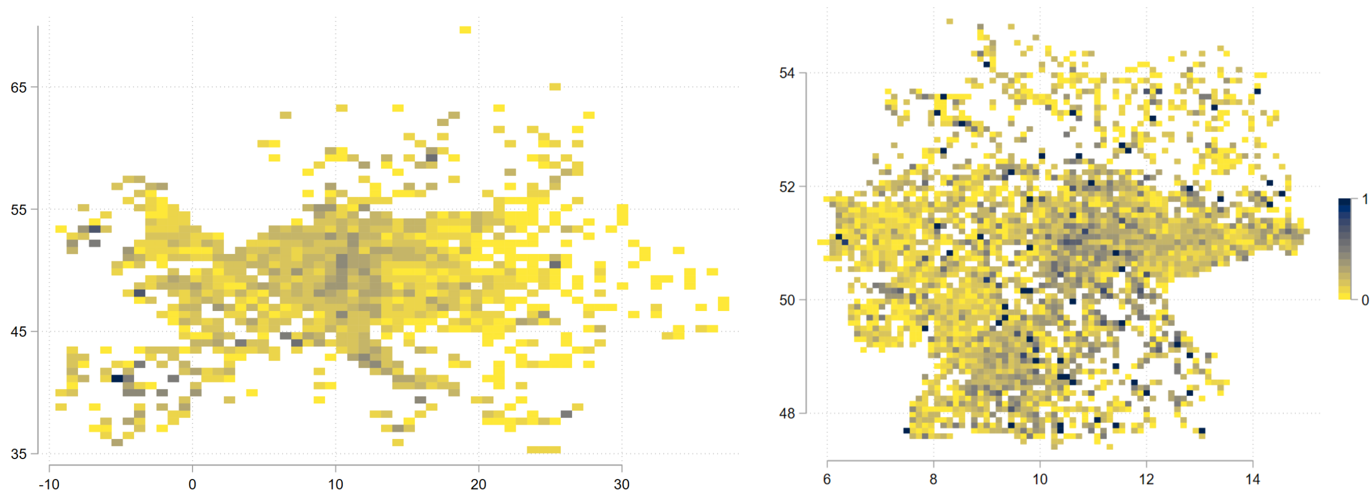
Religious			Non-religious		
Profession	Number	Share	Profession	Number	Share
Pastor	17,148	0.18	Writer	33,250	0.05
Theologian	16,399	0.17	Painter	24,822	0.04
Catholic	4,170	0.04	Lawyer	24,248	0.04
Protestant	3,911	0.04	Doctor	24,080	0.04
Priest	2,756	0.03	Politician	16,664	0.03
Writer	2,589	0.03	Artist	15,326	0.02
University teacher	1,932	0.02	Composer	14,167	0.02
Bishop	1,866	0.02	University teacher	11,720	0.02
Principal	1,622	0.02	Historian	11,557	0.02
Teacher	1,543	0.02	Actor	10,895	0.02
Total	93,970	1.00	Total	616,259	1.00

Top-10 most frequent occupations among authors with at least one occupation that we define as religious in the first columns and among those who do not have a religious occupation among any of their occupations in the last columns. The numbers reflect the number of times the particular occupation is listed. The shares reflect the number of times a particular occupation is listed as a share of the total number of author-occupations. As each author may have more than one occupation, the total numbers exceed the number of authors.

Figure A1: Share of religious occupations

(a) Full sample

(b) Example: Germany



Share of authors in a grid-cell with at least one religious occupation for the full sample in panel a and a zoom in on the country with most authors, Germany, in panel b to help visualize the detail of the data. A dark blue grid indicates that all authors born in this cell had at least one religious occupation, while no authors in clear yellow grids have a religious occupation. Any tones of blue or yellow between these extremes indicate that a mix of authors with and without religious occupations were born here. The sample is restricted to grids with at least 10 authors.

Natural scientific occupation: A dummy variable equal to one if the author had an occupation belonging to natural science, zero otherwise. Our definition of a natural scientific occupation includes all the jobs listed in the "List of scientific occupations" in Wikipedia. This means that if any of the following strings appear in the "occupations" variable in the DNB dataset we classify the occupation as scientific: anatom, archaolog, arzt, astronaut, astronom, biochemiker, biolog, botanik, chemiker, chirurg, demograph, demograf, erdolgeolog, forster, geoberuf, geograph, geograf, geolog, gynakolog, herpetolog, informati, kadiolog, mathematik, medizin, naturforsch, naturwissenschaft, neuro, onkolog, ornitholog, ozeanograph, ozeanograf, palaontolog, pathologe, pharma, physik, physiolog, polarforsch, psychia, statistiker, zoolog, and okologe.

Engineering occupation: A dummy variable equal to one if the author had an occupation relating to engineering, zero otherwise. We define an occupation relating to engineering as occupations including the strings ingenieur, techn, erfinder.

Science or engineering occupation: Our main measure of scientific or engineering occupations is a dummy variable equal to one if the occupation is either related to natural science or engineering, zero otherwise. Table A3 shows the top-10 occupations belonging to these categories and the top-10 occupations that are not either scientific or related to engineering.

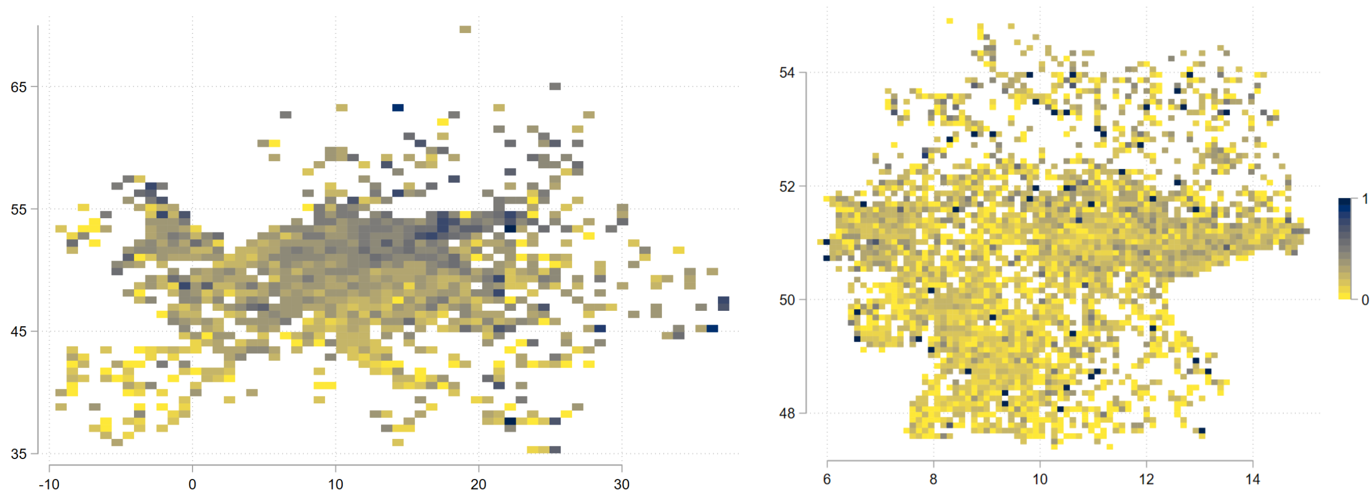
Table A3: Top-10 occupations among authors with and without one science occupation

Science			Non-science		
Profession	Number	Share	Profession	Number	Share
Doctor	24,239	0.21	Writer	34,047	0.06
Chemist	5,167	0.04	Painter	24,651	0.04
Engineer	4,906	0.04	Lawyer	24,478	0.04
University teacher	4,663	0.04	Pastor	16,955	0.03
Mathematician	4,028	0.04	Politician	16,118	0.03
Physicist	3,639	0.03	Theologian	15,910	0.03
Botanist	2,965	0.03	Artist	15,228	0.03
Archaeologist	2,747	0.02	Composer	14,703	0.02
Geologist	2,211	0.02	Historian	11,772	0.02
Veterinarian	2,140	0.02	Actor	10,861	0.02
Total	114,923	1.00	Total	595,406	1.00

Top-10 most frequent occupations among authors with at least one occupation that we define as scientific or engineering in the first columns and among those who do not have a scientific or engineering occupation among any of their occupations in the last columns. The numbers reflect the number of times the particular occupation is listed. The shares reflect the number of times a particular occupation is listed as a share of the total number of author-occupations. As each author may have more than one occupation, the total numbers exceed the number of authors.

Figure A2: Share of scientific occupations

(a) Full sample (b) Example: Germany



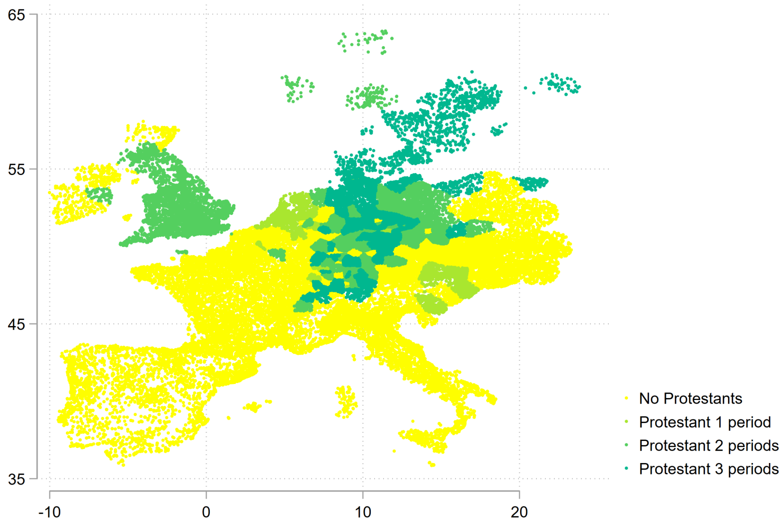
Share of authors in a grid-cell with at least one scientific or engineering occupation for the full sample in panel a and a zoom in on the country with most authors, Germany, in panel b to help visualize the detail of the data. A dark blue grid-cell indicates that all authors born in this cell had at least one scientific or engineering occupation, while no authors in clear yellow grids have a scientific or engineering occupation. Any tones of blue or yellow between these extremes indicate that a mix of authors with and without scientific or engineering occupations were born here. The sample is restricted to grids with at least 10 authors.

Social science occupation: A dummy variable equal to one if the author had an occupation relating to social science, zero otherwise. We define an occupation relating to social science as all occupations categorized as ”social science” in the ”List of scientific occupations” in Wikipedia. This means that if any of the following strings appear in the “occupations” variable in the DNB dataset we classify the job as belonging to social science: anthropolog, ethnolog, historiker, politikwissenschaft, staatswissenschaft, soziolog, stadtplaner, okonom, volkswirt, and wirtschaftswissenschaft.

Protestant area

We gather information on whether an area was Protestant from Rubin (2014). Rubin provides city-level information on whether or not a city was Protestant in years 1530, 1560, or 1600. For authors born within 100 km of one of these cities, we attach them to the information for the closest city. Figure A3 shows the spread of the birthplaces based on whether the closest city was Protestant or not and for how long.

Figure A3: Birthplaces of authors by Protestant areas



The map shows the birthplaces of authors, where colors indicate whether the author was born within 100 km of a city that was Protestant in 1530, 1560, and/or 1600.

Religious name

In individual-level analyses, this variable is a dummy variable equal to one if the person in question has a religious name and zero otherwise. We define religious names as those that are shared with i) a the patron saint of a church in the Buringh *et al.* (2020) dataset. Table 1 in the main text shows the top-10 names among churches. Around one out of four churches were dedicated to a central Catholic figure, the Virgin Mary. After Mary comes Peter, arguably the most prominent of the apostles of Jesus and according to tradition the first leader of the Catholic church. Next comes John, referring to John the Apostle or John the Baptist. The name branch Martin rose to fame due to Martin of Tours, one of the most popular saints in France and patron saint of many communities and organizations across Europe. Thereafter comes Paul, referring to Paul the Apostle, also known as Saint Paul, sometimes referred to as "the second founder of Christianity" for his role in forming the doctrines of the Church and disseminating it throughout the Roman world. Then comes Nicholas, referring to St Nicholas, a central saint who later became part of the basis for Santa Claus. ii) Jesus and his relatives and apostles: Andrew, Jacob, Judah, Peter, Philip, Simon, Talmai, Thomas, Barnabas, Paul, John, Cleopatra, Jesus, Joseph, Mary, Simon,[83] iii) name branches associated with the Christian God, such as Yahweh, Godfrey, and Amadeus, or iv) the name branch "Cristos" which comprises a clear reference to Jesus, even though "Cristos" is not his name, but his title. A religious

[83]The apostles of Jesus (and their name branch in squared brackets) were Simon whom Jesus named Peter [Peter]; his brother Andrew [Andrew]; James [Jacob]; John [John]; Philip [Philip]; Bartholomew [Talmai]; Matthew [Matthew]; Thomas [Thomas]; James, son of Alphaeus [Jacob]; Simon who was called the Zealot [Simon], Judas son of James [Judah], and Judas Iscariot [Judah]. Regarding the origin of "Bartholomew" in the name branch "Talmai", Bartholomew is the English form of the an Aramaic name meaning "Son of Talmai". "Talmai" itself comes from the Hebrew word for "furrowed". The Relatives of Jesus include Jesus [Jesus]; his mother Mary [Mary]; his (legal) father Joseph [Joseph]; his brothers James [Jacob], Joseph [Joseph], Judas [Judah], and Simon [Simon]; as well as his Uncle Cleopas [Cleopatra] and his aunt Mary (of Cleopas) [Mary].

name is thus defined as a name that is either shared by one of the patron saints or one of the central figures in the Bible.

Royal name

An indicator equal to one if the author shared name with any monarchs who reigned the country at any time before the person was born. Data on the names of European monarchs is retrieved from the Genealogy.eu database (`http://w.genealogy.euweb.cz/`). The country is defined based on contemporary borders. Links between current and former country borders were retrieved from the same Genealogy.eu database. In addition, we link the United Kingdom to England and the Czech Republic to the Kingdom of Bohemia. The variable is missing if the country was never a monarchy. 11% of the sample has a name shared by a monarch from their country.

RNI indices

The Religious Name Index, RNI, measures how frequent the name is among the particular religious entity, compared to the population as a whole. We compute these measures using a formula by Fryer & Levitt (2004) who used it to measure the relative "blackness" of a name. The formulas and religious entities are described below.

<u>Saint RNI</u>: Measures how important the name is as a saint name relative to its importance among the full population, using the formula:

$$SaintsRNI_j = \frac{Pr(name_j|Cathedrals)}{Pr(name_j|Cathedrals) + Pr(name_j|Population)}, \tag{4}$$

where $Pr(name_j|Cathedrals)$ is the percentage of churches that were dedicated to a saint with the name branch j. This proxies the importance of the name among Catholics. $Pr(name_j|Population)$ is the percentage of all persons in the sample who have the name. This measure accounts for the concern that some names are simply more common among saints because they are more common in general. The most common names of Saints in this sample are Mary, Peter, John, Martin, Paul, Nicholas, Stephen, Jacob, Michael, and Francis. The top-10 name branches ranked based on the Saint RNI are Afra, Mungo, Elmo, Nazarius, Secundus, Juvenal, Widogast, Gervasius, Domitius, and Brice. These names are held by very few people in our sample and we would not want to construct variables based on this top-10.

<u>Bible RNI</u>: Measures how important a name is as a Biblical name, relative to its general popularity. It does so by comparing how often a name is used in the Bible, relative to its general popularity in the sample, using the following formula:

$$BibleRNI_j = \frac{Pr(name_j|Bible)}{Pr(name_j|Bible) + Pr(name_j|Population)}, \tag{5}$$

where $Pr(name_j|Bible)$ is the number of times name branch j occurs in the Bible divided by the total number of times any name was used in the Bible. To calculate, we use the standard King James version of the bible (`https://www.sacred-texts.com/bib/osrc/index.htm`). All cases where "lord" is used to refer to the Christian God, we count it as an instance of "Yahweh" since it refers to the same person and since this is the term used in the original Hebrew Bible. $Pr(name_j|Population)$ is the percentage of people in the sample who have the name. The variable correlates with the Saint RNI with a ρ of 0.2.

Old testament RNI: The Bible RNI for the Old Testament. The most frequent name branches in the Old Testament are Yahweh (or names referring to God, such as Theodor and Gottfried), David, Seth, Moses, Saul, Judah, Aaron, Abram (including mainly Abraham), Benjamin, and Jacob. The top-10 name branches ranked based on the Old Testament RNI are Seth, Jehoshaphat, Rama, Manasseh, Abiathar, Saul, Jehoram, Micaiah, Uzziah, and Hiram. These names are held by very few people in our sample.

New testament RNI: The Bible RNI for the New Testament. The most frequent name branches in the New Testament are Jesus, Yahweh (or names referring to God, such as Theodor and Gottfried), Peter, Simon, Paul, John, Seth, David, Elijah, and Abram. Instead, the top-10 name branches ranked based on the New Testament RNI are Seth, Herod, Festus, Hananiah, Hosanna, Trophimus, Hagar, Philemon, Abiathar, and Talitha. These names are held by very few people in our sample. The New Testament RNI correlates with the Old Testament RNI with a ρ of 0.6.

Religious occupation RNI: Measures how prominent a name is among authors with religious occupations, relative to authors without such occupations (as defined in Section A.4). We use the following formula:

$$RNI_j^{relprof} = \frac{Pr(name_j|Population_{relprof})}{Pr(name_j|Population_{relprof}) + Pr(name_j|Population_{no-relprof})}, \tag{6}$$

where $Pr(name_j|Population_{relprof}$ is the percentage of authors with a name belonging to name branch j among authors with a religious occupation. $Pr(name_j|Population_{no-relprof})$ is the percentage of authors named j among authors without a religious occupation. The most frequent name branches among individuals with religious occupations are John, Charles, Henry, Joseph, Francis, George, Frederick, William, Ludwig, and Peter.

SNI indices

Instead of attempting to measure the actual social status of parents, we develop Social Status Name Indices, which we term SNI. The rationale is that if parental socioeconomic status biases our results, we should be able to remove the bias by controlling for the part of a name that reflects parents' socioeconomic status. Note that if socioeconomic status is not reflected in a name, then the results were not biased to start. To capture the part of a name that reflects the social status of parents, we compute measures similar to our RNI measures.

Advanced degree NI: The Advanced Degree Name Index measures the relative frequency of a name among individuals with advanced university degrees. We calculate using the following equation that is similar to Equations (4) and (5) where we computed RNIs based on frequencies of the names among churches and biblical figures:

$$SNI_j^{AdvDeg} = \frac{Pr(name_j|AdvDeg)}{Pr(name_j|AdvDeg) + Pr(name_j|Population)}, \tag{7}$$

where $Pr(name_j|AdvDeg)$ is the percentage of university students in the RAG database with a name belonging to name branch j who completed an advanced degree. $Pr(name_j|Population)$ is the percentage of authors with the name. For these calculations, we restrict the sample to authors with a name that also occurred among the advanced university students. The sample shrinks by 10%.

EducNI: The High Education Name Index measures the relative frequency of a name among individuals born in areas with otherwise high education. To calculate, we use a formula similar to Equation (6) where we calculated RNI based on frequencies of names among persons with religious professions:

$$SNI_{jc}^{educ} = \frac{Pr(name_{jc}|Population_{above-educ})}{Pr(name_{jc}|Population_{above-educ}) + Pr(name_{jc}|Population_{below-educ})}, \tag{8}$$

where $Pr(name_{jc}|Population_{above-educ})$ is the percentage of authors with name branch j who live in areas of country c with above-median education levels. $Pr(name_{jc}|Population_{below-educ})$ is the percentage of authors with name branch j who live in areas of country c with below-median education levels. The education levels are calculated using the EVS data. The sample is thus restricted to the parts of the sample where we were able to match the author birth place to a NUTS2 region, cf. Section F. For each NUTS2 region, we calculated the average level of education across EVS individuals born before 1940 based on the EVS variable X025. To ensure some overlap in time periods, these education levels are calculated across EVS individuals born before 1940 (Section F). Note that we do not interpret the education levels of these individuals as proxies for the education levels of the authors.

Instead, the assumption is that the distribution of which NUTS2 regions had above and below median education levels did not shift excessively over the time period within one country. The median birth year among authors is 1884 and the median birth year among the EVS respondents born before 1940 with information on education is 1937. Thus, we judge that the actual distribution of above vs below median NUTS2 regions correlates with the EVS distribution. We allow education medians to differ within countries, as the intention of this measure is to capture the different naming conventions across education groups within countries.

Nobility NI: The Nobility Name Index measures the relative frequency of a name among nobles. We calculate using an equation that is similar to Equations (4) and (5) used to calculate RNIs based on frequencies of names among churches and biblical figures:

$$SNI_j^{noble} = \frac{Pr(name_j|Noble)}{Pr(name_j|Noble) + Pr(name_j|Population)},$$

where $Pr(name_j|Noble)$ is the percentage of university students in the RAG database that had a name belonging to namebranch j and whom were nobles. $Pr(name_j|Population)$ is the percentage of authors with the name. For these calculations, we restrict the sample to authors with a name that also occurred among the university students. The sample shrinks by 15%.

UrbanNI: The Urban Name Index measures the "urbanness" of a name by the frequency of a name in urban areas, compared to the full population. To calculate we again use a formula similar to Equation (6) used to calculate RNI based on frequencies of names among persons with religious professions:

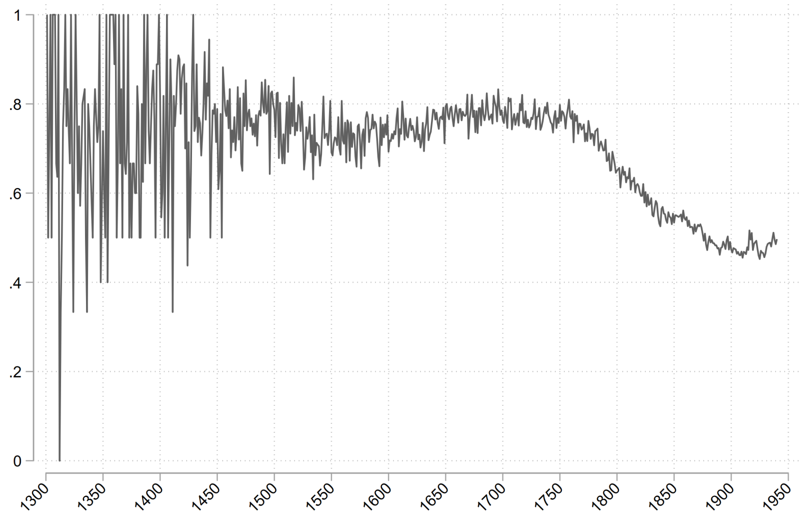
$$SNI_{jc}^{urban} = \frac{Pr(name_{jc}|Population_{urban})}{Pr(name_{jc}|Population_{urban}) + Pr(name_{jc}|Population_{rural})}, \tag{9}$$

where $Pr(name_{jc}|Population_{urban}$ is the percentage of authors with name branch j who were born within 5 km of a city center of country c. $Pr(name_{jc}|Population_{rural})$ is the percentage of authors with namebranch j born further away than 5 km of a city center in country c. Again, we allow the definition of urban vs rural names to differ within each country. For instance, what distinguishes a name as rural is different in France versus Germany.

B Additional material for validity checks of religiosity measures

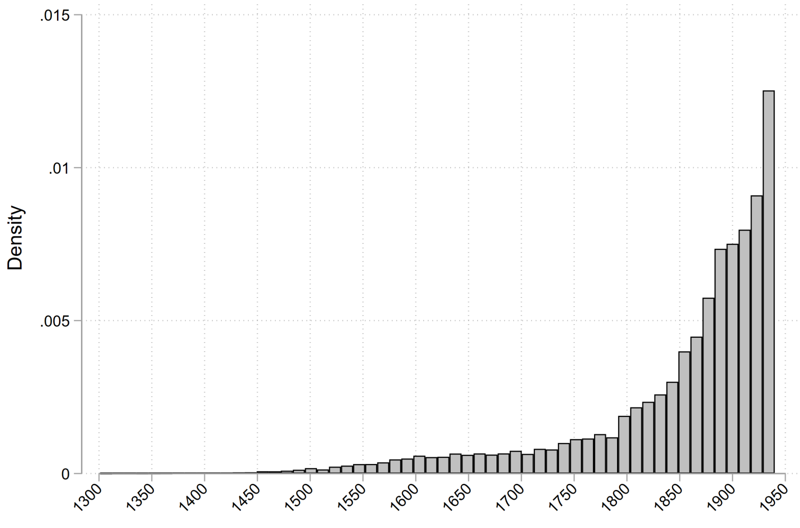
This appendix includes additional figures and tables for the data description of the author database and for robustness checks of the validity checks based on authors.

Figure B1: The share of authors who share name with a major religious figure over time



The share of authors with a name shared with a prominent religious figure throughout the sample period. The year refers to their birth year.

Figure B2: The distribution of birth years among authors



The distribution of birth years for the 487,156 authors in the sample.

Table B1: Religious occupations and names, excluding top-10 names

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Removing top-10 religious names												
Sample excludes:	None	John	Charles	Joseph	Francis	George	Ludwig	Peter	Paul	Anthony	Christos	Top-10
Religious name	0.014***	0.012***	0.016***	0.013***	0.014***	0.015***	0.014***	0.014***	0.014***	0.014***	0.014***	0.011***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
R-squared	0.29	0.28	0.30	0.30	0.30	0.29	0.30	0.29	0.29	0.29	0.29	0.29
Observations	325462	286223	307354	316652	316617	316249	318403	318953	319562	320922	320080	212703
Mean dep var	0.10	0.090	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.100	0.081
Panel B: Removing top-10 non-religious names												
Sample excludes:	None	Henry	Frederick	William	Ernest	Rudolf	Herman	Otto	Walter	Conrad	Gerard	Top-10
Religious name	0.014***	0.014***	0.013***	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.29	0.30	0.30	0.30	0.30	0.29	0.30	0.29	0.29	0.29	0.29	0.32
Observations	325462	314332	314802	316436	320259	320843	320707	321223	321466	321732	322498	265108
Mean dep var	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.11

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The sample is the full sample in columns (1). Panel A excludes the top-10 most frequent religious names one-by one in columns (2)-(11) and excludes them all in column (12). Panel B excludes the top-10 most frequent non-religious names one-by one in columns (2)-(11) and excludes them all in column (12). All regressions include the baseline controls for gender, top-10 most common names, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with a religious name are more likely to hold a religious occupation, even after excluding the most common religious and non-religious names.

Table B2: Religious occupations and names, excluding least and most common names

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)
Common names excluded:	<1 pctile	<10 pctile	<25 pctile	>99 pctile	>90 pctile	>75 pctile
Religious name	0.014***	0.016***	0.019***	0.014***	0.014***	0.014***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
R-squared	0.29	0.29	0.29	0.29	0.29	0.30
Observations	321913	289603	237431	325462	325462	246415

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The samples in the first three columns exclude authors with rare names, which means excluding authors with names held by less than the 1st percentile in the distribution of the number of authors with each name in column (1), less than the 10th percentile in column (2), and less than the 25th percentile in column (3). The samples in the last three columns exclude authors with the most common names, which means excluding authors with names held by more than the 99th percentile in column (4), 90th percentile in column (5), and 75th percentile in column (6). The samples in columns (4) and (5) are the same, since 11 pct of the sample is named John. All regressions include the baseline controls for gender, top-10 most common names, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The results are robust to removing authors with most rare or most common names.

Table B3: Religious occupations and different types of religious names

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	0.014***							
	(0.001)							
Saint name		0.011***						
		(0.001)						
Biblical name			0.019***					
			(0.002)					
Saints RNI				0.029***				
				(0.003)				
New Testament RNI					0.023***			
					(0.004)			
Old Testament RNI						0.018***		
						(0.004)		
Average RNI							0.038***	
							(0.004)	
Maximum RNI								0.030***
								(0.003)
R-squared	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
Observations	325462	325462	325462	325462	325462	325462	325462	325462
Mean dep var	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure in column (1), a saint in column (2), or a biblical figure in column (3). In column (4) the religiosity measure is instead the RNI based on the frequency of the name among churches. All regressions include controls for gender, top-10 most common names, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with a religious name are more likely to hold a religious occupation, independent of the definition of a religious name.

Table B4: Religious occupations and names - different levels of clustering

Dep var: Religious occupation						
Clustering level:	1x1 grid	2x2 grid	5x5 grid	country	city	name branch
	(1)	(2)	(3)	(4)	(5)	(6)
Religious name	0.014***	0.014***	0.014***	0.014***	0.016***	0.014***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.003)
R-squared	0.29	0.29	0.29	0.29	0.27	0.29
Observations	325462	325462	325462	325462	122810	325462
Number clusters	652	225	52	41	1654	1789

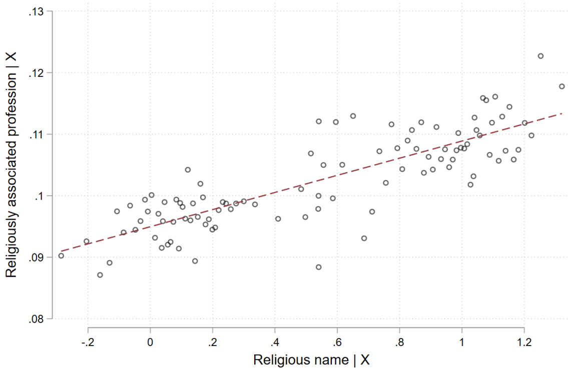
OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All regressions include controls for gender, top-10 most common names, number of occupations, and birth year by 1x1 degree grid cell fixed effects. Robust standard errors in parentheses clustered at the 1x1 grid cell level in column (1), 2x2 grid cell level in column (2), 5x5 grid cell level in column (3), country level in column (4), city level in column (5), and the name branch level in column (6). *, **, and *** indicate significance at the 10%, 5%, and 1% level. The sample in column (5) includes only authors born within 100 km of a city in the Buringh (2021) dataset.
Result: The result is not sensitive towards the particular level of clustering.

Table B5: Religious occupations and names across European regions

Dep var: Religious occupation				
	(1)	(2)	(3)	(4)
	North	South	East	West
Religious name	0.011***	0.017***	0.014***	0.014***
	(0.001)	(0.002)	(0.002)	(0.002)
R-squared	0.31	0.29	0.34	0.26
Observations	162034	161819	162531	161632
Mean dep var	0.10	0.099	0.095	0.11
Number grids	272	409	389	276

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The sample is restricted to North of the median latitude in column (1), South of the median latitude in column (2), East of the median longitude in column (3), and West of the median longitude in column (4). All regressions include controls for gender, top-10 most common names, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The relation between religious names and religious occupations holds in the North, South, East, and West, but is slightly weaker in the North and stronger in the South.

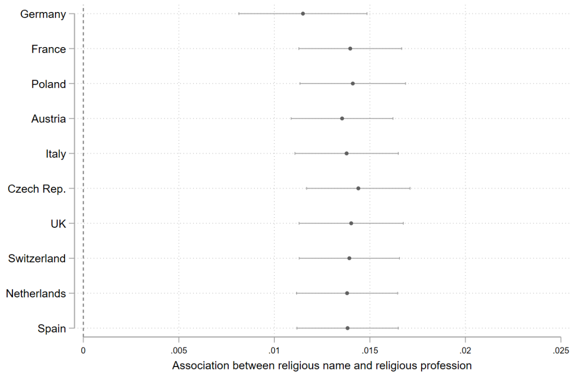
Figure B3: Religious names and religious occupations



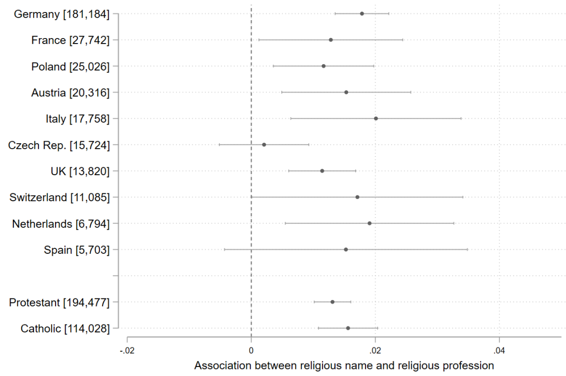
Binned added variables plot where the 325,462 observations are binned into 100 equally sized bins. The line represents the OLS estimate of β in Equation (2), corresponding to the specification in column (8) of Table 3.
Result: No group of observations drive the results.

Figure B4: Religious names and occupations across countries, denominations, and time

(a) Excluding individual countries



(b) Within countries/denominations



(c) The impact over time

The dots reflect the point estimate of β from Equation (2) in regressions, where we exclude each of the countries separately in panel a, restrict the sample to each of the countries or Protestant vs Catholic areas in panel b, and restrict to the time periods 1300-1600, 1600-1700, 1700-1750, 1750-1800, and twenty-year intervals thereafter in panel c. The numbers in squared brackets in panel b are the number of authors born in each country or denomination. The regressions include the baseline controls for gender, number occupations, dummy for top-10 most common names, and 1x1 degree grid cell by birth year fixed effects, except in panels b and c where the fixed effects include exclusively year or grid cell fixed effects, respectively. The 95% confidence bounds with standard errors clustered at the 1x1 grid cell level are shown throughout. **Result:** The association between religious names and occupations is not specific to any country, nor to Protestant or Catholic areas. The association exists before 1600, after 1750, and lasts up until the most recent part of the period.

Table B6: Religious occupations and names split by Protestant areas

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)
Sample:	Full	Full	Full	Cath	Prot
Religious name	0.014***	0.014***	0.013***	0.016***	0.013***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
Protestant area		0.0018			
		(0.005)			
Protestant Name Index			-0.014***		
			(0.005)		
R-squared	0.29	0.29	0.29	0.30	0.31
Observations	313296	313296	313296	114028	194477
Mean dep var	0.10	0.10	0.10	0.083	0.11
Number grids	423	423	423	309	185

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. Protestant Name Index measures the frequency of the name in areas that were Protestant in the 16th century, compared to the population at large. The sample is restricted to the sample with information on whether the area was Protestant in the 16th century in columns (1)-(3), areas that were not Protestant at any point in time in the 16th century in column (4), and areas that were Protestant at any point in time in the 16th century column (5). All regressions include baseline controls for gender, top-10 most common names in each country, number of occupations, and 1x1 degree grid cell by birth year fixed effects, except that the set of fixed effects in column (3) includes 5x5 degree grid cell by birth year fixed effects. The latter to probe the more realistic estimate on the Protestant dummy, as the 1x1 grid cell fixed effects may be too restrictive as Protestant areas are defined in circles of 100 km radius. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The relation between religious names and religious occupations is nearly identical in Protestant and non-Protestant areas.

Table B7: Religious occupations and different controls for name commonness

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	0.013***	0.014***	0.013***	0.014***	0.012***	0.013***	0.013***	0.014***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
First name share within country	0.075							
	(0.056)							
Common first name country dummy		-0.00079						
		(0.002)						
First name share within country-decade			0.016					
			(0.010)					
Common first name country-decade dummy				-0.0057***				
				(0.002)				
First name share within country-gender					0.095**			
					(0.044)			
Common first name country-gender dummy						0.000037		
						(0.001)		
First name share within country-gender-decade							0.016	
							(0.010)	
Common first name country-gender-decade dummy								-0.0046***
								(0.002)
R-squared	0.29	0.29	0.29	0.29	0.29	0.29	0.29	0.29
Observations	325462	325462	325462	325462	325462	325462	325462	325462

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The measure of first name shares controlled for varies across columns. The measure in column (4) is that used throughout the paper. In all columns, the first name share measures the number of persons with the particular first name as a share of persons within the particular group. The first name dummy equals one for the top-10 names in the particular group. The groups are: Country of birth (column 1 and 2), country of birth and 10 years prior to birth (3 and 4), country of birth and gender (5 and 6), country of birth, gender, and 10 years prior to birth (7 and 8). All regressions include controls for the gender, number of occupations and 1x1 degree grid cell fixed effects, year fixed effects, and grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with a religious name are more likely to hold a religious occupation, independent of the control for name commonness.

Table B8: Religious occupations and religiosity of authors accounting for ethnicity

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ethnicity:		Germanic	Celtic	Italic	Slavic	Baltic	Basque	Uralic	Greek	Turkic	Semitic
Religious name	0.014***	0.014***	0.014***	0.013***	0.013***	0.013***	0.013***	0.012***	0.011***	0.012***	0.0089***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Ethnicity		0.0013	0.0015	0.00076	0.0074**	0.0072***	0.0029	0.0070**	0.0088***	0.0089***	0.021***
		(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.003)	(0.003)
R-squared	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30	0.30
Observations	298430	298430	298430	298430	298430	298430	298430	298430	298430	298430	298430

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All baseline controls included throughout: Top-10 most common names, gender, number occupations, and grid cell by year fixed effects. In addition, controls for the ethnicity of a name are added across the columns. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and occupation holds when accounting for ethnicity.

Table B9: Religious occupations and religiosity of authors - additional controls

Dep var: Religious occupation	(1)	(2)	(3)	(4)	(5)	(6)
Religious name	0.014***	0.014***	0.014***	0.015***	0.013***	0.012***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Birthplace latitude	0.0071					
	(0.008)					
Birthplace longitude		-0.0049				
		(0.007)				
Distance to Wittenberg (1000 km)			-0.0016			
			(0.074)			
Migration distance				-0.0014***		
				(0.001)		
Urban					-0.054***	
					(0.005)	
R-squared	0.29	0.29	0.29	0.33	0.30	0.37
Observations	325462	325462	325462	179491	325462	286105
Grid x Year FE	Y	Y	Y	Y	Y	N
City x Year FE	N	N	N	N	N	Y

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All baseline controls included throughout: Top-10 most common names, gender, number occupations, and grid cell by year fixed effects, except that the fixed effects in column (5) are instead city by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses, except for column (6), where standard errors are clustered at the city level. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Result: The association between religious names and occupation holds when adding controls for latitude, longitude, distance to Wittenberg, urban status, and city x year fixed effects.

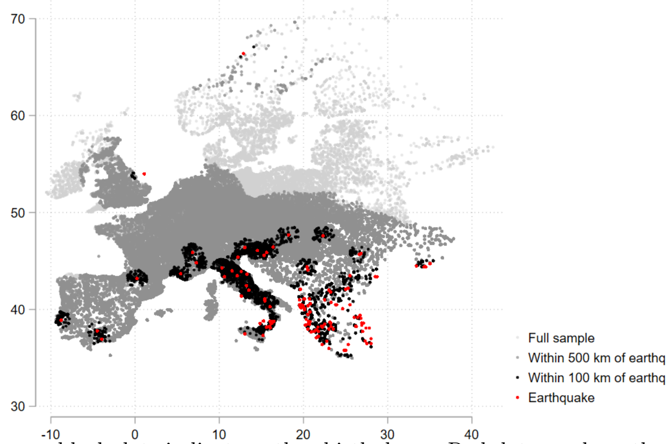
C Earthquakes as a natural experiment

The data on earthquakes is the Significant Earthquake Database created by the National Oceanic and Atmospheric Administration (NOAA)'s National Centers for Environmental Information (`https://www.ngdc.noaa.gov/`). The database is a global listing of significant earthquakes from 2,100 BC to the present. A significant earthquake is defined in the database as one that meets at least one of the following criteria: caused deaths, caused moderate damage (approximately $1 million or more), magnitude 7.5 or greater, Modified Mercalli Intensity (MMI) X or greater, or the earthquake generated a tsunami. As there are very few earthquakes early in the sample and as earthquake detection technology improved greatly over time, we restrict the sample to earthquakes after 1700 and we account for year fixed effects and location-specific time-trends throughout. We further restrict the sample to earthquakes of magnitude 5 or larger, following Bentzen (2019a).

A total of 292 such earthquakes hit within 100 km of a birthplace of an author within 60 years of their birth. The choice of 100 km follows Bentzen (2019a). The choice of 60 years follows the analysis in Figure 4. 63,723 authors in the sample were born within 60 years of one of these earthquakes. 131 of the earthquakes hit within 20 years before an author's birth. These are the earthquakes that we categorize as our main treatment, based on the analysis in Figure 4. According to this definition, 11,119 authors were treated. These individuals experienced earthquakes ranging from magnitude 5 to 8 with a mean (median) magnitude of 6.15 (6). To increase similarity between control and treatment groups, we restrict the control group to individuals who were born within 60 years and 500 km of one or more earthquakes.

Figure C1 depicts the 131 earthquakes and the distribution of treated and untreated authors. The majority hit Southeastern Europe where only few authors were born, meaning that these earthquakes will not greatly influence the analysis. Results are robust to restricting to the 48 major earthquakes that hit Europe to the West of longitude 20 (Table C1). Results are also robust to excluding the two earthquakes that hit to the North of latitude 50 (Table C1). One of these earthquakes was the 5.8 magnitude Luroy earthquake that hit Northern Norway in 1819, known as the largest historical earthquake that hit Fennoscandia. The earthquake was felt throughout a large district in Norway, especially at Salten and Helgeland in the province of Nordland (Mallet & Mallet, 1858). The other was the largest known earthquake in the North Sea region, a 6.0 magnitude earthquake that hit the Dogger Bank area in the open North Sea. It was felt throughout the UK, Belgium, the Netherlands, northern France and Germany, southwestern Norway, and Denmark (Yalçıner, 2009).

Figure C1: Distribution of authors and earthquakes



Grey or black dots indicate author birthplaces. Red dots mark earthquakes.

To examine whether earthquakes increase religious naming, we estimate equations of the form:

$$relname_{irgt} = \sum_{x=5}^{90} \beta_x earthquake_{rgt-x} + \sum_{x=5}^{60} \delta_x earthquake_{rgt+x} + \gamma_g + \gamma_t + \kappa t_r + \omega X_{irgt} + \varepsilon_{irgt} \quad (10)$$

where $relname_{irgt}$ is a dummy equal to one if author i born in year t in 2x2 degree grid cell r within 5x5 degree grid cell g shared name with a major patron saint or biblical figure, zero otherwise. $earthquake_{rt-x}$ is a dummy equal to one if an earthquake hit within 100 km of an author's birthplace within x years before birth, during the birth year or within one year after the birth of individual i. $earthquake_{rt+x}$ is a dummy variable equal to one if an earthquake hit within one plus x years after a person's birth. These latter earthquakes should not matter for name giving and serve as a placebo check. Choosing the length of intervals x is a weighing between getting as close to the year of birth as possible, obtaining long enough placebo intervals, and ensuring that one or more earthquakes hit within the intervals. To produce Figure 4, we choose intervals 5, 10, 20, 30, 60, and 90 years prior to birth and post birth intervals up to 60 years after birth.[84] If we expect that treatment occurs during the 30 years leading up to one's birth, this setup leaves us with 60 years before treatment and 60 years after treatment. The omitted category for the estimation in Figure 4 consists of authors who were not hit by earthquakes neither within 90 years before birth or 60 years after birth, but who are born within 500 km of an earthquake. These comprise 87.3% of the sample. Figure 4 reveals that religious naming is more likely within 30 years prior to birth and thus the robustness checks in Table C1 are conducted using 30 years before birth and a placebo period of 30 years after birth. The omitted category for the estimations in Table C1 consists of authors who were not

[84]For estimations with more than one interval before or after birth, the intervals do not overlap. Thus, the impact of an earthquake that hit within 5 years before birth is caught by β_x, while the impact of an earthquake that hit between 5-10 years before birth is caught by β_{10}.

hit by earthquakes within 30 years before or after birth, but who are born within 500 km of an earthquake. These comprise 94.8% of the sample.

γ_g are 5x5 degree grid cell fixed effects of the author's birthplace location. Since we define a birthplace as having been hit if an earthquake hit within 100 km, we would not want to restrict comparison to 100x100 km grid cells as done in the analysis so far. The main variation in such a regression would have come from comparing persons born within 100 km of an earthquake to those living, say on average 40 km further away. We would not expect large differences in the tendency for religious coping between these individuals, cf Bentzen (2019a). Instead, we compare the impact of earthquakes within 500x500 km grids, where variation in earthquake impact exists. The sample spans 36 such grid cells. To further account for location-time confounders in the absence of grid-cell by year fixed effects, we add 2x2 degree grid cell specific trends, t_r. γ_t are birth year fixed effects and X_{irt} are baseline control variables for top-10 most common names and gender. To ensure that results are not contaminated by physical destruction by earthquakes, we exclude individuals whose birthplace was hit by an earthquake within 20 kilometers 30 years before or after birth. Results are robust to including these (column 7 of Table C1). Standard errors are clustered at the 2x2 degree grid cell level.

Note that the dependent variable - sharing name with a religious figure - does not vary over time, meaning that we have only one observation per author. The time-variation in earthquakes instead produces variation across the variables measuring the timing of earthquakes relative to the timing of author birth. An alternative data structure where information for each author was artificially replicated 240 or 87600 times (the number of years/days between 1700 and 1940) and regressed on a yearly or daily earthquake dummy is less attractive for two reasons. First, how to construct the dependent variable containing the naming decision is ambiguous and second, such a specification would produce estimates biased towards zero, akin to any other regression where the dependent variable is more aggregated than the explanatory variable.

Figure 4 shows the estimates of Equation (10), revealing that parents who experienced an earthquake within 5 years prior to giving birth are more likely to name their child after a religious figure. The same is true for earthquakes that hit up to 20 (and partly up to 30) years before birth, which most likely all hit within the lifespan of the parents.[85] Reassuringly, earthquakes more than 30 years prior to birth did not influence name giving and neither did earthquakes up to 60 years after birth.

Figure 4 includes baseline controls for top-10 most common names, gender, year fixed effects, 5x5 degree grid cell fixed effects, and 2x2 degree grid cell specific trends. Results are not sensitive towards the inclusion of controls nor to the different samples (Table C1).

[85]The slight (insignificant) rise in the estimate moving from 10 to 20 years before birth may reflect earthquakes that hit just before the parents were born and thus may capture the naming decision of grandparents rather than parents. This is not a problem for the purpose with the exercise, which is to probe whether name giving responds to earthquake events in consistency with religious coping.

The raw correlation is significant only at the 12% level, exclusively due to the tendency for the number of detected earthquakes to rise over time with improved earthquake detection technology and the share of religious names to fall. When adding year of birth fixed effects in column (2), significance increases well beyond the 1% level.

One concern is that the impact of earthquakes is biased due to the staggered design of treatment (De Chaisemartin & d'Haultfoeuille, 2020; Callaway & Sant'Anna, 2021). Since already-treated observations enter the control group, the estimate of the treatment effect may be biased towards zero if the impact of treatment falls with time since treatment and away from zero if the impact rises with time since treatment. The problem is larger the more persistent the treatment effect and the larger the share of the sample being treated. In our case, we expect the impact of treatment to fall with time since treatment, consistent with findings by Bentzen (2019a). This would bias our estimates towards zero. However, since the impact of treatment is expected to be rather short-lasting and since the vast majority of our control-group remains untreated, we judge this potential bias to be small.

Table D4 replicates columns (4)-(12) of Table 8, but instead of the direct impact of earthquakes, the table includes IV estimations where the religious name dummy is instrumented with the baseline earthquake dummy in columns (1) and (5) and the earthquake dummy which only turns on for earthquakes that hit the country of birth. These specifications are crucial for the validity of the exclusion restrictions. While the specifications in columns (1) and (5) allow for controlling for earthquakes after birth, the specifications in the remaining columns also allow for controlling for earthquakes before birth in neighboring countries and for the general distance to the earthquakes.

Table C1: Religious names and earthquakes - robustness checks

Dependent variable: Religious name	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Earthquake within 20 years before birth	0.018	0.048***	0.028**	0.038***	0.032**	0.033***	0.031**	0.029**	0.042***
	(0.014)	(0.016)	(0.014)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
Earthquake within 20 years after birth	-0.014	0.00013	-0.014	-0.0077	-0.015	-0.015	-0.013	-0.0096	-0.0031
	(0.016)	(0.015)	(0.012)	(0.011)	(0.011)	(0.011)	(0.012)	(0.011)	(0.012)
Common name					0.16***	0.18***	0.18***	0.13***	0.18***
					(0.028)	(0.028)	(0.027)	(0.023)	(0.029)
Female						0.083***	0.082***	0.096***	0.091***
						(0.009)	(0.009)	(0.007)	(0.008)
R-squared	0.000031	0.032	0.042	0.049	0.072	0.075	0.075	0.067	0.079
Observations	273562	273562	273562	273558	273558	273558	274863	436343	258229
Mean dep var	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.53	0.55
Year FE	N	Y	Y	Y	Y	Y	Y	Y	Y
5x5 grid FE	N	N	Y	Y	Y	Y	Y	Y	Y
Grid-cell trends	N	N	N	Y	Y	Y	Y	Y	Y
Excluding hit	Y	Y	Y	Y	Y	Y	N	Y	Y
Similar controlgroup	Y	Y	Y	Y	Y	Y	Y	N	Y
Geographic spread	Full	Full	Full	Full	Full	Full	Full	Full	$< lon20$
Number 2x2 grids	198	198	198	194	194	194	194	282	139
Number 5x5 grids	44	44	44	44	44	44	44	59	31

OLS across authors. The baseline sample is used throughout (including authors born within 500 km of an earthquake that hit within 60 years before or after their birth and excluding authors within 20 km of an earthquake that hit 20 years before or after their birth), except for columns (7)-(9), where authors hit by an earthquake within 20 km within 20 years before or after birth are included (column 7), the control group includes all other authors in the sample (8), and where authors born to the East of longitude 20 are excluded. Robust standard errors clustered at the 2x2 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Result: The finding that religious naming rises for authors born after earthquakes is robust to the inclusion of different control variables, fixed effects, and samples.

D Additional material for Further Identification

Table D1: Placebo: Increasing potential bias from name changing

Dependent variable:	Religious occupation			Scientific occupation		
	(1)	(2)	(3)	(4)	(5)	(6)
Religious name	0.012***	0.013***	0.013***	-0.015***	-0.012***	-0.012***
	(0.003)	(0.001)	(0.001)	(0.005)	(0.002)	(0.002)
Religious name x Number first names	0.00074			0.0038		
	(0.003)			(0.005)		
Religious name x > one first name dummy		0.0014			0.0047	
		(0.003)			(0.005)	
R-squared	0.29	0.29	0.29	0.16	0.16	0.17
Observations	325462	325462	254841	325462	325462	254841
Sample	Full	Full	One name	Full	Full	One name

The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion in columns (1)-(3) and a dummy equal to one if one of the author's occupations is associated with science in columns (4)-(6). Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All baseline controls included throughout: Top-10 most common names, gender, number occupations, and grid cell by year fixed effects. In addition, columns (1) and (4) include a control for the number of first names held by the person and columns (2) and (5) include a dummy equal to one if the person has more than one first name. The sample includes the full sample, but is restricted to the sample with individuals holding only one name in columns (3) and (6). Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and occupations does not strengthen for those with many names. We interpret this as indicating that results are not inflated due to name changing.

Table D2: Placebo: Increasing potential bias from selection

Dependent variable:	Religious occupation				Scientific occupation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	0.014***	0.013***	0.014***	0.013***	-0.0099***	-0.0086***	-0.011***	-0.0095***
	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.003)	(0.003)	(0.003)
Religious name x Number publications			-0.0057				0.062***	
			(0.007)				(0.018)	
Religious name x Number citations				-0.0039				0.34***
				(0.056)				(0.130)
R-squared	0.29	0.34	0.34	0.34	0.16	0.24	0.25	0.25
Observations	325462	107861	107861	107279	325462	107861	107861	107279
Sample	Full	Wcat	Wcat	Wcat	Full	Wcat	Wcat	Wcat

The dependent variable is a dummy equal to one if one or more of the author's occupations is associated with religion in columns (1)-(4) and a dummy equal to one if one or more of the author's occupations is associated with science in columns (5)-(8). Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All baseline controls included throughout: Top-10 most common names, gender, number occupations, and grid cell by year fixed effects. In addition, columns (3) and (7) include a control for the person's number of publications, recorded by Worldcat and columns (4) and (8) include the number of citations recorded by Worldcat. The sample is the full sample in columns (1) and (5) and the sample restricted to the Worldcat sample in remaining columns. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and occupations is not stronger for more successful authors (authors with information in Worldcat or authors with more publications or citations). We interpret this as an indication that results do not strengthen when increasing potential selection bias. On the contrary, the relation between names and scientific occupations is weaker among the more successful authors.

Figure D1: Illustration of collider bias

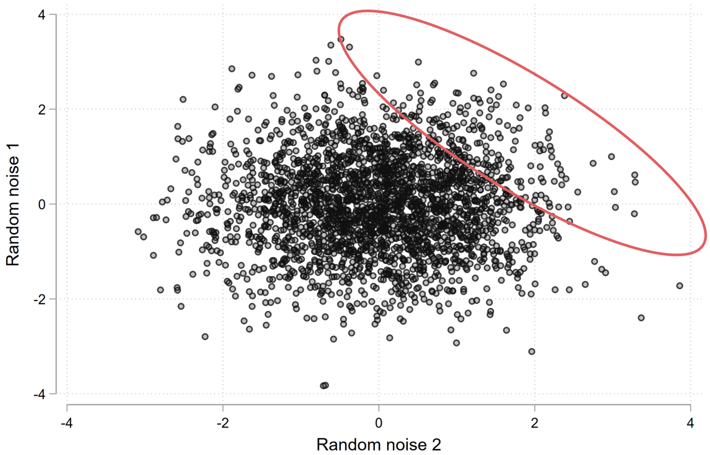


Illustration of collider bias. If selection into authorship is higher for both collider variables, we could expect to see a spurious negative relation between the two.

Table D3: Religious occupations and names - split by gender

Dependent variable:	Religious occupation		Scientific occupation	
	(1)	(2)	(3)	(4)
Sample:	Female	Male	Female	Male
Religious name	0.0061***	0.015***	-0.011**	-0.0095***
	(0.002)	(0.001)	(0.005)	(0.002)
R-squared	0.29	0.30	0.27	0.17
Observations	25052	290861	25052	290861
Mean dep var	0.022	0.11	0.11	0.19
Number grids	314	640	314	640

The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion in columns (1)-(2) and a dummy equal to one if one of the author's occupations is associated with science in columns (3)-(4). Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. Controls included throughout: Top-10 most common names, number occupations, and grid cell by year fixed effects. The sample includes the sample of women in columns (1) and (3) and the sample of men in columns (2) and (4). Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and occupations holds within samples of men and women separately.

Table D4: Earthquakes as shocks to religiosity, IV regressions

Dependent variable:	Religious occupation				Scientific occupation			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	0.087	0.31*	0.31**	0.45+	-0.89***	-0.74**	-0.75**	-0.89*
	(0.215)	(0.159)	(0.158)	(0.313)	(0.315)	(0.366)	(0.370)	(0.458)
Earthquake before birth		-0.0066	0.017	0.025		-0.0048	-0.024	-0.0080
		(0.008)	(0.033)	(0.062)		(0.016)	(0.040)	(0.078)
Earthquake after birth	0.0055	-0.0072	-0.023	0.018	-0.036**	-0.00013	0.0067	-0.042
	(0.009)	(0.013)	(0.037)	(0.047)	(0.015)	(0.016)	(0.054)	(0.064)
Earthquake after birth same country		0.022*	0.021*	0.024		-0.040*	-0.040*	-0.052+
		(0.012)	(0.012)	(0.025)		(0.021)	(0.021)	(0.035)
Avg earthquake distance before birth			0.12	0.17			-0.10	-0.099
			(0.176)	(0.329)			(0.246)	(0.422)
Avg earthquake distance after birth			-0.089	0.11			0.039	-0.25
			(0.205)	(0.220)			(0.294)	(0.324)
Observations	196000	196000	196000	88383	196000	196000	196000	88383
Mean dep var	0.076	0.076	0.076	0.092	0.18	0.18	0.18	0.15
Kleibergen-Paap first stage F	5.24	7.91	8.85	5.55	5.24	7.91	8.85	5.55
Cragg-Donald first stage F	13.7	14.5	14.1	4.74	13.7	14.5	14.1	4.74
5x5 grid and Year FE	Y	Y	Y	Y	Y	Y	Y	Y
2x2 grid trends	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y
Sample	Full	Full	Full	Migrant	Full	Full	Full	Migrant

IV regressions across authors. The dependent variable is a dummy equal to one if the author has a religious occupation in columns (1)-(4) or a dummy equal to one if the author has a scientific occupation in columns (5)-(8). The endogenous variable is a dummy equal to one if the author shares name with a significant religious figure. The excluded instrument is a dummy equal to one if an earthquake hit within 100 km of the author's birthplace within 20 years before their birth in columns (1) and (5) and a dummy equal to one if the earthquake hit in the author's country of birth in columns (2)-(4) and (6)-(8). Baseline controls included throughout: Top-10 most common names, gender, number occupations, 5x5 grid cell fixed effects, year of birth fixed effects, and 2x2 grid cell trends. The baseline sample includes authors born within 500 km of an earthquake that hit within 60 years before or after their birth and excluding authors within 20 km of an earthquake that hit within 20 years before or after their birth. The sample includes the full sample throughout, except from columns (4) and (8), where it is restricted to authors who die in a different location than they were born. Robust standard errors clustered at the 2x2 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Having a religious name increases the likelihood of ending up in a religious occupation and decreases the likelihood of ending up in a scientific occupation, even when religious naming is restricted to names given after exogenous shocks.

E Refractory clergy and religious naming

In this section, we check whether our names-based measure of religiosity correlates with the only other measure of historic religiosity that we know of; the share of refractory clergy. The measure is available across French departments. The refractory clergy share measures the percentage of priests in a French department who refused to swear an oath to the civil constitutions of the clergy. This variable has been used in other research to indicate the general level of religiosity in the society (Squicciarini, 2020). The variable is measured in 1791. To match with the author names, we restrict the sample of authors to those born within 30 years of the refractory clergy measure, ie. authors born between 1761 and 1821. We restrict the sample to departments where more than 10 authors were born during the period. This means excluding Lozere, which is currently the least populous department of France.

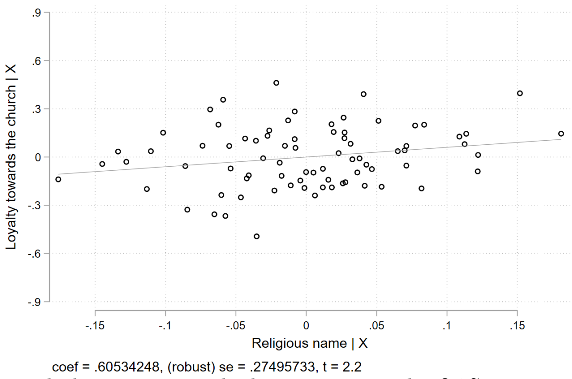
We run regressions of the form:

$$clergy_d = \alpha + \beta relname_d + \omega X_d + \varepsilon_d, \tag{11}$$

where $relname_d$ is the share of authors born during the period 1761-1821 in department d who shared name with a major patron saint, zero otherwise. As France was predominantly Catholic during this period, we use the definition of a religious name based on patron saints. $clergy_d$ measures the share of clergymen in department d who swore an oath of loyalty towards the church. Control variables X_d include the share of authors in department d with common names and the share of female authors and the full set of controls used by Squicciarini (2020): Population, temperature, precipitation, wheat suitability, pre-industrial activities, distance from Paris, a measure of the power of the king, knowledge elites, and the enrollment rate.

Figure E1 shows the added variables plot of the association between the likelihood of sharing name with a patron saint and the share of refractory clergy in the department of the author's birth, including the full set of controls from Squicciarini (2020).

Figure E1: Religious names and a loyal clergy



Added variables plot across 78 French departments. The line represents the OLS estimate of β in Equation (11), including controls for population, temperature, precipitation, wheat suitability, pre-Industrial activities, distance from Paris, power of the king, number Encyclopedie subscribers, and elementary school enrollment rates (from Squicciarini (2020)), corresponding to column (3) of Table E1.
Result: The share of authors with religious names is higher in French departments where the clergy was more likely to be loyal to the Church.

Table E1 shows the sensitivity of the results towards inclusion of the various controls. The relation between refractory clergy and the share of authors with religious names becomes insignificant when no controls are included (column 1). The main reason is that more populous departments have larger shares of refractory clergy, but lower shares of authors with religious names (column 2). The relation again becomes insignificant when adding the share of authors with common names and the share of female authors to the extensive list of controls (column 4). Most controls are insignificant, though, and including only the significant controls retains significance in the relation between refractory clergy and religious names shares (column 5).

Table E1: Religious names and refractory clergy

Dependent variable: Share refractory clergy	(1)	(2)	(3)	(4)	(5)
Saint name	0.38	0.59**	0.61**	0.46+	0.54**
	(0.295)	(0.289)	(0.275)	(0.308)	(0.260)
Population 1800		0.22***	0.29***	0.26***	0.29***
		(0.067)	(0.076)	(0.079)	(0.062)
Mean Temperature			0.071	0.034	
			(0.167)	(0.182)	
Mean Precipitation			-0.099	-0.11	
			(0.140)	(0.141)	
Wheat Suitability			-0.019	-0.017	
			(0.023)	(0.022)	
Pre-Industrial Activities			0.0064	0.024	
			(0.045)	(0.045)	
Distance from Paris			0.064**	0.067**	0.065***
			(0.029)	(0.029)	(0.022)
Pays d'Elections			-0.12**	-0.11*	-0.11**
			(0.055)	(0.061)	(0.054)
Knowledge Elites			-0.0082	-0.0087	
			(0.013)	(0.014)	
Enrollment Rate 1851			-0.071	-0.12	
			(0.112)	(0.117)	
Common name share				0.29	
				(0.265)	
Female share				0.83	
				(0.783)	
R-squared	0.015	0.14	0.33	0.35	0.31
Observations	78	78	78	78	78
Mean dep var	0.42	0.42	0.42	0.42	0.42

OLS across departments of France with more than 10 authors born during the period 1761-1821. The controls are those used by Squicciarini (2020) in addition to our two main control variables measuring the share of female authors and authors with common names: Population represents (log) total department population in 1800. Temperature and Precipitation measure the (log) average precipitation and temperature in the 1700–1800 period. Wheat suitability is wheat soil suitability. Pre-Industrial activities is an index of pre-industrial activities in France that includes the number of mines, forges, iron trading locations, and textile manufactures before 1500. Distance from Paris measures the (log) distance from Paris (in km). Pays d'election is a dummy for departments where the king, before 1789, exerted particularly strong power (especially in terms of fiscal and financial matters). Knowledge elites reflects the (log) number of subscribers to the Encyclopedie in 1777–1780. Enrollment rate represents the ratio of students to school-age population (5 to 15 years) in 1851. Robust standard errors in parentheses clustered at the departmental level. +, *, **, and *** indicate significance at the 15%, 10%, 5%, and 1% level.
Result: The likelihood that the clergy were loyal to the church is higher in districts with more authors who share name with a major patron saint.

F European Values Study and religious naming

In this section, we check whether our names-based religiosity measure correlates with information on religious values available from contemporary surveys, conducted by the European Values Study (EVS, 2022). We combine all five waves currently available (1981-1984, 1990-1993, 1999-2001, 2008-2010, 2017-2021). These data encompass responses from 224,434 respondents from 48 countries/regions. Of these, we have information on the NUTS region of interview for 92,879 of them. We know the birth year for 92,429 of them. These respondents were born between 1905 and 2003 with an average (median) birth year of 1963 (1964). Since our data on authors stops in 1940, we restrict to individuals born no later than 1940.

This shrinks the data to 10,313 individuals for whom we have information on self-reported values. These individuals were located in 257 regions of Europe. Eleven questions relate to the religiosity of respondents. We include all eleven. The questions are:[86] "How important is religion in your life?" (on a scale from 0 to 3, where 0 is not at all important and 3 is very important), "How important is God in your life?" (on a scale from 1 to 10, where 0 is not at all important and 10 is very important), "Do you believe in God?" (0-1), "Do you believe in an Afterlife?" (0-1), "Do you get comfort and strength from religion?" (0-1), "How often do you pray to God outside religious services?" (on a scale from 0 to 6, where 0 is never and 6 is every day), "Do you belong to a religious denomination?" (0-1), "How often do you attend religious services?" (on a scale from 0 to 6, where 0 is never, practically never and 6 is more than once a week), "How often did you attend religious services when you were 12 years old?" (on a scale from 0 to 6, where 0 is never, practically never and 6 is more than once a week), "Are you a religious person?" (on a scale from 0 to 2, where 0 is a convinced atheist, 1 is not a religious person, and 2 is a religious person), and the last question is based on a list of features in their child that the person can mark as important to learn at home. If the person mentions religious faith, this last variable takes the values one, zero otherwise.

We match an EVS respondent to an author if the respondent was interviewed in the same NUTS2 region as the author was born. We are able to match 472,857 authors to a NUTS2 region. Of these, 471,636 were matched to a NUTS2 that was also included in the EVS sample.

To increase the comparability between the two datasets, we restrict to authors born after 1900, which reduces the number of authors to 182,595. To estimate the relation between responses in the EVS and the names of the authors, we run two types of regressions. In one set, we estimate the relation across authors with the explanatory variable measuring the average value of the particular question on religion in the particular re-

[86] Several of the variables in the original data are higher the lower the religiosity. We recode these variables, so that higher values indicate higher religiosity. We report here the recoded values. The original variables are A006, F063, F050, F051, F064, F066, F024, F028, and F034.

gion (panel A of Table F1). For this set, we calculate averages across EVS respondents after removing variation in common confounders used in the literature using the surveys; respondent age, age squared, gender, marital status, and birth- and interview year fixed effects.[87] The regression includes our baseline author controls for country of birth and authors' gender, commonality of first name, country- and birth year fixed effects. The results reveal that regions across Europe with higher religiosity levels measured using eleven different questions in surveys are more likely to be inhabited by authors with religious names. Nine out of eleven estimates are significant at conventional levels.

For the other set of estimations, we estimate the relation across respondents in the EVS surveys and regress the degree of religiosity revealed by their responses on the average share of authors with religious names in the particular region (panels B, C, and D). When calculating the religious name share, we first remove variation in the baseline author controls for gender, name commonality, and birth year fixed effects.[88] The regression in panels B, C and D include EVS respondent-level controls for marital status, gender, age, age squared, birth year, and survey year fixed effects. Consistent with the results in panel A, we find in panel B that EVS respondents reveal higher religiosity in all eleven questions in regions where more authors were born with religious names. The estimates are significant for nine of eleven measures.

Both types of regressions include NUTS2 regions from 30-31 countries.

The EVS surveys come with one additional piece of information that is useful for the analysis; measures of respondent income and education and respondents' fathers' education. We add the former two to the set of controls in panel C. Respondent income and education are negatively associated with religiosity in the responses, but adding them as controls does not alter the correlation between the share of religious names and the likelihood of responding with a more religious answer to survey questions significantly. While respondent income and education may be both causes and consequences of respondent religiosity, these controls may bias results. Instead, we add a control for father's education in panel D. Respondents with more educated fathers tend to be less religious, but accounting for this does not influence the relation between religious name shares and religious survey respondents.

In sum, regions where more authors have religious names are also more likely to be populated by individuals who categorize themselves as being more religious, even after accounting for respondent and parent socioeconomic status.

[87]These averages are calculated by first running a regression of the particular EVS response on measures of respondent age, age squared, gender, marital status, and birth- and interview year fixed effects. We then save the residuals and calculate NUTS2 region averages for these residuals. Results are robust to not conducting this adjustment and using the raw EVS responses instead.

[88]These religious name shares are calculated by first running a regression of the religious name dummy on measures of the author's gender, name commonality, and birth year fixed effects. We then calculate NUTS2 region averages for the residuals. Results are robust to not conducting this adjustment and using the raw religious name averages instead.

Table F1: Religious names and survey measures of religiosity

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
EVS measure of religiosity	Importance of religion	Importance of God	Believe in God	Believe in Afterlife	Comfort from religion	Pray outside service	Belong to rel denom	Church attendance	Attendance child	Religious person	Faith child
Panel A: Across authors											
Dep var: Religious name											
Average EVS religiosity	0.008**	0.003**	0.018**	0.012	0.028***	0.003**	0.021***	0.004**	0.004**	0.015**	0.010
	(2.28)	(2.18)	(2.29)	(1.18)	(3.58)	(1.98)	(3.17)	(2.17)	(2.17)	(2.44)	(0.69)
R squared	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
Observations	155078	154623	155078	153217	125447	155078	155078	155078	155078	155078	155078
Country and Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number regions	157	156	157	150	122	157	157	157	157	157	157
Panel B: Across EVS respondents											
Dep var: EVS religiosity measure											
Religious names share	0.086***	0.066*	0.103***	0.072**	0.106***	0.067*	0.145***	0.092***	0.064+	0.080**	0.040
	(2.65)	(1.83)	(2.73)	(2.47)	(2.71)	(1.69)	(4.18)	(2.73)	(1.53)	(2.22)	(0.95)
R squared	0.20	0.19	0.13	0.15	0.16	0.23	0.16	0.18	0.27	0.12	0.15
Observations	8472	8344	8077	7245	5285	8375	8530	8522	8319	8310	8331
Country and Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number regions	241	241	241	239	154	241	241	241	241	241	241
Panel C: Across EVS respondents including income and education controls											
Dep var: EVS religiosity measure											
Religious names share	0.082**	0.075**	0.088**	0.083***	0.122***	0.068+	0.140***	0.100***	0.062	0.065*	0.013
	(2.46)	(2.04)	(2.29)	(2.84)	(2.88)	(1.64)	(3.80)	(2.72)	(1.39)	(1.91)	(0.29)
Income	-0.019	-0.022	-0.020	-0.016	-0.010	-0.031**	-0.007	-0.012	-0.013	-0.050***	-0.035**
	(-1.24)	(-1.40)	(-1.41)	(-1.00)	(-0.50)	(-2.00)	(-0.48)	(-0.77)	(-0.81)	(-3.12)	(-2.40)
Education	-0.067***	-0.098***	-0.110***	-0.045**	-0.090***	-0.086***	-0.093***	-0.029*	-0.019	-0.122***	-0.073***
	(-3.91)	(-5.43)	(-7.10)	(-2.15)	(-4.28)	(-4.87)	(-5.59)	(-1.87)	(-1.23)	(-7.95)	(-4.53)
R squared	0.20	0.20	0.14	0.15	0.17	0.23	0.17	0.17	0.27	0.14	0.15
Observations	6835	6752	6522	5882	4248	6784	6882	6891	6743	6711	6710
Country and Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number regions	239	239	239	236	153	239	239	239	239	239	239
Panel D: Across EVS respondents including control for father's education											
Dep var: EVS religiosity measure											
Religious names share	0.080**	0.063*	0.097**	0.094***	0.099**	0.063+	0.138***	0.098***	0.067	0.080**	0.039
	(2.59)	(1.72)	(2.50)	(3.12)	(2.34)	(1.63)	(3.73)	(2.71)	(1.41)	(2.47)	(1.00)
Father's education	-0.048***	-0.073***	-0.075***	-0.012	-0.060***	-0.081***	-0.061***	-0.024+	-0.021	-0.088***	-0.061***
	(-3.34)	(-4.50)	(-4.56)	(-0.70)	(-2.94)	(-5.69)	(-3.70)	(-1.57)	(-1.34)	(-5.48)	(-4.56)
R squared	0.19	0.19	0.14	0.15	0.16	0.23	0.16	0.18	0.26	0.12	0.14
Observations	6898	6802	6587	5931	4114	6832	6950	6945	6790	6773	6784
Country and Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Number regions	237	237	237	233	154	237	237	237	237	237	237

OLS across authors in panel A and across EVS respondents in panels B, C, and D. The estimates are beta estimates, t-statistics in parenthesis. The dependent variable is a dummy equal to one if the author shares name with a significant religious figure in panel A and the particular EVS religiosity measure in panels B, C, and D. The main independent variable is the NUTS2 regional average (after accounting for respondent gender, age, age squared, marital status, birth year, and survey year) of the EVS measure in panel A and the NUTS2 regional average (after accounting for author gender, name normality, and birth year) share of religious names in panels B, C, and D. In addition to the controls included before aggregating the data, all regressions in panel A include authors' country of birth and birth year fixed effects, and controls for gender and the top-10 most common names and all regressions in panels B-D include respondent country of birth, birth year and survey year fixed effects, and controls for gender, age, age squared, and marital status. Robust standard errors clustered at the NUTS2 regional level. Estimates are the standardized beta estimates and t-statistics are presented in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Result: Regions with higher shares of religious names are more likely to be populated by individuals who categorize themselves as more religious, even after accounting for respondent and parent socioeconomic status.

G Additional material for the analysis of scientific occupations

This appendix includes additional figures and tables for robustness checks of the analysis of the impact of religious names on scientific occupations in Section 3. The checks mirror the checks performed for the analysis of the impact of religious names on religious occupations.

Table G1: Scientific occupations and names, excluding top-10 names

Dep var: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Removing top-10 religious names												
Sample excludes:	None	John	Charles	Joseph	Francis	George	Ludwig	Peter	Paul	Anthony	Christos	Top-10
Religious name	-0.0099***	-0.011***	-0.010***	-0.0085***	-0.0091***	-0.010***	-0.0099***	-0.0095***	-0.0098***	-0.0097***	-0.0100***	-0.0057***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
R-squared	0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.16	0.17	0.16	0.16	0.18
Observations	325462	286223	307354	316652	316617	316249	318403	318953	319562	320922	320080	212703
Mean dep var	0.18	0.19	0.18	0.19	0.19	0.18	0.19	0.18	0.18	0.19	0.19	0.19
Panel B: Removing top-10 non-religious names												
Sample excludes:	None	Henry	Frederick	William	Ernest	Rudolf	Herman	Otto	Walter	Conrad	Gerard	Top-10
Religious name	-0.0099***	-0.011***	-0.010***	-0.010***	-0.0095***	-0.0098***	-0.0097***	-0.0097***	-0.0093***	-0.0098***	-0.0097***	-0.0084***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
R-squared	0.16	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.18
Observations	325462	314332	314802	316436	320259	320843	320707	321223	321466	321732	322498	265108
Mean dep var	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The sample is the full sample in columns (1). Panel A excludes the top-10 most frequent religious names one-by one in columns (2)-(11) and excludes them all in column (12). Panel B excludes the top-10 most frequent non-religious names one-by one in columns (2)-(11) and excludes them all in column (12). All regressions include the baseline controls for the top-10 most common names in each country, number of occupations, 1x1 degree grid cell fixed effects, birth year fixed effects, and grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with a religious name are less likely to hold a scientific or engineering occupation, even after excluding the most common religious and non-religious names.

Table G2: Scientific occupations and names, excluding least and most common names

Dep var: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)
Common names excluded:	<1 pctile	<10 pctile	<25 pctile	>99 pctile	>90 pctile	>75 pctile
Religious name	-0.010***	-0.012***	-0.015***	-0.0099***	-0.0099***	-0.012***
	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
R-squared	0.16	0.17	0.17	0.16	0.16	0.18
Observations	321913	289603	237431	325462	325462	246415

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The samples in the first three columns exclude authors with rare names, which means excluding authors with names held by less than the 1st percentile in the distribution of the number of authors with each name in column (1), less than the 10th percentile in column (2), and less than the 25th percentile in column (3). The samples in the last three columns exclude authors with the most common names, which means excluding authors with names held by more than the 99th percentile in column (4), 90th percentile in column (5), and 75th percentile in column (6). The samples in columns (4) and (5) are the same, since 11 pct of the sample is named John. All regressions include the baseline controls for gender, top-10 most common names, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The results are robust towards removing authors with most rare or most common names.

Table G3: Scientific occupations and different types of religious names

Dep var: Scientific and engineer occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Religious name	-0.0099***								
	(0.001)								
Saint name		-0.0065***							
		(0.001)							
Biblical name			-0.0098***						
			(0.002)						
Saints RNI				-0.014***					
				(0.003)					
New Testament RNI					-0.011***				
					(0.004)				
Old Testament RNI						-0.013***			
						(0.004)			
Average RNI							-0.018***		
							(0.004)		
Maximum RNI								-0.013***	
								(0.003)	
Religious job RNI									-0.088***
									(0.019)
R-squared	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Observations	325462	325462	325462	325462	325462	325462	325462	325462	325462
Mean dep var	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. The measure of a religious name varies across columns. A name is defined as religious if it is shared by a significant saint or biblical figure (our baseline measure) in column (1), a saint in column (2), or a biblical figure in column (3). The following columns use instead RNI, calculated based on Saints to which churches are dedicated (column 4), names in the New Testament (column 5), names in the Old Testament (column 6), the average of these RNI (column 7), the maximum values of the RNIs (column 8), and last the names of individuals with religious occupations. All regressions include controls for the top-10 most common names in each country, number of occupations and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with a religious name are more likely to hold a scientific or engineering occupation, independent of the definition of a religious name.

Table G4: Scientific occupations and names - different levels of clustering

Dep var: Scientific occupation Clustering level:	1x1 grid (1)	2x2 grid (2)	5x5 grid (3)	country (4)	city (5)	name branch (6)
Religious name	-0.0099***	-0.0099***	-0.0099***	-0.0099***	-0.016***	-0.0099***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.003)	(0.003)
R-squared	0.16	0.16	0.16	0.16	0.17	0.16
Observations	325462	325462	325462	325462	122810	325462
Number clusters	652	225	52	41	1654	1789

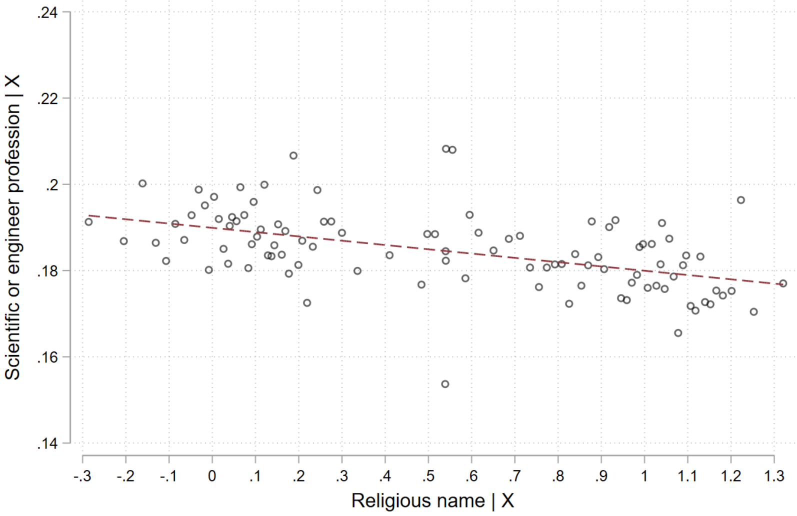
OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All regressions include controls for gender, top-10 most common names, number of occupations, and birth year by 1x1 degree grid cell fixed effects. Robust standard errors in parentheses clustered at the 1x1 grid cell level in column (1), 2x2 grid cell level in column (2), 5x5 grid cell level in column (3), country level in column (4), city level in column (5), and the name branch level in column (6). *, **, and *** indicate significance at the 10%, 5%, and 1% level. The sample in column (5) includes only authors born within 100 km of a city in the Buringh (2021) dataset.
Result: The result is not sensitive towards the particular level of clustering.

Table G5: Scientific occupations and names across European regions

Dep var: Scientific or engineering occupation	(1) North	(2) South	(3) East	(4) West
Religious name	-0.0062***	-0.013***	-0.0093***	-0.010***
	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.17	0.16	0.18	0.15
Observations	162034	161819	162531	161632
Mean dep var	0.21	0.16	0.20	0.17
Number grids	272	409	389	276

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with religion. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The sample is restricted to North of the median latitude in column (1), South of the median latitude in column (2), East of the median longitude in column (3), and West of the median longitude in column (4). All regressions include controls for gender, top-10 most common names in each country, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The negative association between religious names and scientific occupations holds in the North, South, East, and West, but is slightly weaker in the North and stronger in the South.

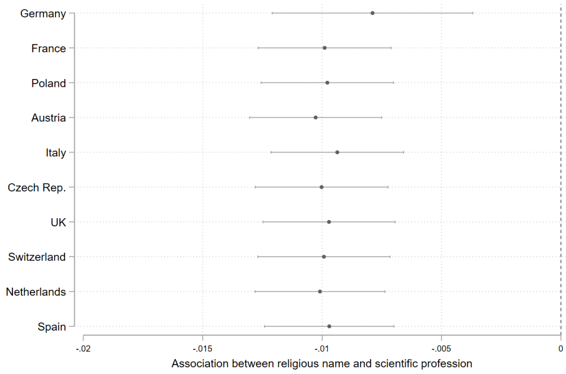
Figure G1: Religious names and science



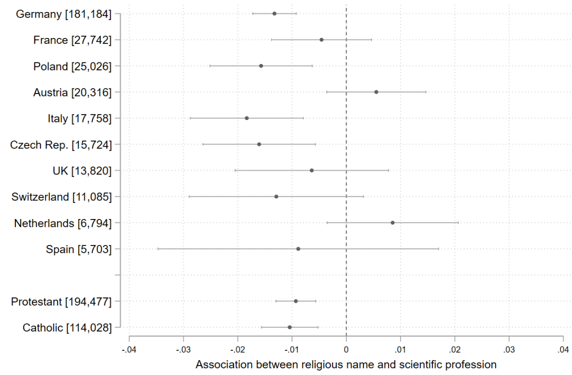
Binned added variables plot where the 325,462 observations are binned into 100 equally sized bins. The line represents the OLS estimate of β in Equation (2), corresponding to the specification in column (8) of Table 5.
Result: Authors with religious names are less likely to have occupations related to science or engineering. No group of observations drive the results.

Figure G2: Religious names and scientific occupations across countries, denominations, and time

(a) Excluding individual countries



(b) Within countries/denominations



(c) The impact over time

The dots reflect the point estimate of β from Equation (2) where occupations relate to science or engineering. We exclude each of the top-10 countries separately in panel a, restrict the sample to each of the countries or Protestant vs Catholic areas in panel b, and restrict to the time periods 1300-1600, 1600-1700, 1700-1750, 1750-1800, and twenty-year intervals thereafter in panel c. The numbers in squared brackets in panel b are the number of authors born in each country or denomination. The regressions include the baseline controls for gender, number occupations, dummy for top-10 most common names, and 1x1 degree grid cell by birth year fixed effects, except in panels b and c, where the fixed effects include exclusively year fixed effects or grid cell fixed effects, respectively. The 95% confidence bounds with standard errors clustered at the 1x1 grid cell level are shown throughout.
Result: The result is not specific to any country nor to Protestant or Catholic areas. The gap in scientific occupations emerges exclusively after 1820 and may be vanishing again for authors born between 1921 and 1940.

Table G6: Scientific occupations and religious names - Protestants vs Catholics

Dep var: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)
Sample:	Full	Full	Full	Full	Cath	Prot
Religious name	-0.010***	-0.010***	-0.0048***		-0.010***	-0.0094***
	(0.001)	(0.001)	(0.001)		(0.003)	(0.002)
Protestant area		0.0080*				
		(0.005)				
Protestant Name Index			0.061***	0.064***		
			(0.007)	(0.007)		
Religious job Name Index				-0.083***		
				(0.018)		
R-squared	0.16	0.16	0.16	0.16	0.20	0.16
Observations	313296	313296	313296	313296	114028	194477
Mean dep var	0.19	0.19	0.19	0.19	0.16	0.20
Number grids	423	423	423	423	309	185

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. Protestant Name Index measures the frequency of the name in areas that were Protestant in the 16th century, compared to the population at large. Religious job Name Index measures the frequency of the name among individuals with religious occupations, compared to the rest of the population. The sample is restricted to the sample with information on whether the area was Protestant in the 16th century in columns (1)-(4), areas that were not Protestant at any point in time in the 16th century in column (5), and areas that were Protestant at any point in time in the 16th century column (6). All regressions include baseline controls for gender, top-10 most common names in each country, number of occupations, and 1x1 degree grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The relation between religious names and scientific occupations is nearly identical in Protestant and non-Protestant areas.

Table G7: Religious occupations and different controls for common names

Dep var: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	-0.011***	-0.0099***	-0.010***	-0.0099***	-0.011***	-0.010***	-0.010***	-0.0099***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
First name share within country	0.19***							
	(0.059)							
Common first name country dummy		0.0070***						
		(0.002)						
First name share within country-decade			0.037***					
			(0.007)					
Common first name country-decade dummy				0.011***				
				(0.002)				
First name share within country-gender					0.20***			
					(0.049)			
Common first name country-gender dummy						0.0082***		
						(0.002)		
First name share within country-gender-decade							0.037***	
							(0.007)	
Common first name country-gender-decade dummy								0.010***
								(0.002)
R-squared	0.16	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Observations	325462	325462	325462	325462	325462	325462	325462	325462

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. The measure of first name shares controlled for varies across columns. The measure in column (4) is that used throughout the paper. In all columns, the first name share measures the number of persons with the particular first name as a share of persons within the particular group. The first name dummy equals one for the top-10 names in the particular group. The groups are: Country of birth (column 1 and 2), country of birth and 10 years prior to birth (3 and 4), country of birth and gender (5 and 6), country of birth, gender, and 10 years prior to birth (7 and 8). All regressions include controls for the gender, number of occupations and 1x1 degree grid cell fixed effects, year fixed effects, and grid cell by birth year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Authors with a religious name are less likely to hold a scientific occupation, independent of the control for name commonness.

Table G8: Scientific occupations and religiosity of authors accounting for ethnicity

Dep var: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Ethnicity:		Germanic	Celtic	Italic	Slavic	Baltic	Basque	Uralic	Greek	Turkic	Semitic
Religious name	-0.0097***	-0.0056***	-0.0096***	-0.0052***	-0.0084***	-0.0100***	-0.0085***	-0.0089***	-0.0052***	-0.0098***	-0.0041***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)
Ethnicity		0.026***	0.0034	-0.023***	-0.0092**	0.0030	-0.0047	-0.0028	-0.015***	0.00080	-0.025***
		(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
R-squared	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
Observations	298430	298430	298430	298430	298430	298430	298430	298430	298430	298430	298430

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All baseline controls included throughout: op-10 most common names, gender, number occupations, and grid cell by year fixed effects. In addition, controls for the ethnicity of a name are added across the columns. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and scientific occupation holds when accounting for ethnicity.

Table G9: occupations and religiosity of authors - additional controls

Dep var: Scientific or engineering occupation	(1)	(2)	(3)	(4)	(5)	(6)
Religious name	-0.0099***	-0.0100***	-0.0099***	-0.0066***	-0.0098***	-0.0098***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
Birthplace latitude	0.026***					
	(0.007)					
Birthplace longitude		-0.0013				
		(0.005)				
Distance to Wittenberg (1000 km)			-0.23***			
			(0.065)			
Migration distance				0.0072***		
				(0.001)		
Urban					0.0095**	
					(0.004)	
R-squared	0.16	0.16	0.16	0.20	0.16	0.24
Observations	325462	325462	325462	179491	325462	286105
Grid x Year FE	Y	Y	Y	Y	Y	N
City x Year FE	N	N	N	N	N	Y

OLS across authors. The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure. All baseline controls included throughout: Top-10 most common names, gender, number occupations, and grid cell by year fixed effects, except that the fixed effects in column (5) are instead city by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses, except for column (6), where standard errors are clustered at the city level. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The association between religious names and occupation holds when adding controls for latitude, longitude, distance to Wittenberg, urban status, and city x year fixed effects.

Table G10: Religious names and occupations in science and social science

Dependent variable:	Science and engineer	Natural science	Engineer	Social science
	(1)	(2)	(3)	(4)
Religious name	-0.0100***	-0.0057***	-0.0044***	-0.00016
	(0.001)	(0.001)	(0.001)	(0.001)
R-squared	0.16	0.16	0.11	0.15
Observations	325462	325462	325462	325462
Mean dep var	0.18	0.16	0.025	0.070

The dependent variable is equal to one if one of the author's occupations was associated with natural science or engineering in column (1), corresponding to our main measure of science, natural science in column (2), engineering in column (3), and social science in column (4). All baseline controls included throughout: Top-10 most common names, gender, number occupations, grid cell by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Result: Authors with religious names have a lower likelihood of taking a job within natural sciences or engineering. The likelihood of taking jobs within social science is unaffected.

Table G11: Religious names and occupations - name-changes

Dependent variable: Scientific occupation	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Replaced name share:	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
Religious name	-0.0068***	-0.0054***	-0.0048***	-0.0049***	-0.0052***	-0.0036**	-0.0027*	-0.0028	-0.0023	-0.0026
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
R-squared	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.17
Observations	288138	288138	288138	288138	288138	288138	288138	288138	288138	288138
Mean dep var	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20

The dependent variable is a dummy equal to one if one of the author's occupations is associated with science or engineering. Religious name is a dummy equal to one if the name is shared by a significant saint or biblical figure, except that we re-code the variable across columns in the following way: In column (1), for a random 5% of the sample, we re-code the name as being religious, despite its true status, in column (2), 10% and so on. The sample excludes authors with at least one religious occupation. All baseline controls are included throughout: Top-10 most common names, gender, number occupations, and grid cell by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.

Result: Even in a sample without religious occupations, we can replace 50% of the non-religious names with religious names and retain the sign of the estimate. Significance remains until we replace 40% with religious names to remove significance.

H Additional material for the growth analysis

This appendix includes additional robustness checks for the main analysis of economic growth using authors in Section 5.1. Additional tables and figures for the growth analysis based on medieval university students are found in Appendix I.

H.1 Variable descriptions

All author level variables

To aggregate variables from the author level to city i, we calculate the average across all authors born within 100 km of the center of city i within the specified time intervals. If an author is born within 100 km of more than one city center, we attach him or her to the closest city. We set the particular variable to missing if less than 10 authors were born within 100 km of the city center in the particular time interval.

(log) City population size

The natural logarithm of one plus the number of people estimated to be living in a city in the period. The data is from Buringh (2021).

City population growth

The difference between the (log) city population size in period $t-1$ and period t.

Coast

A dummy equal to one if the city is located within 100 km of any ocean, zero otherwise.

Atlantic

A dummy equal to one if the city is located within 100 km of the Atlantic ocean, zero otherwise.

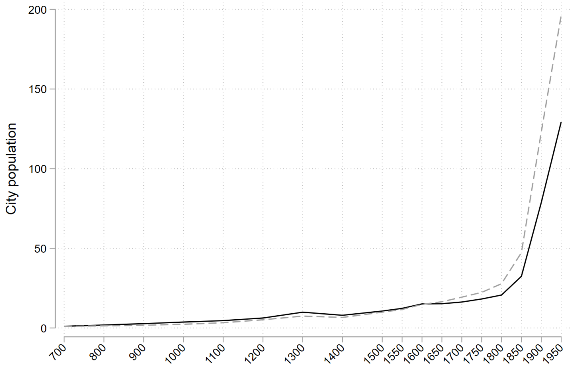
Earthquakes

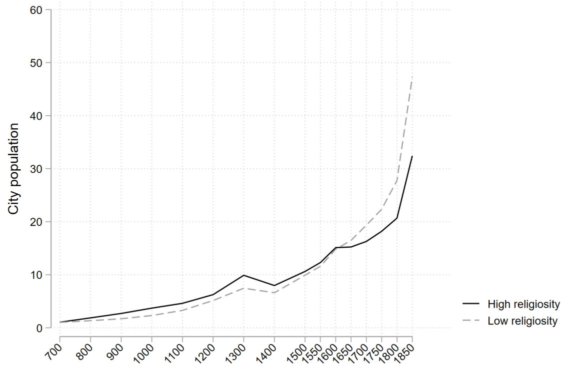
We calculate two type of earthquake exposure. One is author-specific and the other is city specific. For the author-specific measure, we compute a dummy equal to one if one or more authors born within 100 km of a city center was hit by an earthquake within x years prior to his/her birth. As in the previous analysis, earthquakes are defined as hitting a location if it hit within 100 km of that location. For the city-specific measure, we compute a dummy equal to one if an earthquake hit within 100 km of a city center within x years before the date of the population measure for that city.

H.2 Tables and figures

Figure H1: Simple relation between religiosity and modern growth 700-1950

(a) 700-1950 (b) 700-1850





Average city size over time. The cities were split in two by the median level of religiosity during the period 1300-1600.

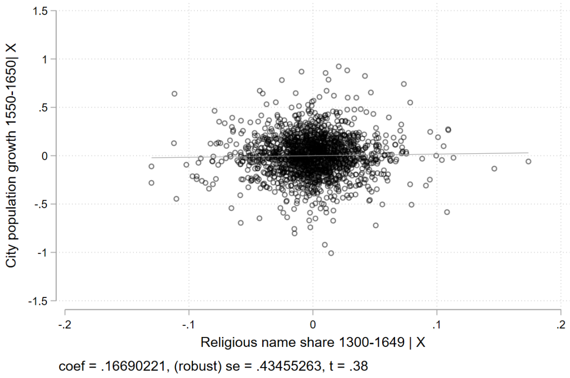
Table H1: Pre-1300 balancing tests

Controls:	(1) none	(2) basic	(3) +ini	(4) +year fe	(5) +5x5 grid fe
Dependent variable:					
(log) City population size	0.13***	0.11	0.0028	0.0059	0.013
	(0.034)	(0.070)	(0.011)	(0.012)	(0.012)
City population growth	0.0062	0.0056	0.0028	0.0059	0.013
	(0.011)	(0.011)	(0.011)	(0.012)	(0.012)
Share female 1300-1600	0.0019	0.0052	0.0053	0.0054	0.0032
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Share common names 1300-1600	0.034	0.036	0.030	0.028	0.015
	(0.023)	(0.022)	(0.022)	(0.021)	(0.016)
Coast	-0.074*	-0.088**	-0.090**	-0.091**	-0.031
	(0.043)	(0.042)	(0.042)	(0.042)	(0.042)
Atlantic	-0.010	-0.012	-0.014	-0.015	-0.00073
	(0.010)	(0.010)	(0.010)	(0.010)	(0.007)
Protestant	-0.16***	-0.14***	-0.11***	-0.11***	-0.014
	(0.049)	(0.046)	(0.041)	(0.041)	(0.038)
Latitude	-1.83***	-1.64***	-1.43***	-1.35***	0.12
	(0.387)	(0.372)	(0.321)	(0.315)	(0.130)
Longitude	0.76	1.03*	1.19**	1.24**	0.020
	(0.593)	(0.571)	(0.551)	(0.550)	(0.139)

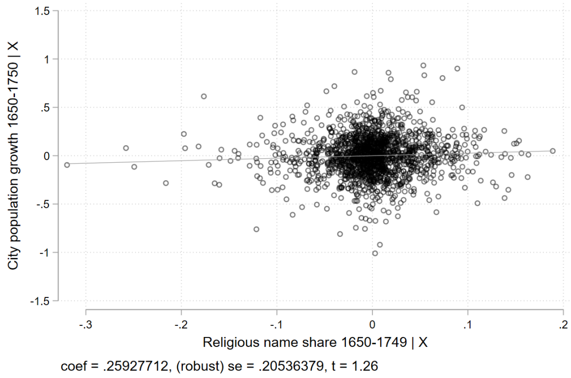
OLS estimates across city-centuries. The sample is restricted to the seven centuries between 700 and 1300 for the 417 cities where at least 10 authors were born between 1300 and 1600. The main explanatory variable is a dummy variable equal to one if the city had above median share authors with religious names in 1300-1600. The dependent variable is (log) city population size in row 1, city population growth (row 2), share women in the period 1300-1600 (row 3), share individuals with common names (row 4), an indicator of whether the city is located within 50 km of the coast (row 5), an indicator of whether the city is located within 50 km of the Atlantic coast (row 6), an indicator of whether the city had a Protestant majority (row 7), the latitude of the city center (row 7), and the longitude of the city center (row 9). The main explanatory variable is a dummy equal to one for the cities with above median religiosity in the period 1300-1600, zero otherwise. Each estimate represents one regression and thus reflects the difference in the dependent variable between cities with religiosity levels above and below what the median level was in years 1300-1600. The estimates in column (1) reveal the simple difference without controls. Controls included additively after that are the share of women and individuals with common names in column (2) (except that only common name share is included in row 3 and only the share of women in row 4), the initial (log) population size (column 3), century fixed effects (column 4), and 5x5 degree grid cell fixed effects (column 5). Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities that were more religious in 1300-1600 were larger prior to 1300. They were also further from the coast, less likely to be Protestant, and located further to the South. Adding initial level population size removes the size differences. Adding 5x5 grid cell fixed effects removes the geographic differences.

Figure H2: Religiosity and city growth

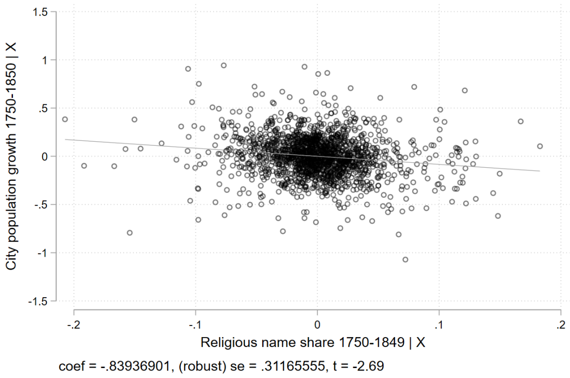
(a) 1300-1649, 435 cities



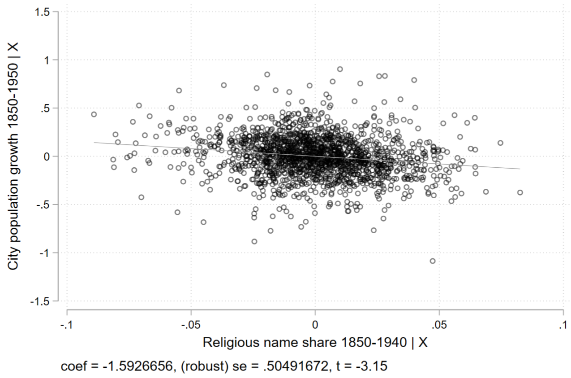
(b) 1650-1749, 435 cities



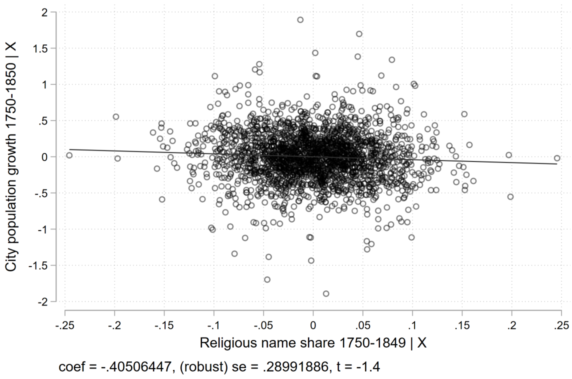
(c) 1750-1849, 435 cities



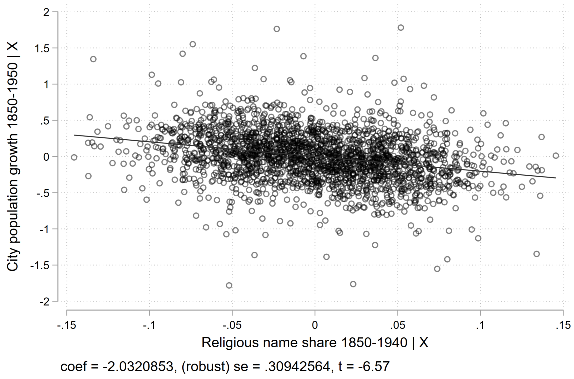
(d) 1850-1940, 435 cities



(e) 1750-1849, 1104 cities



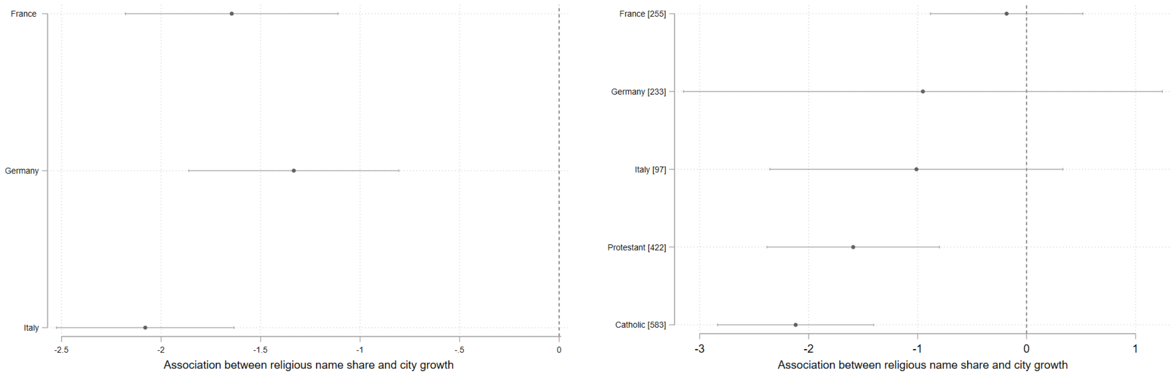
(f) 1850-1940, 1104 cities



Added variables plots corresponding to column (7) of Table 9 in panels a-d and corresponding to column (9) of Table 9 in panels e-f.

Figure H3: Religious names and city growth - heterogeneity across countries and denominations

(a) Excluding countries (b) Restricting to countries or denominations



The dots reflect the point estimate of β_{1950} from Equation (3) in regressions corresponding to column (9) of Table 9, where we exclude each of the three most represented countries separately in panel a and restrict the sample to each of the countries or Protestant vs Catholic areas in panel b. The numbers in squared brackets in panel b are the number of cities in each country or denomination. The 95% confidence bounds with standard errors clustered at the 1x1 grid cell level are shown throughout.
Result: The result is not specific to any of the most represented countries, nor to Protestant or Catholic areas.

Table H2: Religiosity and economic growth - sensitivity to levels or growth rates

Dependent variable: City population	growth (1)	level (2)	level (3)	growth (4)	level (5)	level (6)
(log) Initial population size	-1.19***	-0.19**		-0.88***	0.12**	
	(0.075)	(0.075)		(0.046)	(0.046)	
Religious name 1300-1649 x 1650	0.17	0.17	0.18			
	(0.354)	(0.354)	(0.359)			
Religious name 1650-1749 x 1750	0.26	0.26	0.27			
	(0.168)	(0.168)	(0.173)			
Religious name 1750-1849 x 1850	-0.84***	-0.84***	-0.88***	-0.42**	-0.42**	-0.31
	(0.254)	(0.254)	(0.260)	(0.205)	(0.205)	(0.202)
Religious name 1850-1940 x 1950	-1.59***	-1.59***	-1.39***	-2.04***	-2.04***	-2.10***
	(0.412)	(0.412)	(0.404)	(0.218)	(0.218)	(0.214)
R-squared	0.76	0.95	0.95	0.47	0.82	0.82
Observations	1740	1740	1740	2208	2208	2208
Mean dep var	0.55	2.89	2.89	0.86	3.23	3.23
City-trends	Y	Y	Y	N	N	N
Number cities	435	435	435	1104	1104	1104

OLS estimates across cities and time. The dependent variable is the growth of city population size in columns (1) and (4) and the (log) city population size in columns (2)-(3) and (5)-(6). The sample is the balanced sample of 437 cities for which at least 10 authors were born within 100 km of the city center in each of the four periods in columns (1)-(3) and the balanced sample of 1105 cities for which at least 10 authors were born within 100 km of the city center in the periods 1750-1849 and 1850-1940 in columns (4)-(6). The baseline controls are included throughout: The top-10 most common names share and female share, city fixed effects, time fixed effects, and columns (1)-(3) further include city-specific time-trends. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities that were more religious experienced lower growth rates from 1750 onward, independently of whether the dependent variable is the growth rate or the population level.

Table H3: Religiosity and economic growth - sensitivity to initial levels

Dependent variable: City population growth	(1)	(2)	(3)
Religious name 1300-1649 x 1650	0.40	-0.38**	
	(0.443)	(0.170)	
Religious name 1650-1749 x 1750	0.37	0.040	
	(0.237)	(0.194)	
Religious name 1750-1849 x 1850	-1.29***	-1.24***	-1.07***
	(0.309)	(0.209)	(0.226)
Religious name 1850-1940 x 1950	-1.41***	-1.75***	-1.48***
	(0.513)	(0.349)	(0.240)
R-squared	0.61	0.51	0.29
Observations	1740	1740	2208
Mean dep var	0.55	0.55	0.86
City-trends	Y	N	N

OLS estimates across cities and time. The dependent variable is the growth of city population size. The sample is the balanced sample of 437 cities for which at least 10 authors were born within 100 km of the city center in each of the four periods in columns (1)-(2) and the balanced sample of 1105 cities for which at least 10 authors were born within 100 km of the city center in the periods 1750-1849 and 1850-1940 in column (3). The baseline controls are included throughout: The top-10 most common names share and female share, city fixed effects, time fixed effects, and column (1) further include city-specific time-trends. In addition, the level of (log) city population in the first period interacted with time dummies is included throughout. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities that were more religious experienced lower growth rates from 1750 onward, also when the initial population accounted for is the (log) population size in the first period interacted with time-dummies.

Table H4: Religiosity and economic growth - time-varying impact of controls

Dependent variable: City population growth	(1)	(2)	(3)	(4)
(log) Initial population size	-1.22***		-0.90***	
	(0.074)		(0.046)	
Religious name 1300-1649 x 1650	0.25	0.059		
	(0.341)	(0.335)		
Religious name 1650-1749 x 1750	0.27	0.36**		
	(0.169)	(0.165)		
Religious name 1750-1849 x 1850	-0.75***	-0.59**	-0.43**	-0.43**
	(0.258)	(0.245)	(0.209)	(0.210)
Religious name 1850-1940 x 1950	-1.17***	-1.57***	-1.97***	-1.97***
	(0.451)	(0.483)	(0.226)	(0.226)
R-squared	0.76	0.79	0.47	0.47
Observations	1740	1740	2208	2208
Mean dep var	0.55	0.55	0.86	0.86
Common and female x time	Y	Y	Y	Y
Initial city size x time	N	Y	N	Y
Number cities	435	435	1104	1104

OLS estimates across cities and time. The dependent variable is the growth of city population size. The sample is the balanced sample of 435 cities for which at least 10 authors were born within 100 km of the city center in each of the four periods in columns (1)-(2) and the balanced sample of 1105 cities for which at least 10 authors were born within 100 km of the city center in the periods 1750-1849 and 1850-1940 in columns (3)-(4). City and time fixed effects are included throughout. In addition, initial (log) population size is included linearly in columns (1) and (3) and interacted with time in columns (2) and (4). The top-10 most common names share and female share are included throughout interacted with time. Columns (1)-(2) further include city-specific time-trends. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities that were more religious experienced lower growth rates from 1750 or 1850 onward, independently of whether or not the impact of initial levels, common name shares or female shares are allowed to vary over time.

Table H5: Religiosity and economic growth accounting for ethnicity

Dependent variable: City population growth										
Ethnicity:	Germanic	Celtic	Italic	Slavic	Baltic	Basque	Uralic	Greek	Turkic	Semitic
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: 1650-1950										
(log) Initial population size	-1.18***	-1.19***	-1.19***	-1.19***	-1.19***	-1.19***	-1.19***	-1.19***	-1.18***	-1.19***
	(0.076)	(0.075)	(0.076)	(0.076)	(0.075)	(0.075)	(0.076)	(0.075)	(0.075)	(0.076)
Religious name 1300-1649 x 1650	0.17	0.16	0.17	0.19	0.18	0.11	0.20	0.15	0.16	0.25
	(0.355)	(0.354)	(0.357)	(0.357)	(0.357)	(0.361)	(0.360)	(0.361)	(0.353)	(0.360)
Religious name 1650-1749 x 1750	0.28	0.25	0.26	0.27	0.27	0.16	0.30	0.23	0.23	0.41**
	(0.169)	(0.174)	(0.172)	(0.169)	(0.171)	(0.182)	(0.200)	(0.193)	(0.171)	(0.191)
Religious name 1750-1849 x 1850	-0.83***	-0.84***	-0.84***	-0.82***	-0.83***	-0.92***	-0.79***	-0.88***	-0.87***	-0.66**
	(0.253)	(0.257)	(0.254)	(0.265)	(0.262)	(0.259)	(0.293)	(0.286)	(0.264)	(0.277)
Religious name 1850-1940 x 1950	-1.54***	-1.59***	-1.58***	-1.56***	-1.58***	-1.69***	-1.55***	-1.62***	-1.61***	-1.49***
	(0.424)	(0.412)	(0.425)	(0.424)	(0.420)	(0.414)	(0.426)	(0.413)	(0.412)	(0.417)
Ethnicity	0.17	0.043	-0.023	-0.16	-0.059	0.29	-0.14	0.085	0.10	-0.48*
	(0.272)	(0.289)	(0.252)	(0.316)	(0.242)	(0.217)	(0.301)	(0.251)	(0.189)	(0.258)
R-squared	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76	0.76
Observations	1740	1740	1740	1740	1740	1740	1740	1740	1740	1740
Panel B: 1850-1950										
(log) Initial population size	-0.87***	-0.88***	-0.88***	-0.88***	-0.88***	-0.88***	-0.88***	-0.89***	-0.89***	-0.88***
	(0.046)	(0.045)	(0.046)	(0.046)	(0.046)	(0.046)	(0.045)	(0.046)	(0.045)	(0.046)
Religious name 1750-1849 x 1850	-0.37*	-0.41**	-0.35*	-0.43**	-0.42**	-0.47**	-0.55**	-0.36	-0.49**	-0.23
	(0.207)	(0.204)	(0.208)	(0.210)	(0.205)	(0.215)	(0.222)	(0.233)	(0.224)	(0.229)
Religious name 1850-1940 x 1950	-1.94***	-2.04***	-1.92***	-2.08***	-2.03***	-2.10***	-2.20***	-1.99***	-2.11***	-1.89***
	(0.227)	(0.215)	(0.230)	(0.243)	(0.235)	(0.231)	(0.235)	(0.239)	(0.229)	(0.232)
Ethnicity	0.63*	0.68**	-0.64**	0.14	-0.044	0.25	0.45	-0.19	0.35	-0.62**
	(0.330)	(0.280)	(0.318)	(0.363)	(0.301)	(0.284)	(0.281)	(0.279)	(0.354)	(0.288)
R-squared	0.47	0.48	0.47	0.47	0.47	0.47	0.47	0.47	0.47	0.47
Observations	2208	2208	2208	2208	2208	2208	2208	2208	2208	2208

OLS estimates across cities and time. The dependent variable is city population growth. Religious name share is the share of authors born within 100 km of the city center who share name with a religious figure. The sample is the balanced sample of 435 cities for which at least 10 authors were born within 100 km of the city center in each of the four periods in panel A and the balanced sample of 1104 cities for which at least 10 authors were born within 100 km of the city center in the periods 1750-1849 and 1850-1940 in panel B. Baseline controls included throughout: city and time fixed effects, top-10 common names share, and female share. Panel A includes also city-specific trends. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The negative impact of religiosity on city growth is not caused by differences in ethnicity.

Table H6: City growth and religiosity controlling for migration

Dependent variable: City population growth				
Authors attached to location of:	Birth		Death	
	(1)	(2)	(3)	(4)
(log) Initial population size	-0.56***	-0.54**	-0.91***	-0.91***
	(0.211)	(0.217)	(0.076)	(0.076)
Religious name 1750-1849 x 1850	1.22	1.07	0.025	0.026
	(0.939)	(0.985)	(0.292)	(0.292)
Religious name 1850-1940 x 1950	-1.55*	-1.43*	-0.95***	-0.95***
	(0.874)	(0.855)	(0.342)	(0.342)
Migrant share		0.84		0.0022
		(0.997)		(0.204)
R-squared	0.46	0.46	0.54	0.54
Observations	166	166	840	840
Mean dep var	0.80	0.80	0.98	0.98
Number cities	83	83	420	420

OLS estimates across cities. The dependent variable is population growth. Religious name share is the share of authors who share name with a religious figure and who were born within 100 km of the city center in columns (1)-(2) or died within 100 km of the city center in columns (3)-(4). Migrant share measures the share of authors who did not die in the same location as they were born, where the share is attached to their closest city of birth in columns (1)-(2) and to their closest city of death in columns (3)-(4). The sample is restricted to authors with information on location of birth and death. The sample further includes the balanced sample of cities for which at least 10 authors were born within 100 km of the city center in each of the periods 1750-1849 and 1850-1940 in columns (1)-(2). The sample in columns (3)-(4) is the balanced sample of cities for which at least 10 authors died within 100 km of the city center in each of the periods 1750-1849 and 1850-1940. Baseline controls included throughout: (log 1+) initial population size, city and time fixed effects, top-10 common names share, and female share. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities with more religious authors grew slower, accounting for migration.

Table H7: Religiosity and economic growth accounting for Atlantic trade

Dependent variable: City population growth						
Coast:	Any		Atlantic		Any	Atlantic
	(1)	(2)	(3)	(4)	(5)	(6)
(log) Initial city size	-0.43***	-1.18***	-0.42***	-1.17***	-0.88***	-0.86***
	(0.037)	(0.076)	(0.038)	(0.071)	(0.045)	(0.044)
Religious name 1300-1649 x 1650	-0.0079	0.18	-0.098	0.022		
	(0.200)	(0.359)	(0.195)	(0.359)		
Religious name 1650-1749 x 1750	0.067	0.24	0.18	0.25		
	(0.190)	(0.170)	(0.175)	(0.164)		
Religious name 1750-1849 x 1850	-1.32***	-0.84***	-1.21***	-0.71***	-0.37*	-0.25
	(0.194)	(0.256)	(0.183)	(0.245)	(0.206)	(0.207)
Religious name 1850-1940 x 1950	-1.95***	-1.59***	-1.84***	-1.35***	-1.97***	-1.84***
	(0.317)	(0.428)	(0.313)	(0.419)	(0.219)	(0.221)
Coast x 1750	-0.12***		0.11			
	(0.039)		(0.098)			
Coast x 1850	-0.080	0.040	0.25*	0.10		
	(0.050)	(0.059)	(0.143)	(0.154)		
Coast x 1950	-0.11	0.055	-0.22	-0.36	-0.11***	-0.35***
	(0.073)	(0.125)	(0.143)	(0.293)	(0.041)	(0.065)
R-squared	0.56	0.76	0.57	0.76	0.48	0.49
Observations	1740	1740	1740	1740	2208	2208
City-trends	N	Y	N	Y	N	N
Number cities	435	435	435	435	1104	1104

OLS estimates across cities and time. The dependent variable is city population growth. Religious name share is the share of authors born within 100 km of the city center who share name with a religious figure. The sample is the balanced sample of 435 cities for which at least 10 authors were born within 100 km of the city center in each of the four periods in columns (1)-(4) and the balanced sample of 1104 cities for which at least 10 authors were born within 100 km of the city center in the periods 1750-1849 and 1850-1940 in columns (5)-(6). Baseline controls included throughout: city and time fixed effects, top-10 common names share and female share. Columns (2) and (4) include also city-specific trends. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: The main result is unchanged after accounting for coastlines. Cities located along the coast apparently do not experience higher economic growth towards the end of the period, the reason being mainly that the largest cities in the sample are not located within 100 km of the Atlantic ocean (for instance Berlin, Paris, London, and Vienna).

Table H8: Occupations and economic growth

Dependent variable: City population growth						
Occupation:		Religious occupation			Scientific occupation	
	(1)	(2)	(3)	(4)	(5)	(6)
(log) Initial population size	-1.14***	-0.47***	-0.91***	-1.14***	-0.46***	-0.88***
	(0.077)	(0.051)	(0.056)	(0.078)	(0.050)	(0.056)
Occupation share 1300-1649 x 1650	-0.080	-0.44**		-0.38	-0.22	
	(0.291)	(0.185)		(0.580)	(0.347)	
Occupation share 1650-1749 x 1750	-0.21	-0.13		0.15	0.18	
	(0.158)	(0.149)		(0.252)	(0.335)	
Occupation share 1750-1849 x 1850	-0.33	-0.30	-0.13	0.59	0.52	0.49
	(0.438)	(0.299)	(0.426)	(0.425)	(0.359)	(0.354)
Occupation share 1850-1940 x 1950	-0.97	-2.57***	-2.00***	0.49	0.75*	0.93**
	(1.134)	(0.962)	(0.705)	(0.557)	(0.416)	(0.381)
R-squared	0.79	0.62	0.48	0.79	0.62	0.48
Observations	1164	1164	1260	1164	1164	1260
Mean dep var	0.55	0.55	0.96	0.55	0.55	0.96
City-trends	Y	N	N	Y	N	N
Number cities	291	291	630	291	291	630

OLS estimates across cities and time. The dependent variable is population growth. Occupation share is the share of authors born within 100 km of the city center who have a religious occupation in columns (1)-(3) or a scientific occupation in columns (4)-(6). The sample is the balanced sample of 293 cities for which at least 10 authors with information on occupation were born within 100 km of the city center in each of the four periods in columns (1)-(2) and (4)-(5) and the balanced sample of 635 cities for which at least 10 authors with information on occupation were born within 100 km of the city center in the periods 1750-1849 and 1850-1940 in columns (3) and (6). Baseline controls included throughout: (log 1+) initial population size, city and time fixed effects, top-10 common names share, and female share. In addition, all columns include controls for the average number of occupations interacted with the time. Columns (1) and (4) include also city-specific trends. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Cities with a larger share of individuals with scientific occupations experienced higher growth, while cities with more individuals in religious occupations grow slower. The effects are only significant at conventional levels in the specifications without city-specific trends.

Table H9: Innovation and religiosity

Dependent variable:	(log) Number inventors						City pop growth
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Religious name share	-3.95***	-1.19***	-1.19***	-1.01***	-0.88***		
	(0.406)	(0.288)	(0.289)	(0.315)	(0.319)		
(log) Population size	0.73***	0.74***	0.75***	0.79***	0.79***	0.80***	
	(0.051)	(0.030)	(0.030)	(0.030)	(0.030)	(0.031)	
Female share		2.85***	2.92***	2.58***	2.61***	2.87***	1.16***
		(0.574)	(0.576)	(0.614)	(0.608)	(0.608)	(0.347)
Common name share		-0.72***	-0.70***	-0.26	-0.35	-0.51*	-0.36**
		(0.226)	(0.226)	(0.282)	(0.287)	(0.285)	(0.150)
Coastline			-0.14**	-0.27***	-0.31***	-0.29***	0.018
			(0.071)	(0.074)	(0.075)	(0.074)	(0.045)
Atlantic coast			-0.060	-0.029	-0.040	-0.13	0.32*
			(0.258)	(0.264)	(0.271)	(0.300)	(0.178)
Protestant city				0.62***	0.53***	0.52***	-0.25***
				(0.097)	(0.098)	(0.099)	(0.062)
Latitude					0.10***	0.100***	-0.024*
					(0.025)	(0.024)	(0.014)
Longitude					-0.0083	-0.0075	0.031***
					(0.022)	(0.022)	(0.012)
Scientist share						1.21***	
						(0.385)	
(log) Number inventors							0.26***
							(0.018)
(log) Initial population size							-0.21***
							(0.035)
R-squared	0.22	0.79	0.79	0.78	0.78	0.78	0.38
Observations	1528	1523	1523	1304	1304	1304	1304
Mean dep var	3.89	3.89	3.89	4.25	4.25	4.25	1.00
5x5 grid FE	N	Y	Y	Y	Y	Y	Y
Number 5x5 grids	.	44	44	30	30	30	30

OLS estimates across cities. The dependent variable is (log 1+) Number of inventors within a radius of 100 km of a city registered in a patent office between 1850 and 1950 in columns (1)-(6) and the city population growth rate from 1850 to 1950 in column (7). The sample is includes all cities where at least 10 authors were born within 100 km of the city center between 1850 and 1950. Columns (1)-(5) probes the link between religiosity and innovators, while columns (6) and (7) are meant as consistency checks of the measure of innovators against the science share and city population growth. Robust standard errors clustered at the city level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level. **Result**: Cities with larger shares of religious names have more inventors registered in a patent office. At the same time, cities with larger shares of scientists have more innovators and cities with more innovators grow more.

I The RAG database of medieval university students

I.1 RAG Database

Our data on university students comes from the Repertorium Academemicum Germanicum (RAG) project, which records biographies of university scholars in the Holy Roman Empire. The RAG online database has information on 61,380 students, including information on their field and level of study, student names, and the year of birth and birthplace geocodes for 47,125 students. These data enable us to perform a series of additional robustness checks of our names-based religiosity measure.

The medieval university was - in most cases - divided into the three "higher faculties", law, medicine and theology as well as a faculty of arts. In order to study in one of the higher faculties, one had first to finish a master of arts. Although there was some variation between universities, the typical path of a student was first to finish a baccalaureate (or bachelor's) degree which typically took two years. After this, one could proceed with the master's degree which would typically take four additional years (Le Goff, 1993, 75-76). In those six years the student would be schooled in the seven liberal arts starting with the trivium (grammar, logic and rhetoric) and ending with the quadrivium (arithmetic, geometry, music, and astronomy). The student then had the option to continue studying medicine, law or theology.

It is our contention that the choice to study theology rather than law and medicine was partly driven by ideological and religious concerns. The first universities were founded at the turn of the 13th century as corporations of university students and teachers, and were thus part of the broader expansion of corporations across western Europe (Le Goff, 1993, 75-76). Despite being private partnerships, the universities were inherently Christian from the start, especially until the increasing regionalization of the 15th century under the patronage and influence of the pope (Le Goff, 1993, 138). The university was in other words closer to being an ally than a rival to the church and studying medicine or law was far from irreligious at the time. In fact, the studies included several religious ideas and practices, such as morning prayer (Friedman, 2021).

However, this did not mean that all studies were equal in the eyes of ecclesial authorities. Often the study of law, and to a less extent medicine, was seen with religious suspicion. Prestige in the middle ages was closely linked to piety and because theology had a plausible claim to be the religiously most important study, it was often considered to be "above" the other fields awarding them precedence in academic occupations and more prominent placement in royal ceremonies (Brundage, 2008, 274). However, the situation was different when it came to wordly matters. Lawyers tended to be promoted ahead of theologians in high positions in the state as well as in the church, something that was clear for both teachers and students. This meant that studying the law was much more popular than theology. This furthermore meant that professors of law, who were generally paid by their students, became much richer than professors of theology. The relative popularity of law (and to a lesser degree medicine) was lamented by many intellectuals at the time, perhaps most prominently by Dante. Others even argued that taking money to defend the guilty was sinful, comparing lawyers to prostitutes (Brundage, 2008, 274-275). The famous theologian Roger Bacon argued that "Everything in Civil Law has lay character. To address oneself to such a gross art is to leave the church" (Le Goff, 1993, 96). Even some lawyers argued that while the law had merits, it might not be good for the soul to pursue a career in law for the sake of money. Theologians thus complained that the fundamental reason they could not attract students was simple: greed (Brundage, 2008, 274-276).

The historical evidence thus seems consistent with the idea that since law (and medicine) was more lucrative, but at the same time was considered less pious and prestigious than theology, those students (or their parents) to whom piety was a relatively greater concern were more likely to chose theology as their field of study. If this is true, it follows that if religious individuals are more likely to have religious names, then those with religious names should be more likely to study theology.

The RAG database encompasses students who either graduated as masters or at one of the higher faculties (law, theology or medicine) as well as all nobles who attended university, even if they did not finish their degree. The database thus does not include students at the faculty of arts who did not graduate (scholares simplices), nor graduates with lower degrees (baccalaurii artium).[89] Of those that took advanced degrees, 14% took a degree in medicine, 68% in law, and 23% in theology. These sum to more than 100% because some students graduated in multiple subjects. We exclude the nobles from our sample to better interpret results. Results are robust to including the nobles (see the working paper version of this paper where nobles were included throughout, Andersen & Bentzen (2022)).

The RAG database contains 543 distinct first names, which are rooted in 356 original name branches.

[89]There are 5,927 nobles in the main sample, whereof the vast majority (83%) did not proceed with advanced studies. The results are robust to excluding the nobles.

I.2 Variables in the RAG database

Advanced studies dummy

A dummy variable equal to one if the student finished advanced studies in either medicine, law or theology. The variable is zero if the student finished with a master in arts and thus did not proceed with advanced studies.

Female

Dummy variable equal to one if the person is recorded to be female and zero otherwise.

Latitude and longitude

Latitude refers to the latitude of the "place of origin" of each student. In some cases we have direct data on the birthplace of students, while in other cases we only have data on their "geographical origin", whether or not they where also born there. In the case where we have data on the birthplace we use this to calculate location. If not, we use the "geographical origin". For a subset we have data on the "birthplace" and "geographical origin" of the students. For these, the correlation between birthplace and geographical origin is 0.99.

Religious name

We define a dummy variable equal to one if the student's name branch is also the name branch of one of the major patron saints. As these data contain information on individuals before the Protestant Reformation, our preferred measure is based exclusively on the major saints. In this period, the saints-based measure is nearly identical to the measure that also includes biblical names (rho=0.998), which is expected as the renewed interest in other Biblical figures that came with the Protestant Reformation had not yet occurred. 70% of the students share first name with a major patron saint according to our definition.

Studying theology

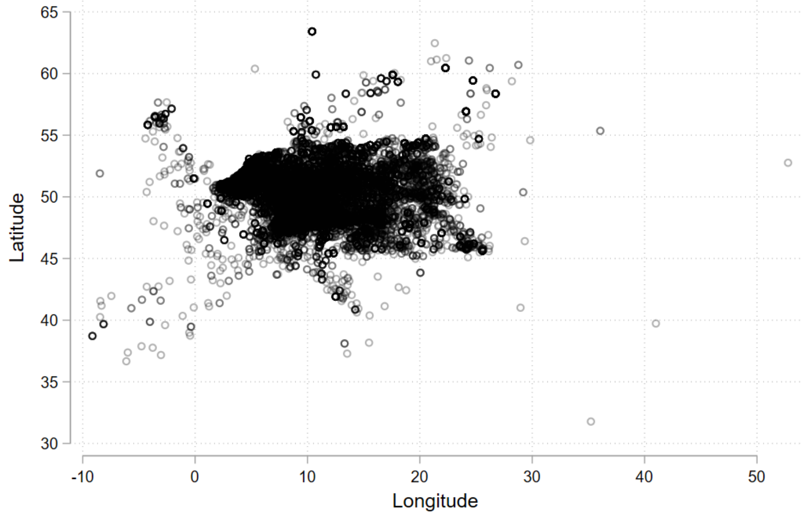
A dummy variable equal to one if the person is recorded to have finished an advanced degree in theology and zero otherwise. We define anyone with a theology degree as having studied theology, including those that also took other degrees. The variable is missing for students who did not finish their advanced studies.

Year

Year refers to the year the student started their studies. For a small portion of the students we do not have such data and instead calculate the likely year they started studies based on their birth year. We do this by calculating the median time from birth until commencement of studies in our sample, which is 17 years, adding it to the recorded birth year.

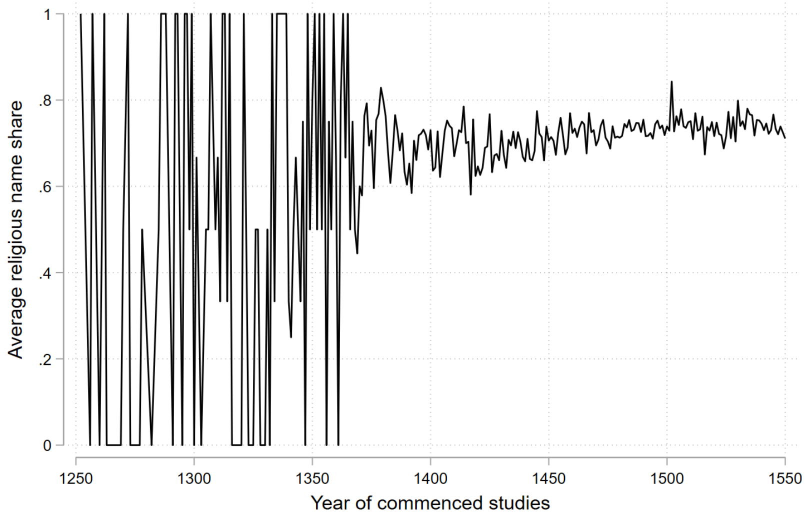
I.3 Tables and figures based on the RAG database

Figure I1: Birthplaces of the medieval university students



The map shows the spread of birthplaces of students in the RAG database throughout Europe.

Figure I2: Share of university students who share name with a prominent religious figure over time



Share of students with a name shared with a prominent religious figure throughout the sample period.

Table I1: Top-10 religious and non-religious name branches among university students

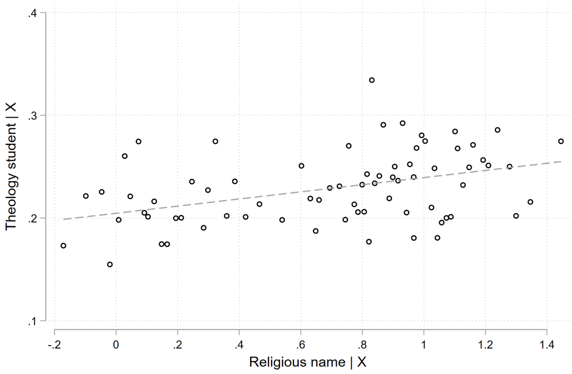
Religious names		Non-religious names	
Name	Share	Name	Share
John	0.24	Henry	0.05
Peter	0.04	William	0.02
Nicholas	0.04	Conrad	0.02
Jacob	0.04	Herman	0.01
George	0.03	Gerard	0.01
Yahweh	0.03	Arnold	0.01
Christos	0.02	Jasper	0.01
Andrew	0.02	Frederick	0.01
Matthew	0.02	Jerome	0.01
Martin	0.01	Hadrian	0.01
Remaining names	0.51	Remaining names	0.83

Table I2: Religious names and studying theology

Dependent variable: Theology student dummy	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	0.051***	0.052***	0.049***	0.026***	0.033***	0.034***	0.035***	0.046***
	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.007)	(0.009)	(0.008)
Female		-0.028	-0.024	-0.033	-0.022	-0.016	-0.022	-0.018
		(0.034)	(0.034)	(0.034)	(0.034)	(0.036)	(0.047)	(0.038)
Common name			0.015*	0.012*	-0.0058	-0.0057	0.0080	-0.0053
			(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Year					-0.0012***			
					(0.000)			
R-squared	0.0029	0.0029	0.0030	0.038	0.051	0.056	0.11	0.043
Observations	14276	14276	14276	14267	14267	14255	10474	14008
Mean dep var	0.25	0.25	0.25	0.25	0.25	0.25	0.23	0.25
Year FE	N	N	N	N	N	Y	Y	Y
1x1 grid FE	N	N	N	Y	Y	Y	Y	Y
1x1 grid x year FE	N	N	N	N	N	N	Y	N
2x2 grid x year FE	N	N	N	N	N	N	N	Y
Number 1x1 grids	190	190	190	181	181	181	129	183
Number 2x2 grids	11	11	11	11	11	11	9	9

OLS across students with an advanced degree. The dependent variable is a dummy equal to one if the student studied theology. Religious name is a dummy equal to one if the name is shared by a major saint. Controls include a dummy for top-10 most common names, gender, birth year fixed effects, 1x1 degree grid cell fixed effects, and 1x1 degree grid cell by birth year fixed effects, included consecutively until column (8). Instead of the latter, column (9) includes 2x2 grid cell by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Students who shared name with a major religious figure were around 15% of the mean more likely to study theology, on average.

Figure I3: Religious names and studying theology



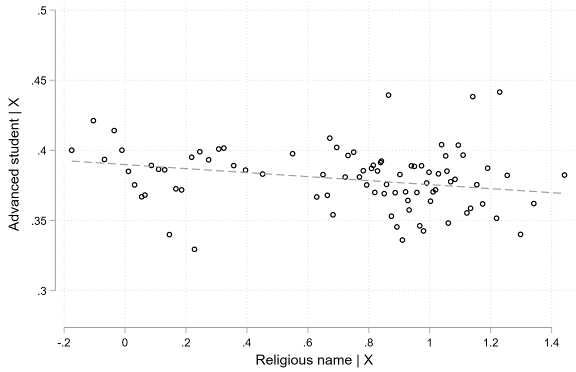
Binned added variables plot where 10,474 observations are binned into 100 equally sized bins. The line represents the OLS estimate of β in Equation (2), corresponding to the specification in column (8) of Table I3.

Table I3: Religious names and taking an advanced degree

Dependent variable: Advanced degree dummy	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Religious name	-0.045***	-0.043***	-0.041***	-0.020***	-0.015***	-0.016***	-0.014**	-0.032***
	(0.009)	(0.009)	(0.010)	(0.007)	(0.005)	(0.006)	(0.006)	(0.007)
Female		-0.11***	-0.11***	-0.091***	-0.085***	-0.078***	-0.074***	-0.088***
		(0.023)	(0.022)	(0.022)	(0.020)	(0.019)	(0.020)	(0.018)
Common name			-0.011	0.0020	-0.0075	-0.0073	-0.010*	-0.017***
			(0.009)	(0.008)	(0.005)	(0.005)	(0.006)	(0.005)
Year					-0.00058*			
					(0.000)			
R-squared	0.0017	0.0024	0.0025	0.049	0.052	0.077	0.17	0.066
Observations	38317	38317	38317	38316	38316	38296	32590	38012
Mean dep var	0.37	0.37	0.37	0.37	0.37	0.37	0.38	0.37
Year FE	N	N	N	N	N	Y	Y	Y
1x1 grid FE	N	N	N	Y	Y	Y	Y	Y
1x1 grid x year FE	N	N	N	N	N	N	Y	N
2x2 grid x year FE	N	N	N	N	N	N	N	Y
Number 1x1 grids	195	195	195	194	194	194	175	195
Number 2x2 grids	11	11	11	11	11	11	11	11

OLS across students. The dependent variable is a dummy equal to one if the student studied advanced studies. Religious name is a dummy equal to one if the name is shared by a major saint. Controls include a dummy for top-10 most common names, gender, birth year fixed effects, 1x1 degree grid cell fixed effects, and 1x1 degree grid cell by birth year fixed effects, included consecutively until column (8). Instead of the latter, column (9) includes 2x2 grid cell by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. +, *, **, and *** indicate significance at the 15%, 10%, 5%, and 1% level.
Result: Students who share name with a major religious figure were less likely to proceed with advanced studies in the Middle Ages.

Figure I4: Religiosity and taking an advanced degree



Binned added variables plot where observations (32,590 university students) are binned into 100 equally sized bins. The line represents the OLS estimate of the religious name dummy, including controls corresponding to column (8) of Table I3. **Result**: Students who shared name with a patron saint were less likely to proceed with advanced studies in the Middle Ages.

Table I4: Religious names and advanced degree for particular fields of study

Dependent variable:	Adv deg (1)	Law (2)	Medicine (3)	Theology (4)
Religious name	-0.026***	-0.041***	0.0034	0.0033
	(0.008)	(0.009)	(0.003)	(0.005)
R-squared	0.075	0.085	0.025	0.048
Observations	32547	27970	21534	22755

OLS across students. The dependent variable is a dummy equal to one if the student studied advanced studies in column (1), advanced law studies in column (2), advanced medical studies in column (3), and advanced theology studies in column (4). The sample is the full sample in column (1), the sample restricted to students who studied either advanced law studies or did not study advanced studies in column (2), students who studied advanced medicine or did not study advanced studies in column (3), and to students who studied advanced theology or did not study advanced studies in column (4). Religious name is a dummy equal to one if the name is shared by a major saint. Controls include a dummy for top-10 most common names, gender, and 2x2 grid cell by year fixed effects. Robust standard errors clustered at the 1x1 grid cell level in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% level.
Result: Students with religious names were less likely to proceed with advanced studies in law. The likelihood of proceeding with advanced studies in medicine or theology were no different from students without religious names in this sample including both students who continued with advanced studies and those who did not.