

DISCUSSION PAPER SERIES

DP15454
(v. 3)

LANGUAGE, KNOWLEDGE, AND GROWTH: EVIDENCE FROM EARLY MODERN EUROPE

Christine Binzel, Andreas Link and Rajesh
Ramachandran

MACROECONOMICS AND GROWTH

CEPR

LANGUAGE, KNOWLEDGE, AND GROWTH: EVIDENCE FROM EARLY MODERN EUROPE

Christine Binzel, Andreas Link and Rajesh Ramachandran

Discussion Paper DP15454
First Published 18 June 2021
This Revision 18 January 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programmes:

- Macroeconomics and Growth

Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Christine Binzel, Andreas Link and Rajesh Ramachandran

LANGUAGE, KNOWLEDGE, AND GROWTH: EVIDENCE FROM EARLY MODERN EUROPE

Abstract

This paper documents a language change in printing from Latin to the vernaculars, the spoken tongues, in the immediate aftermath of the Protestant Reformation of 1517. As a result, the share of vernacular titles in Europe rose from around 30% in 1500 to almost 60% in 1600. With the increased use of the vernaculars in printing, the availability of knowledge and ideas increased at the city level and became more diverse in terms of authors and themes. Finally, we study long-run consequences. Using linguistic differences across cities as a source of exogenous variation in the number of vernacular titles printed in cities, we document a positive effect of vernacular printing output on upper-tail human capital and city growth. This suggests that the turn to the vernaculars in printing was an important driver of European dynamism in the early modern period.

JEL Classification: E02, N13, Z12, Z13

Keywords: Inclusive institutions

Christine Binzel - christine.binzel@fau.de

Friedrich-Alexander-Universität Erlangen-Nürnberg and CEPR

Andreas Link - andreas.link@fau.de

Friedrich-Alexander-Universität Erlangen-Nürnberg

Rajesh Ramachandran - rajesh.ramachandran@awi.uni-heidelberg.de

Monash University Malaysia

Acknowledgements

We thank Luz Marina Arias, Stefan Bauernschuster, Christoph Becker, Sascha Becker, Mathias Bühler, Eltjo Buringh, Davide Cantoni, Jeremiah Dittmar, Dietmar Fehr, Karrar Hussain, David Laitin, Melissa Lee, Hannes Mueller, Markus Nagler, Jean-Philippe Platteau, Marteen Praak, Jared Rubin, Devesh Rustagi, Lucais Sewell, Gaspare Tortorici, Bas van Bavel, Christoph Vanberg, Fabian Wahl, Fabian Waldinger, Noam Yuchtman, and Jan Luiten van Zanden for valuable comments. Our paper has further benefited from helpful feedback and suggestions from audiences at the AALIMS-NYUAD Conference 2019, SDU/DIAS/CAGE/CEPR 2019 Conference, Barcelona Summer Forum 2019, ASREC Europe 2019, 2019 CEPR Conference "The Economics of Religion," #ASREC24h Conference 2020, the 2020 NEUDC Annual Conference, the Virtual Economic History Seminar (Spring 2021), the 28th CEPR European Summer Symposium in International Macroeconomics (ESSIM) 2021, the Workshop "Religion, Culture and Economic Growth in Historical Perspective" at UCLouvain in 2022, as well as seminar participants at the University of Cologne, Bayreuth University, Heidelberg University, University of Namur, Göttingen University, Passau University, Regensburg University, Utrecht University, and Ashoka University. We thank Nakul Heroor and Henrike Claes for excellent research assistance. All errors remain our own.

Language, Knowledge, and Growth: Evidence from Early Modern Europe

Christine Binzel, Andreas Link, and Rajesh Ramachandran*

January 2023

Abstract

This paper documents a language change in printing from Latin to the vernaculars, the spoken tongues, in the immediate aftermath of the Protestant Reformation of 1517. As a result, the share of vernacular titles in Europe rose from around 30% in 1500 to almost 60% in 1600. With the increased use of the vernaculars in printing, the availability of knowledge and ideas increased at the city level and became more diverse in terms of authors and themes. Finally, we study long-run consequences. Using linguistic differences across cities as a source of exogenous variation in the number of vernacular titles printed in cities, we document a positive effect of vernacular printing output on upper-tail human capital and city growth. This suggests that the turn to the vernaculars in printing was an important driver of European dynamism in the early modern period.

Keywords: Institutions, language, diglossia, upper-tail human capital, growth

JEL: E02, N13, Z12, Z13

* Binzel: University of Erlangen-Nuremberg (FAU), CEPR, and IZA, christine.binzel@fau.de; Link: University of Erlangen-Nuremberg (FAU), andreas.link@fau.de; Ramachandran: Monash University Malaysia, rajesh.ramachandran@monash.edu. We thank Luz Marina Arias, Stefan Bauernschuster, Christoph Becker, Sascha Becker, Mathias Bühler, Eltjo Buringh, Davide Cantoni, Jeremiah Dittmar, Dietmar Fehr, Karrar Hussain, David Laitin, Melissa Lee, Hannes Mueller, Markus Nagler, Jean-Philippe Platteau, Marteen Praak, Jared Rubin, Devesh Rustagi, Lucais Sewell, Gaspare Tortorici, Bas van Bavel, Christoph Vanberg, Fabian Wahl, Fabian Waldinger, Noam Yuchtman, and Jan Luiten van Zanden for valuable comments. Our paper has further benefited from helpful feedback and suggestions from audiences at the AALIMS-NYUAD Conference 2019, SDU/DIAS/CAGE/CEPR 2019 Conference, Barcelona Summer Forum 2019, ASREC Europe 2019, 2019 CEPR Conference “The Economics of Religion,” #ASREC24h Conference 2020, the 2020 NEUDC Annual Conference, the Virtual Economic History Seminar (Spring 2021), the 28th CEPR European Summer Symposium in International Macroeconomics (ESSIM) 2021, the Workshop “Religion, Culture and Economic Growth in Historical Perspective” at UCLouvain in 2022, as well as seminar participants at the University of Cologne, Bayreuth University, Heidelberg University, University of Namur, Göttingen University, Passau University, Regensburg University, Utrecht University, and Ashoka University. We thank Nakul Heroor and Henrike Claes for excellent research assistance. All errors remain our own.

1 Introduction

Recent research documents that a lack of equal opportunity leads to a misallocation of resources and is a major impediment to innovation and productivity (e.g. Bayer and Rouse, 2016; Aghion et al., 2018; Bell et al., 2019a; Hsieh et al., 2019). Nonetheless, direct evidence on the economic benefits of more broadly shared opportunity remains scarce. In this article, we study how the removal of language barriers impacted diversity in the marketplace of ideas and promoted economic growth. The context we consider is 16th century Europe, specifically the shift from Latin, a language not understood by large segments of society, to the vernaculars as the preferred language for printed works.

To this end, we make use of a unique database, the Universal Short Title Catalogue (USTC), a repository of all known books and pamphlets published in Europe in the first two centuries after the invention of the movable-type printing press in 1451. For the period from 1451 to 1600, more than 350,000 titles (or, works) – unique at the city-year level – are listed in the database.¹ Panel a. of Figure 1 shows the growth in the number of titles printed in Europe, from less than 2,000 titles printed per year at the beginning of the 16th century to nearly 6,000 titles printed per year at the end of the 16th century. Panel b. of Figure 1 documents that in the decades after the Reformation, the growth in the number of titles was largely due to an increase in works printed in the vernacular: the share of vernacular works increased from 30% of all titles at the beginning of the 16th century to almost 60% towards the end of the 16th century (corresponding to a sixfold increase in absolute levels), with a sharp increase after the Protestant Reformation of 1517.

We provide empirical evidence that the increased use of the vernaculars was accompanied by an increase in the availability and diversity of knowledge and ideas. Drawing on within-city differences in printing in the vernacular and in Latin, we document that following the Reformation, cities’ printing markets changed significantly for works printed in the vernacular compared with those printed in Latin: The number of authors from whom works became available at the city level increased, and a greater number of works from authors with a low socioeconomic background became accessible. We further show that at the city level, titles in the vernacular covered relatively more fields, thereby increasing the number of fields outside religion. In this way, vernacularization in printing likely contributed to the secularization of European societies after the Reformation (Cantoni et al., 2018). While some of the change we document was likely made possible by the broader societal and institutional changes that occurred in many European cities after the Reformation, we also provide evidence for a causal effect of language. Lastly, we turn to longer-run development. We document that an increase in the number of vernacular titles printed at the city level over the period 1518–1600 is associated with an increase in city population growth, a proxy for economic development. We also provide causal estimates by exploiting within-country linguistic

¹ A title shows up multiple times in our database if it is either printed in several cities or reprinted in the same city in another year. Data on the size of print runs is not available in the USTC.

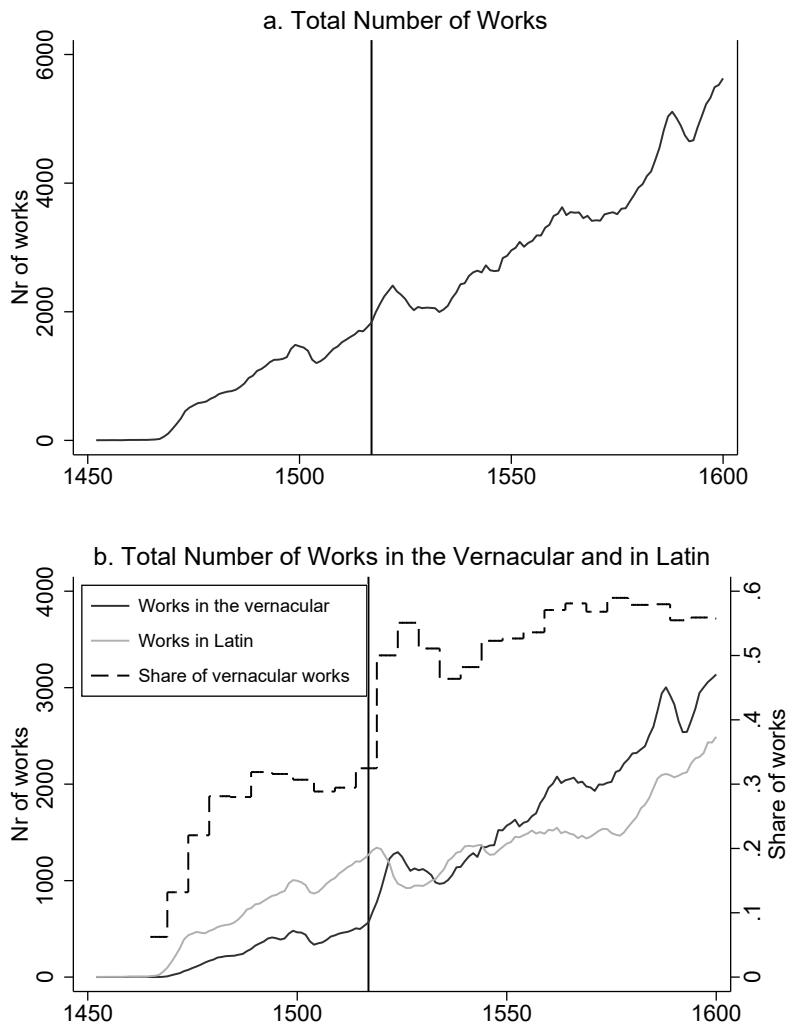


Figure 1: Language Change in Early Modern Europe

Notes: 5-year moving averages. Panel a. shows the total number of titles (unique at the city-year level) printed across cities over time, while Panel b. shows the number of titles printed in Latin and in the vernacular. The vertical line indicates the onset of the Protestant Reformation in 1517. Source: Universal Short Title Catalogue (USTC).

differences in cities' distance to Latin as a source of exogenous variation in the number of vernacular titles printed in cities.

At the beginning of the early modern period in Europe, Latin, while no longer used for everyday spoken communication, was still predominantly used not only in printing but also in written and formal contexts, e.g. in the domains of public administration, law, and politics. At the same time, the varieties of language spoken on a day-to-day basis had remained oral tongues whose standardization was incomplete, resulting in a linguistic situation also known as diglossia

(Ferguson, 1959; Versteegh, 2014). Diglossia has been a stable hallmark of many societies over long periods of time, both past and present. In addition to Latin in medieval and early modern Europe, prominent examples of diglossic situations include Sanskrit in ancient and medieval South Asia, and Standard Arabic in the Arabic-speaking world (Burke, 2004; Pollock, 2006; Versteegh, 2014). In this article, we focus on language change in the field of printing in 16th-century Europe, and its implications for the creation and consumption of knowledge and ideas.²

To study language change in printing, we start by aggregating printing output from the USTC over 10-year intervals, and compare the number of vernacular and Latin titles within cities over time in a difference-in-differences framework, controlling for city and decade fixed effects. We find that, following the Reformation, there was a sharp differential increase in vernacular titles relative to Latin titles, such that vernacular printing output started to dominate Latin printing output after 1517. This language change in printing was first driven by cities located in the Holy Roman Empire, where the Reformation originated. However, just a decade after the Reformation, a differential increase in vernacular titles was also observable in cities outside the Holy Roman Empire. There are several reasons why the Reformation had such rapid effects on language use in printing outside of cities that adopted the new religion or that were located in the Holy Roman Empire or were German-speaking (for details see Section 2.3). First, the Catholic Church’s loss of monopoly power in the market for religion resulted in secular leaders being less reliant on the Church to secure their legitimacy (Ekelund et al., 2002; Rubin, 2017; Cantoni et al., 2018). While the Catholic Church had tried to preserve Latin for centuries, the increase in religious competition implied a loss of Church influence over language use (Graff, 1991; Burke, 2004). Hence, printing – both in terms of content and language – became increasingly determined by market forces. As Luther’s vernacular writings sold extremely well, printers learned that ordinary people were willing to pay for books and pamphlets, insofar as they were printed in the vernacular (Burke, 2004; Pettegree, 2010; Schilling, 2017). Second, the vernacular writings by the reformers and especially the vernacular translations of the Bible played an important role for the standardization of vernaculars not only in the Holy Roman Empire but also outside it (Burke, 2004). Finally, by using the vernacular to discuss religious matters and by translating the Bible into the vernacular, Luther and his followers helped elevate the status of the vernacular in society (Burke, 2004; Pollock, 2006; Versteegh, 2014). Thus, it seems that the conditions in the printing market changed profoundly in the immediate aftermath of the Reformation.

We then use the same difference-in-differences approach to show that the increased use of the vernacular in printing influenced the diversity of knowledge and ideas at the city level. First, following the Reformation, cities saw a differential increase in the number of authors from whom printed works in the vernacular became available at the city level. Thus, readers now had access to a greater variety of works and, consequently, to a broader range of knowledge and ideas. Second, titles

² In other domains, such as academia, Latin continued to dominate (see de la Croix et al., 2022).

available at the city level increasingly came from a more diverse group of authors. We focus on one particular dimension of diversity, namely authors' socioeconomic background.³ Our results show a differential increase in the number of vernacular works from authors with a low socioeconomic background after the Reformation. Third, to provide more direct evidence on changes in book content, we draw on USTC's subject classification information. We document a differential increase in the thematic coverage of vernacular titles at the city level. Previously almost non-existent fields such as agriculture, art and architecture, economics, military handbooks, and news books saw an increase in titles predominantly printed in the vernacular. Overall, vernacular titles outnumbered Latin titles in most fields, and few fields remained almost exclusively Latin.

To be sure, the documented changes that accompanied the increased use of the vernacular in printing were likely related in part to other societal and institutional changes that occurred after the onset of the Reformation. It should be noted, however, that the documented changes appeared immediately or very soon after the Reformation, while at least some of the broader societal and institutional changes took more time to materialize. For example, in a related work, we show that among cities located in the Holy Roman Empire, Protestant cities saw a differential increase in the establishment of German schools only several decades after the Reformation (Binzel et al., 2023). To corroborate a potential causal effect of language, we first examine whether our results remain robust to removing from our sample cities that saw major societal and institutional changes following the Reformation. We start by removing from our sample all printing cities that adopted Protestantism over the course of the 16th century. As Protestantism left its mark not only on cities that adopted the new religion, but also on cities that remained Catholic but were near a Protestant town (Cantoni et al., 2018; Binzel et al., 2023), in a next step we additionally drop Catholic cities located in the vicinity (within 100km) of a Protestant city. We thereby rule out that the effects we observe are largely driven by spillover effects occurring in this subset of Catholic cities. Overall, we find similar (albeit somewhat delayed) patterns after dropping these cities from our sample. That is, within a few decades after the Reformation, the documented changes in printing are observable even for Catholic printing cities located further away from Protestant cities.

Second, we provide more direct evidence on the role of language by exploiting linguistic variation across cities. We posit that the linguistic distance of a city's vernacular to Latin affected a city's ability to use the vernacular in printing. In particular, greater similarity to Latin implied lower costs for borrowing words and grammar, and thus for standardizing the vernacular, which took centuries among the various European ethnic groups (Sasaki, 2017). We therefore hypothesize that cities with a vernacular closer to Latin were better able to respond to the push towards the use of the vernacular in printing triggered by the Reformation (for details see Section 6.2). To test our hypothesis, we restrict our sample to cities located outside the Holy Roman Empire, the birthplace of the Reformation. We then compare changes in vernacular printing output across cities with a

³Note that in the 16th century, almost all authors were male.

low and high linguistic distance to Latin, controlling for city and decade fixed effects. Consistent with our hypothesis, we show that cities with a low linguistic distance to Latin saw a comparatively larger increase in vernacular printing output after (but not before) the Reformation. Moreover, in these cities the relative diversity of titles printed in the vernacular increased. These results hold when we allow geographic and historical controls to have time-varying effects on vernacular printing output. Hence, our city-level measure of linguistic distance to Latin is unlikely to proxy for other city-level factors that might have differentially influenced vernacular printing output over time. Moreover, the lack of a pre-trend supports our contention that the Reformation altered the conditions in the printing market, such that we observe differential changes by cities’ linguistic distance to Latin only after 1517.

Overall, our findings indicate that the increased use of the vernacular in printing profoundly altered the knowledge and ideas that became available to readers in early modern Europe. This transformation in printing likely impacted a variety of professions, including craftsmen and tradesmen. In the remainder of the paper, we therefore study the longer-run implications of this transformation. Our main outcome is city population growth, a proxy for economic development in pre-industrial times. We document that cities with a greater number of vernacular titles – and thus with a greater availability and diversity of knowledge and ideas – after the Reformation experienced stronger city-population growth in the 17th century. This finding holds when controlling for country fixed effects and a set of geographic and historical controls. To address potential concerns about the endogeneity of the number of vernacular titles printed in cities between 1518 and 1600, we instrument for this variable using a city’s linguistic distance to Latin. If taken at face value, our point estimate implies that a one standard deviation increase in the number of vernacular titles is associated with an increase in city growth by 17%, where a one standard deviation increase corresponds to moving from Bamberg with 100 vernacular works to Valencia with 905 vernacular works or from Rotterdam with 163 vernacular works to Milan with 1,436 vernacular works. Finally, we provide evidence for one particular channel through which the availability of more diverse knowledge and ideas influenced city growth: a rise in upper-tail human capital. Using the same specification as before, we show that cities with a higher number of vernacular titles over the period 1518–1600 had a greater number of births of famous individuals in the 17th century, a proxy for upper-tail human capital in this era. This result holds using both OLS and IV estimation.

Our paper contributes to several strands in the literature. First, our study relates to recent work on what Bell et al. (2019b) have coined “lost Einsteins.” If who becomes an inventor depends heavily on “nurture” (the environment) rather than “nature” (innate ability), societies may fail to make optimal use of available talent, resulting in the misallocation of resources and adverse economic effects (Aghion et al., 2018; Bell et al., 2019a; Agarwal and Gaule, 2020). Previous work has shown that improved representation of historically marginalized groups, such as women and ethnic minorities, improves the economic performance of firms (Carter et al., 2003; Roberson and

Park, 2007), financial markets (Levine et al., 2014), public bureaucracies (Rasul and Rogger, 2015) and the economics profession (Bayer and Rouse, 2016). Our work highlights the potential role of institutional barriers by showing that the switch from Latin to the vernaculars allowed broader segments of society to access available knowledge and to contribute to knowledge production. Our study is thus the first to analyze the effect of a major pre-industrial event on equality of opportunity. We also provide evidence regarding the economic benefits of more inclusive institutions by showing the impact on upper-tail human capital and city growth.

Second, our study contributes to the large literature that has tried to explain the economic rise of the West (e.g. McNeil, 1990; Mokyr, 2005; Acemoglu et al., 2005; Voigtländer and Voth, 2013; Stasavage, 2016; de la Croix et al., 2018; Schulz, 2022), including studies that have attempted to explain Europe’s dynamism relative to the Arab World (e.g. Kuran, 2011; Blaydes and Chaney, 2013; Chaney, 2016; Platteau, 2017; Rubin, 2017). This paper complements these studies while also identifying a previously unrecognized factor: the increased use of common tongues in the domain of printing in 16th-century Europe. We argue that the use of a well understood language in printing transformed the way knowledge and ideas were produced and disseminated. This had profound impacts on longer-run development. It is also worth highlighting that the emergence of the spoken tongues as national languages has been viewed as a key factor in the construction of “imagined communities” and nation-building in Europe (Anderson, 2006).

Third, within this broader literature, our paper is related to the literature that ascribes Europe’s rise to knowledge elites (e.g. Mokyr et al., 2002; Dittmar, 2011; Meisenzahl and Mokyr, 2012; Squicciarini and Voigtländer, 2015; Dittmar, 2019; de la Croix et al., 2022; Link, 2023). Our analysis on vernacular printing output and the birth of famous individuals expands upon the studies by Dittmar and Meisenzahl (2020) and Serafinelli and Tabellini (2022). Dittmar and Meisenzahl (2020) show how upper-tail human capital was impacted in Protestant cities by the adoption of church ordinances (*Kirchenordnungen*), which increased the provision of public goods such as schools. Serafinelli and Tabellini (2022) find evidence suggesting that a number of city-level institutions established in free cities, rather than local economic conditions, produced and attracted famous individuals. We in turn posit a complementary channel for the rise of knowledge creators in society: the use of a well understood language in writing and in the public domain, which reduces the cost of participating in social and economic life for the vast majority of the population. A recent work by Dittmar and Seabold (2022) also bears resemblance to our paper. However, Dittmar and Seabold (2022) are concerned with how the thematic contents of books impacted city growth and institutional change, rather than with the impacts exerted by printing language. In this way, our paper is also distinct from Becker et al. (2021), who study the effects of the Catholic Church’s censoring of books across Europe during the 16th century. Drawing on the USTC and information on thinkers and famous people, Becker et al. (2021) find that book censorship made cities less attractive to famous people and reduced city population growth.

Fourth, because the language change we exploit was a consequence of the Reformation, this paper also builds on and contributes to the literature that examines the consequences of the Protestant Reformation (see Becker et al., 2016, 2021, for reviews). Drawing on data from Germany, Cantoni et al. (2018) document a shift towards greater secularization in Protestant cities relative to Catholic ones in the aftermath of the Reformation, whereas Binzel et al. (2023) document changes in language use in printing and in education. Both studies highlight spillovers from Protestant cities to nearby Catholic ones. This paper provides evidence that the Reformation resulted in language change in printing in Europe, and that it thereby contributed to European dynamism in the early modern period.

Finally, this paper contributes to the economics of language literature (Ginsburgh and Weber, 2020). A recent literature examines the effect of specific “top-down” linguistic policies of using non-indigenous languages as the official language on human development outcomes, amongst others in the context of sub-Saharan Africa (e.g. Laitin and Ramachandran, 2016; Laitin et al., 2019; Laitin and Ramachandran, 2022). Our paper instead examines the effect of “bottom-up” efforts to promote the vernacular tongue. We also identify the effect of language change as a catalyst on enabling greater societal participation, an aspect that has received little attention to date.

The rest of this paper is structured as follows: In Section 2 we provide relevant historical background. In Section 3 we describe our main data sources. In Section 4 we document the rise of the vernaculars in printing while in Section 5 we examine how the increased use of the vernacular in printing influenced the knowledge and ideas that became available at the city level. In Section 6 we corroborate a potential causal effect of language. Finally, in Section 7 we examine the longer-run consequences of the increased use of the vernaculars on economic growth. We conclude in Section 8.

2 Historical Background

2.1 Diglossia in Latin Europe

The linguistic situation in medieval Europe can be referred to as diglossia, a term that designates the parallel existence, within a speech community, of a high variety (Latin) and a low variety (the vernaculars such as German, English, and French). The high variety is a codified language used for purposes such as public administration, formal education, law, and literature, whereas the low variety is used for everyday communication, including trade. The low variety is not yet standardized (Ferguson, 1959; Versteegh, 2014). Indeed, among the various European ethnic groups, language standardization took centuries, as the data from Sasaki (2017) reveal: from the first vernacular Bible (which was published in 1380 in Czech and in 1382 in English) to the first comprehensive dictionary (the first was published in English in 1755, and in German in 1852), more than 300

years passed on average.⁴ While some vernaculars – of which there were between 40 and 70 in early modern Europe (Burke, 2004) – developed into (national) languages, others remained oral tongues or even disappeared. One should also note that there existed also different varieties of Latin and, for example, notaries mixed Latin with vernacular terms (Burke, 2004).

In medieval Europe, instruction in Latin was confined to a few institutions such as monasteries and grammar schools (Graff, 1991). Accordingly, literacy in Latin was restricted to elites, or approximately 1% to 2% of the population (Houston, 2011). Also, few were able to comprehend spoken Latin. Representative data on literacy rates in the vernacular do not exist for this time period. However, there is evidence that literacy rates in the vernacular were generally much higher than literacy rates in Latin (see Appendix Section A for more details). At the same time, while literacy rates in the vernacular varied significantly across Europe, they were generally on the rise (Graff, 1991; Houston, 2014). This was likely due in part to the rise of the merchant classes, craftpersons, and artisans from the 11th and 12th centuries onward, who were typically literate in the vernaculars (e.g. Houston, 2011). Further developments in the 13th and 14th centuries increased the demand for vernacular texts, including the increased availability of affordable reading glasses, the emergence of commercial writing workshops, and the switch from parchment to paper (von Polenz, 2013). The use of relatively cheap paper created new opportunities for writing and reading in the vernacular. Paper, which unlike parchment could be produced and stored in large quantities, could now be used for various kinds of secular uses, including accounting, letters, and merchant records (von Polenz, 2013). Moreover, some states decided to use the vernacular in the domain of administration, for example the French royal chancery and the Spanish King Alfonso X in the 13th century. However, the switch from Latin to French in the domain of law occurred not before 1539, and “French” would still encompass different vernaculars (Burke, 2004). From the 14th century onwards, also Italian, Spanish, English, German, and Czech were increasingly used for these purposes (Burke, 2004).

2.2 The Printing Industry Prior to the Reformation

Before the introduction of the movable-type printing press, it was largely the task of monks to copy books manually. This was both time consuming and costly (Eisenstein, 1980). Moreover, it gave the Catholic Church significant power to control the spread of knowledge and ideas (Logan, 1986). For example, historical evidence suggests that the Catholic church significantly contributed to Europe’s loss of ancient Greek knowledge, partly because monks often did not copy works written in Greek (see Link, 2023).

The spread of the printing press fundamentally altered the book market. The movable-

⁴ Burke (2004, p. 81) notes with regard to the German language in early modern Europe: “Strange as it may now seem, German was a language relatively poor in abstractions in this period, creating an obstacle to certain kinds of translation. Although Leibnitz argued that German was an appropriate language for philosophy, it was only in the eighteenth century that his advice began to be taken seriously.”

type printing press was invented by Johannes Gutenberg in Mainz, Germany, in the mid-15th century. From Mainz, the new technology diffused across Europe in a concentric pattern (Dittmar, 2011). Importantly, with the introduction of the movable-type printing press, book prices fell dramatically, namely by two-thirds (Dittmar, 2011). Also, the printing industry was from its very start characterized by profit-orientation and competition (Dittmar, 2011; Dittmar and Seabold, 2022). According to Dittmar and Seabold (2022, p. 21), more than half of the cities that adopted the printing press by 1500 already had more than one printer, with 31% having three or more.

As a result of the “decline of the clerical quasi-monopoly of knowledge” (Burke, 2004, p. 62), vernaculars started to be used more frequently and attempts were made to standardize them. Yet as shown in Figure 1, at the beginning of the 16th century, only about 30% of works were printed in the vernacular, reflecting the still dominant position of Latin at that time.

2.3 The Reformation and the Rise of the Vernaculars

In 1517, Martin Luther started to publicly criticize the practices of the Catholic Church.⁵ While his 95 Theses were written in Latin, as was common at the time, he soon decided to involve the wider public in the religious debate, and thus published the vast majority of his writings in German (Edwards, 2004; Pettegree, 2010). According to data from Edwards (2004, p. 501), some 89% of Luther’s works printed by 1523 were in German. His writings were extremely popular, so printers were keen on printing his works (e.g. Pettegree, 2010; Schilling, 2017). Martin Luther’s ideas and writings soon spread to neighboring countries and all of Europe (Graff, 1991). Luther’s new understanding of the Bible also implied that the individual believer should be able to directly access the word of God. Even more, he saw it as an obligation of the individual to read the Bible. Accordingly, Luther translated the New Testament in 1522, and a translation of both the Old and the New Testament appeared in 1534. With his writings and translation of the Bible, Luther profoundly shaped the German language (e.g. Sonderegger, 1998; von Polenz, 2000; Burke, 2004; Pettegree, 2015; Schilling, 2017).⁶

Drawing on the USTC, in a related work, we document that printing cities located in the Holy Roman Empire saw a sharp rise in vernacular works immediately after the Reformation, not only in cities that would adopt the new religion over the course of the 16th century but also in those that remained Catholic, albeit with some delay (Binzel et al., 2023). Moreover, and importantly, this was not confined to religious writings. Non-religious works in the vernacular also became widely available in the immediate aftermath of the Reformation. The spillovers on Catholic printing cities and on printing on non-religious themes suggest that the general conditions in the printing market changed. Consistent with the idea that the Reformation implied the Catholic Church’s loss of monopoly power in the market of religion (Ekelund et al., 2002; Rubin, 2017; Cantoni et al., 2018)

⁵ Detailed accounts of the origins of the Reformation are provided by Cantoni (2012) and Becker et al. (2021).

⁶ Nevertheless, Luther’s Bible was later translated into Low German because his original translation was less understood in the North (Burke, 2004).

and hence a loss of Church influence over language use (Graff, 1991; Burke, 2004), we document that, initially, religious works in the vernacular increased more in Catholic cities located close to a Protestant city. We also provide evidence on the importance of Luther in the development and spread of the German language. We show that among German-speaking Protestant cities, those that had a dialect closer to the one employed by Luther saw larger increases in vernacular printing output. This is consistent with the interpretation that Luther’s prolific production provided a corpus of material that could be employed especially by cities with a similar German dialect.

Similar mechanisms were likely at play at the European level. With the Catholic Church’s loss of influence after the Reformation, market forces increasingly determined the content and the language of printed works. With higher literacy rates in the vernacular, printers could increase profits by printing in the vernacular and by targeting ordinary people (Pettegree, 2010). In this vein, Burke (2004) notes that individual printers contributed to the standardization of vernaculars. Also, it is only after the Reformation, in the 1530s, that grammars of several European vernaculars were printed, including French, German, Italian, Spanish, Portuguese, and Czech (Burke, 2004). Further, though historical evidence has often highlighted the role of Luther in the standardization of the German dialects, vernacular translations of the Bible played an important role in the standardization process also outside the Holy Roman Empire (Burke, 2004). The reformers invested in promoting and/or creating a supra-regional or national vernacular when possible because the reformers wanted to achieve the widest possible reach. In turn, the influence of vernacular Bible translations after the Reformation (but not before) on the vernacularization process can be traced for many vernaculars, including Czech, Danish, and English. Of course, different versions of the vernaculars continued to exist in parallel, especially in places where printing was polycentric such as Germany and Italy but also, for example, in France (Burke, 2004).

Finally, Luther helped elevate the status of the vernacular in society. In Latin Europe, works written in the vernacular were often perceived as inferior and were thus accorded less esteem than those written in Latin (Ferguson, 1959; Burke, 2004; Versteegh, 2014). The reformers’ use of the vernacular for translating the Bible and for theological discussions legitimated the use of vernaculars to transmit written knowledge. Appendix Table B1 lists a selection of texts that sought to promote the various European vernaculars. All appeared well after the Reformation. The first is a text written by the Italian poet Dante in the early 14th century and printed in 1529. The next set of texts appeared only from the 1540s onwards and were written by preeminent intellectuals at the time, including many humanists, with the aim to advance the vernaculars as literary languages, asserting the vernaculars should be granted the same dignity and legitimacy as Latin. The dearth of “dignity” ascribed to the vernacular has not only been noted in the context of Latin Europe, but also with regard to Asia and the Arabic-speaking world (Pollock, 2006; Versteegh, 2014).

3 Main Data Sources

3.1 The Universal Short Title Catalogue (USTC)

The Universal Short Title Catalogue (USTC) is a repository that aims to include all printed works and pamphlets after the invention of the printing press up to – in its current version – 1650. It includes information on the printing place (city), language, subject classification, and author. The USTC is considered the most comprehensive European database for this time period and has been used in other studies, including Dittmar and Seabold (2022), Becker and Pascali (2019), and Becker et al. (2021).

Note that a particular work may appear multiple times in the USTC (and in our dataset), when a title was reprinted in another city and/or in a different year. Because of the high cost of transportation, print media usually stayed in the city where it was printed; instead of transporting books to faraway places, works were often reprinted (e.g. Febvre and Martin, 1976; Dittmar, 2011; Dittmar and Seabold, 2022). Thus, figures on printing output provide insight into the stock of knowledge and ideas that was available in a given printing city and year, as well as the influence of both supply and demand.

The dataset we retrieved in March 2019 contains a total of 354,354 entries (books and pamphlets) for the period 1451 to 1600.⁷ Breakdowns by country, language, and subject classification are provided in Appendix C.⁸ When classifying the language of a title, the USTC does not distinguish between different varieties of a vernacular. Instead, it uses today’s national European languages. Thus, for example, German dialects are simply classified as German. We refer to any work not printed in Latin as being printed in the vernacular.⁹ Panel b. of Figure 1 shows the total number of titles printed in the vernacular and in Latin over time across all printing cities. Since we identify titles at the city-year level, an increase in the total number of titles does not necessarily imply an increase in the total number of titles at the European level. However, it does imply an increase in the number of titles printed in cities, and, by extension, in the knowledge available to readers.¹⁰

Appendix Figures C1 and C2 disaggregate the data based on the major European language in the cities under consideration. For all language communities, except Spanish, we find that Latin works dominated vernacular works prior to the Reformation. Note that in Spanish-speaking cities, overall output was very low, with a total of approximately 20 works printed annually each in the vernacular and in Latin just before the Reformation. Immediately or soon after the Reformation,

⁷ We exclude works written in classical Greek (3,427 works).

⁸ For 15,109 observations (4.26%) we lack information about the printing city; for 5,511 observations (1.56%) the USTC does not provide information about the language in which the work was printed; for 96,578 observations (27.25%) the USTC lacks information about the author; and for 50,087 observations (14.13%) the USTC does not provide any subject classification.

⁹ We exclude works printed in Arabic or Hebrew.

¹⁰ Identifying the same title in different cities and over time is not straightforward. Vernaculars were not standardized yet, and titles were often long and not maintained consistently.

the share of vernacular works rises for all language communities, with the pattern being most pronounced for German and English.¹¹

We aggregate these book and pamphlet entries by city and year to obtain a city-year panel dataset with information on, amongst others, the number of works printed in the vernacular/Latin, the number of authors from whom works were available in the vernacular/Latin, and the number of subject classifications for vernacular/Latin works in a given city and year. We then limit our sample to cities with some printing output over the period 1478–1577, for which geographic location and additional city-level information is available from Rubin (2014). This results in a total sample of 323 printing cities (see Figure 2 for a map).

3.2 Further City-Level Variables

We draw on several sources to obtain additional city-level information.

Population data and information about famous people. As is common in historical work, we use city population figures from Bairoch et al. (1988). City population data for 1500 and 1600 is available for 232 out of the 323 printing cities. Data on notable, or famous, individuals comes from Laouenan et al. (2022). Their cross-verified database draws on several Wikipedia editions and Wikidata, including non-English editions of Wikipedia. It contains a total of 2.2 million entries, significantly more than, for example, Schich et al. (2014), who using Freebase provide birth and death data for 150,000 notable individuals.

Geographic and historical control variables. We primarily draw on available city-level information from Rubin (2014). This includes dummy variables for whether a city was Protestant in 1600, whether it had a university in 1450, and whether it hosted a bishop or archbishop before or in 1517. The dataset further provides a city’s geographic coordinates, and it includes information about the nation (as of 1500) to which a city belonged.

In addition to the variables from Rubin (2014), we geocode data from Ciolek (2004) to determine a city’s nearest distance to the main trade routes in Europe for the period 1300–1500. From Nüssli (2008) we take information on each city’s territorial affiliation in 1500, which we use, among others, to cluster standard errors. We further use data from Link (2023) to identify cities with a known presence of Greek migrants in the second half of the 15th century, and data from Büntgen et al. (2012) to obtain information about which cities were affected by plagues in the 16th century.

The summary statistics for the various city-level variables are shown in Appendix Tables D1 and D2.

¹¹ The decline in the share of Dutch and of German titles towards the end of the 16th century is attributable to a steep rise of academic dissertations in Latin (see Figure 5). Note that we restrict our empirical analysis to the year 1577.

4 The Rise of the Vernaculars in Printing in Early Modern Europe

We start by documenting the rise of the vernaculars over Latin in printing over the period 1478–1577 for the 323 printing cities in our sample. To do so, we first aggregate vernacular and Latin printing output over 10-year intervals, because annual printing output in cities was at times low, especially in the first few decades following the invention of the printing press. We also convert our data into long form, such that each city i appears twice in every decade t , once for works printed in Latin and once for works printed in the vernacular. We then estimate the following model:

$$Y_{i,t,l} = \alpha + \sum_{t=1478}^{1577} \beta_t (\text{Vernacular}_{i,t,l} * \text{decade}_t) + \gamma \text{Vernacular}_{i,t,l} + \delta_i + \phi_t + \epsilon_{i,t,l}, \quad (1)$$

where $Y_{i,t,l}$ is the natural logarithm of one plus the number of works printed in city i , decade t , and language l . *Vernacular* is an indicator variable equal to one for works printed in the vernacular, and zero for works printed in Latin. *decade* _{t} are dummy variables for each decade t – from 1478–1487 to 1568–1577 – with the omitted period being 1508–1517. δ_i are city fixed effects and ϕ_t are decade fixed effects. Standard errors are clustered at the territory \times decade level (with $N=52$ territories). Our coefficients of interest are the β_t s, which measure how, on average, vernacular printing output changes compared with Latin printing output *within* cities, relative to the omitted period (1508–1517).

We additionally estimate a difference-in-differences model where we aggregate printing output over three decades: 1478–1507 (pre-Reformation) as well as 1518–1547 and 1548–1577 (post-Reformation):

$$Y_{i,t,l} = \alpha + \sum_{h=1}^3 \beta_t (\text{Vernacular}_{i,t,l} * \text{period}_h) + \gamma \text{Vernacular}_{i,t,l} + \delta_i + \phi_t + \epsilon_{i,t,l}, \quad (2)$$

where *period* _{h} refers to 1478–1507 (1), 1518–1547 (2), and 1548–1577 (3).¹² All other variables are defined as before.

The results from estimating equations 1 and 2 are shown in Figure 3 and reported in column (1) of Appendix Tables E1 and E2. They show no differential trend in vernacular and Latin printing output within cities prior to the Reformation. Following the Reformation, however, there is a differential increase in vernacular compared with Latin works. This differential increase is not only statistically but also economically significant: in the first decade following the Reformation, vernacular works increase by 25% relative to Latin works, compared with 1508–1517 (column 1 of Appendix Table E1). This differential rise continues into the last two decades of the period under

¹²This is similar in spirit to the strategy employed by Cantoni et al. (2018).

study (1558–1577), when vernacular works increase by over 50% relative to Latin works. Further, the differential increase stems from an increase in vernacular works, rather than from a decline in Latin works. Latin works first remain at levels similar to those prior to the Reformation, and only begin to rise in the second half of the 16th century.

Appendix Figure E1 explores whether the differential rise in vernacular works is initially strongest in the Holy Roman Empire, the birthplace of the Reformation. We consider cities as being part of the Holy Roman Empire if the territory they are located in was part of the Holy Roman Empire at the beginning of our study period (the mid-15th century). At that time, the Holy Roman Empire included areas of today’s Germany, Austria, Belgium, Czech Republic, other parts of Central Europe, and Switzerland.¹³ The figure confirms that the rise of the vernaculars in printing occurred first in cities located in the Holy Roman Empire: in the first post-Reformation decade, a differential increase in vernacular works occurred only for cities located in the Holy Roman Empire. From then on, however, language change became a European-wide phenomenon with no statistically significant difference between cities in the Holy Roman Empire and those outside it.

Overall, the results imply that with the increased use of the vernacular in printing after the Reformation, the availability of knowledge and ideas increased at the city level.

5 Vernacularization in Printing and the Diversity of Printed Works

In this section, we first show that with the increase in printed titles in the vernacular after the Reformation, more diverse knowledge and ideas became available, measured as an increase in the number of authors from whom works were available at the city level (Section 5.1). Since market forces increasingly determined the production of books and pamphlets, it is likely that much of this increase was due to authors and printers responding to the fact that vernacular works were accessible to a broader section of the population, since the literacy rate in the vernaculars was higher than in Latin. In Section 5.2 we consider an additional mechanism, namely that the reduction in language barriers likely also allowed broader segments of society to engage in writing and printing activities. We consider one particular dimension of diversity, namely authors’ socioeconomic background. Finally, in Section 5.3 we provide more direct evidence that book diversity increased by drawing on USTC’s subject classification information.

5.1 Increase in the Number of Authors from Whom Works Became Available

We first examine whether the increased use of the vernaculars was associated with an increase in the total number of authors from whom works became available at the city level. We think of the number of authors from whom works were available at the city level as a measure of the diversity

¹³ Our results are robust to the dropping of cities located in Switzerland, which broke away from the Holy Roman Empire in 1499.

of works and hence of the diversity of the knowledge and ideas that were available to a city in a given time period.

To this end, we reestimate equations 1 and 2 where the dependent variable $Y_{i,t,l}$ is now defined as the natural log of one plus the total number of authors from whom at least one title is available in city i , decade t , and language l (vernacular/Latin). Panel a. of Figure 4 plots the difference-in-differences estimates on the interactions between the dummy for vernacular printing output and decade fixed effects (for the regression results, see column (2) of Appendix Tables E1 and E2). Prior to the Reformation, we observe similar trends. Also, as expected, the number of authors from whom works were available in a given city is significantly higher for Latin works relative to vernacular works (column (2) of Appendix Table E1). Following the Reformation, we see a strong and statistically significant differential increase in the total number of authors from whom vernacular works were available. The estimates suggest a differential increase by 22% (30%) for the first (second) post-Reformation decade. This is driven by an increase in the number of authors at the city level for vernacular works; the number of authors for Latin works increases only with some delay. Consequently, by the end of the 16th century, the number of authors from whom works were available is significantly higher for vernacular works. Thus, the evidence presented in this section suggests that the use of the vernaculars increased the diversity of knowledge and ideas available in a city.

Appendix Figure E2 shows that the total number of authors (across all cities) did not increase in the first post-Reformation period. This suggests that the differential increase in the number of authors from whom vernacular works were available in a city was originally due to an increase in reprinted works rather than the entry of new authors.

5.2 Increase in the Number of Works from Authors with a Low Socioeconomic Background

We now examine whether books and pamphlets printed at the city level increasingly came from a more diverse group of authors. To this end, we collect background information on all prolific authors, drawing on Wikipedia and other online sources. We define prolific authors as those with a minimum of 50 works in the USTC.¹⁴ We do so successfully for 246 out of the 468 authors (53%), totaling 41,433 titles. We restrict our sample to publications from these authors and classify authors as coming from a high socioeconomic background if their family was noble; if they were literate in Latin; if their parents were university educated; or if their fathers were churchmen. All other authors are considered as coming from a low socioeconomic background. These include, for example,

¹⁴ Using this definition, we obtain a sample of 720 authors. We drop authors who lived before the period under consideration or for whom birth and death data are not available (mostly authors from antiquity). This leaves 468 authors, accounting for 25.09% of works where the author is known. Note that we exclude the works from Martin Luther in order to prevent his works from driving the results.

authors whose parents were farmers or craftsmen.¹⁵ Using this definition, 28% of authors have a low socioeconomic background. Considering all vernacular (Latin) works of these authors that were printed prior to the Reformation, 15% (16%) were written by authors with a low socioeconomic background. This suggests that it was generally difficult for such authors to enter the printing market. Following the Reformation, this share increases to 53% for vernacular works and to 42% for Latin works.

We then estimate equations 1 and 2, where $Y_{i,t,l}$ is now defined as the natural log of one plus the total number of works from authors with a low socioeconomic background in city i , decade t , and language l (vernacular/Latin). The difference-in-differences estimates are shown in Panel b. of Figure 4 and reported in column (3) of Appendix Tables E1 and E2. Before the Reformation, we again observe parallel trends. Following the Reformation, vernacular works from authors with a low socioeconomic background show a greater differential increase. For the first (second) post-Reformation decade, we observe a differential increase by 5% (9%). This rises further to over 15% for the last two decades of the period under study. From the second post-Reformation decade onwards, the differential rise is also statistically significant. Hence, merely within two decades after the Reformation, there is a sizable increase in works from authors with a low socioeconomic background. The fact that the differential increase becomes statistically significant only from the second post-Reformation period onward is consistent with our finding in the previous section, namely that initially the increase in the diversity of titles is attributable to an increase in reprinted works rather than the entry of new authors. Finally, the regression results show an increase in diversity for Latin works immediately after the Reformation, which suggests that the Reformation also reduced the barriers to entry in the market for Latin books.

5.3 Increase in the Number of Subject Classifications

In this section, we provide more direct evidence on changes in book content following the increased use of the vernacular. To this end, we draw on the subject classifications provided by the USTC. We reestimate equations 1 and 2, where $Y_{i,t,l}$ is now defined as the natural log of one plus the average number of subject classifications in city i and decade t for works printed in language l (vernacular/Latin).

Panel c. of Figure 4 plots the difference-in-differences estimates and column (4) of Appendix Tables E1 and E2 reports the regression results. Before the Reformation, we observe a similar trend in the thematic coverage of books written in Latin and in the vernacular, with Latin works covering more fields than vernacular works. This changes with the Reformation, after which we observe an immediate differential increase in the average number of subject classifications for vernacular works. In the first post-Reformation decade, the average number of subject classifications for vernacular

¹⁵ The results presented below are similar if we keep 64,690 titles from all 468 authors in the sample and consider authors with missing information as authors coming from a low socioeconomic background. Note that nearly all authors in the USTC are male.

relative to Latin works increased by 7%. Relative book diversity increased further over the next decades, such that in the last two decades, the average number of subject classifications was 18% higher for vernacular than for Latin works.

Did language change occur equally across all fields? To shed light on this question, Figures 5 and 6 show the evolution of the number of titles printed in the vernacular and in Latin for each subject classification.¹⁶ They indicate that certain fields, such as agriculture, art and architecture, economics, military handbooks, music, and news books, became meaningful categories only after the Protestant Reformation, with the majority of works printed in the vernacular. Furthermore, even in traditionally Latin fields, such as science and mathematics, astrology and cosmography, and philosophy and morality, in which printing in Latin by far outweighed printing in the vernacular prior to the Reformation, a significant share of works were soon printed in the vernacular. Finally, only a few fields remained almost exclusively Latin, such as academic dissertations and classical authors. This suggests that once the barriers to using the vernacular were reduced and there was room for language choice in printing, writers and printers perceived the use of the vernaculars as economically beneficial.

6 Corroborating a Causal Effect of Language

The preceding sections present evidence that after the Reformation and with the increased use of the vernacular in printing, a larger number of works became available at the city level, and printed works became more diverse. Although we consider this to be an important finding in itself, in this section we want to assess the extent to which the documented changes would have also materialized in the absence of other important societal and institutional changes that occurred in many cities in the aftermath of the Reformation. We conduct two exercises. We first examine whether our results persist when we drop an increasing subset of cities from our sample – cities that experienced institutional and broader societal changes following the Reformation (Section 6.1). We then provide direct evidence of the role of language by exploiting variation in cities’ linguistic distance to Latin (Section 6.2).

6.1 Dropping Subsets of Cities from the Sample

In this section we aim to shut down interactions between the language effect and the effect of other changes that occurred in cities in the aftermath of the Reformation. To this end, we identify subsets of cities in our sample that experienced institutional and broader societal changes following the Reformation. We then drop these cities and reestimate equations 1 and 2.

¹⁶There are in total 37 subject classifications, see Appendix Table C3. The classification “wedding pamphlets” has only two entries and is hence not shown.

We first drop all Protestant printing cities from our sample. The Reformation led to fundamental changes in the societal and religious order in cities that adopted the new religion. To rule out that these societal and institutional changes interacted with the language effect, we drop all printing cities from our sample that had adopted the new religion by 1600.¹⁷ This leaves us with 207 (Catholic) printing cities out of a total sample of 323 printing cities. We then additionally drop Catholic cities from our sample that are located in the vicinity of a Protestant city (irrespective of whether the city engaged in printing). This reduces our sample to 160 Catholic printing cities. For cities located in the Holy Roman Empire, Cantoni et al. (2018) and Binzel et al. (2023) provide evidence that the effects of the Reformation were not confined to Protestant cities but soon spilled over to neighboring Catholic ones. The explanation for these spillovers is that the Reformation increased competition in the market for religion, which also had implications for the political market, as it changed the bargaining power of the political elites vis-à-vis the religious elites (Ekelund et al., 2002; Rubin, 2017; Cantoni et al., 2018). By dropping these Catholic cities, we rule out that the documented changes are driven by this subset of Catholic cities. As variation in cities' religious denomination is largely present within the Holy Roman Empire, dropping Catholic cities located in close vicinity to a Protestant city means dropping almost all Catholic cities located in the Holy Roman Empire.

Figure 7 shows the differential change in vernacular printing output relative to Latin printing output over time and across the three samples: a. all cities (Protestant and Catholic), b. Catholic cities, and c. Catholic cities located far from Protestant cities. While, compared with our main results (Panel a.), the differential increase in vernacular printing output is smaller in the decades after the Reformation, the overall results are not only qualitatively but also quantitatively similar.

Figure 8 turns to our main outcomes of interest, the natural log of one plus the number of authors from whom at least one vernacular/Latin title is available in city i and decade t (Panel a.), the number of vernacular/Latin works from prolific authors with a low socioeconomic background in city i and decade t (Panel b.), and the average number of subject classifications for vernacular/Latin works in city i and decade t (Panel c.). Reassuringly, for both subsamples and all three outcomes, we find no pre-trend. Panel a. indicates that following the Reformation, the differential increase in the number of authors from whom works were available at the city level also holds in the two subsamples. The main difference is that early on, in the first two decades after the Reformation, the differential increase is less pronounced, suggesting that Protestant cities were the first to observe a differential increase in the number of authors. The estimates also become somewhat more noisy when dropping an increasing number of cities, though this is expected. Nonetheless, 30 years after the Reformation, even Catholic cities located further away from a Protestant city experienced a statistically and economically significant differential increase in the number of authors from whom vernacular titles were available.

¹⁷ Four cities adopted Protestantism but later converted back to Catholicism. We classify these cities as Catholic cities.

When considering the number of works from prolific authors with a low socioeconomic background (Panel b.), we find that once all Protestant printing cities are dropped from the sample, the differential increase observed in the first 30-year period becomes small and statistically insignificant. This suggests that the documented change in the diversity of works in terms of authors' socioeconomic background was initially limited to cities that embraced the new religion and experienced various institutional and broader social changes. Yet 40 to 50 years after the Reformation, even Catholic cities located further away from Protestant cities observe a significant differential increase in the number of works from authors with a low socioeconomic background.

Panel c. documents the results for our last outcome, the thematic coverage of books. Here, the pattern is again similar to the one observed for the full sample of cities. Only, for the first decade following the Reformation, the differential increase in the average number of subject classifications is smaller and statistically insignificant.

In summary, the results of this exercise support the idea that at least part of the documented change can be interpreted as being independent of other societal and institutional changes put in motion by the Reformation.

6.2 Heterogeneous Effects by Cities' Linguistic Distance to Latin

We now provide evidence that allows us to more directly capture the role of language. To this end, we drop cities located in the Holy Roman Empire, the birthplace of the Reformation. We then examine whether cities with a dialect closer to Latin were better able to take advantage of the exogenous push to use the vernacular in printing triggered by the Reformation, and whether in these cities, in turn, more varied knowledge and ideas became available. Put differently, by exploiting heterogeneity in cities' distance to Latin, we aim to create variation in the number of vernacular titles that is orthogonal to the variation that is the result of other changes following the Reformation.

A vernacular's proximity to Latin may affect the process of vernacularization in various ways. First, the closer a vernacular is to Latin, the lower the costs to learn Latin and hence the lower the incentive to vernacularize (Laitin and Ramachandran, 2016). Consistent with this, the first languages to be written down using the Greek or Latin alphabet were Gothic in the 5th century, Irish in the 6th century, German and English in the 8th century, and the Slavic languages in the middle of the 9th century. By contrast, the creation of alphabets for the Romance languages can be dated closer to the early/mid 12th century (Wright, 1982, 1991; Colin, 1999). However, once the vernacular is to be used in writing and printing, cities with a vernacular closer to Latin face lower costs for borrowing words and grammar and hence for standardizing the vernacular. Also recall from Section 2.1 that in Europe, the standardization of the different vernaculars took centuries. Hence, the cost of standardization, and thus the potential benefit of borrowing from Latin, was

sizable.¹⁸

As argued in Section 2.3 above, the use of the vernacular in printing was strongly stimulated as a result of the Reformation. The Reformation reduced the Catholic Church’s influence on language use in printing. At the same time, the reformers’ use of the vernacular for religious matters contributed to the standardization of the vernaculars and helped elevate their status in society. Their vernacular works sold well, creating new opportunities for profit in the printing industry. Thus, the Reformation likely increased writers’ and printers’ ability and willingness to use the vernacular in printing. We therefore hypothesize diverging trends in vernacular printing output by cities’ linguistic distance to Latin after, but not before, the Reformation.

6.2.1 Measuring a City’s Linguistic Distance to Latin

To create a city-level measure for linguistic distance to Latin in the 16th century, we make use of the book and pamphlet titles provided in the USTC. Following Tambovtsev (2007), we define linguistic distance in terms of phonological distance. Phonological distance is determined as follows: First, for each city i , we compute the frequencies of eight different types of consonants¹⁹ using all titles printed in city i in the vernacular between 1451 and 1600.²⁰ We also compute the frequencies for Latin works based on all texts printed between 1451 and 1600. For each city, we then determine its distance to Latin by computing the sum of the squared differences in relative frequencies between a city’s vernacular titles and all Latin titles over all eight types of consonants. For the subsample of cities located outside the Holy Roman Empire, we then classify cities into low linguistic distance cities if their linguistic distance to Latin is at the 50th percentile or below; all other cities are classified as high linguistic distance cities. Considering only cities outside the Holy Roman Empire means dropping the vast majority of German-speaking cities as only five cities outside the Holy Roman Empire were German-speaking.

6.2.2 Empirical Specification

To study changes in the number of vernacular titles over time by cities’ linguistic distance to Latin, we estimate the following difference-in-differences model:

¹⁸ That the cost of standardization was sizable is also supported by historical work and linguistic research, which documents the important role Luther’s writings played for the standardization of the German language, see Section 2.3.

¹⁹ These types are labials, forelinguals, mediolinguals, guttural, sonorant, occlusive non-sonorant, fricative non-sonorant, and voice non-sonorant consonants.

²⁰ We also consider publications post-1517 as only part of the cities were printing prior to the Reformation and printing output was typically low. Note that for the analysis, we will classify cities into low and high linguistic distance cities. To address the concern of potential misclassification of cities, we will examine the robustness of our results when dropping cities from the sample whose linguistic distance to Latin is close to the cutoff (see Section 6.2.4).

$$Y_{i,t} = \alpha + \sum_{t=1478}^{1577} \beta_t (LowLingDistancetoLatin_{i,t} * decade_t) + \phi_t + \delta_i + \epsilon_{i,t}, \quad (3)$$

where $Y_{i,t}$ is the natural log of one plus the number of vernacular works in city i and decade t and $LowLingDistancetoLatin$ is an indicator variable equal to one for cities with a low linguistic distance to Latin and 0 otherwise. As before, $decade_t$ are dummy variables for each decade t – from 1478–1487 to 1568–1577 – with the omitted period being 1508–1517. ϕ_t are decade fixed effects and δ_i are city fixed effects. Note that $LowLingDistancetoLatin$ does not enter the model as it is absorbed by the city fixed effects. Standard errors are clustered at the territory \times decade level (with $N=23$ territories). Our coefficients of interest are the β_t s, which measure how vernacular printing output changes over time for cities with a low linguistic distance to Latin compared with those with a high linguistic distance to Latin, relative to the omitted period (1508–1517).

We additionally aggregate the number of vernacular titles over three decades (1478–1507, 1518–1547, and 1548–1577) and estimate the following model:

$$Y_{i,t} = \alpha + \sum_{h=1}^3 \beta_t (LowLingDistancetoLatin_{i,t} * period_h) + \phi_t + \delta_i + \epsilon_{i,t}, \quad (4)$$

where $period_h$ refers to the periods 1478–1507 (1), 1518–1547 (2), and 1548–1577 (3), and all other variables are defined as before.

In order to interpret any differential increase post-Reformation as a causal effect of language requires that cities’ linguistic distance to Latin is not correlated with factors that may also influence cities’ vernacular printing output over time. We address this in two ways. First, we examine pre-trends. Second, we add interactions between geographic and historical controls and decade fixed effects to equations 3 and 4 (see Section 6.2.4). We thereby allow the influence of geographic and historical controls to vary over time in a flexible manner.

6.2.3 Results

Figure 9 shows the results from estimating equations 3 and 4 for cities located outside the Holy Roman Empire. Panel A shows the difference-in-differences estimates while Panel B shows the predicted values of the dependent variable for cities with a low and a high linguistic distance to Latin. The regression results are reported in column (1) of Appendix Tables E3 and E4. Prior to the Reformation, we observe parallel trends. Following the Reformation, cities with a low linguistic distance to Latin experienced a stronger increase in vernacular printing output compared with those with a high linguistic distance to Latin. The differential increase is economically and, from the second post-Reformation decade onwards, statistically significant. They suggest that in the last decade under consideration, vernacular printing output was 40% higher for low linguistic distance cities compared with high linguistic distance ones, relative to the omitted decade. Hence, a city’s

linguistic distance to Latin created persistent differences in the number of titles printed in cities.

We next examine changes in the knowledge and ideas available to cities. We reestimate equations 3 and 4 where the dependent variable now is (a) the natural log of one plus the total number of authors from whom printed works in the vernacular are available in city i and decade t , (b) the natural log of one plus the total number of vernacular works from prolific authors with a low socioeconomic background in city i and decade t , and (c) the natural log of one plus the average number of subject classifications for vernacular works in city i and decade t . Panel A of Figure 10 shows the difference-in-differences estimates while Panel B shows the predicted values of the dependent variables for cities with a low and high linguistic distance to Latin. The regression results are reported in columns (2) to (4) of Appendix Tables E3 and E4.

For all three outcomes, we find parallel trends prior to 1517. Following the Reformation, we observe divergent trends by a city’s linguistic distance to Latin. More precisely, cities with a low linguistic distance to Latin see a stronger increase in the number of authors from whom vernacular works were available (on average 37% for the last 30-year period), in the number of vernacular works from authors with a low socioeconomic background (on average 26% for the last 30-year period), and in the average number of subject classifications covered by vernacular works (on average 32% for the last 30-year period). These results thus suggest persistent heterogeneous effects on the diversity of works by a city’s linguistic distance to Latin.

6.2.4 Robustness

We first address the concern that the measurement of cities’ linguistic distance from Latin may be subject to some error, leading to misclassification in our *LowLingDistancetoLatin* variable. Under the assumption that any misclassification is concentrated among cities close to the cutoff, we drop cities whose linguistic distance to Latin is between the 40th and 60th percentiles. This implies dropping 34 cities from the sample. Appendix Tables E5 and E6 present the results. Overall, our results are robust to the dropping of cities close to the cutoff.

Another concern may be that the heterogeneous effects we document pick up other geographic or historical city-level differences. While the time-invariant influence of these factors on our outcomes of interest are captured by the inclusion of city fixed effects, their influence may have changed over time, following the Protestant Reformation. To address this concern, we add interactions between various geographic and historical variables and decade fixed effects to equations 3 and 4. Geographic characteristics include the natural log of one plus a city’s distance to a trade route, the natural log of one plus a city’s distance to the coast, as well as a city’s latitude and longitude. Historical characteristics include dummy variables for whether a city had a university in 1450, was Protestant in 1600, and hosted a bishop or archbishop before or in 1517. The results are shown in Appendix Tables E7 and E8. We find little evidence for pre-trends prior to the Reformation.²¹

²¹ Merely the interaction between *LowLingDistancetoLatin* and the decade 1488–1497 is statistically significant

Following the Reformation, low linguistic distance cities see a differential rise in vernacular printing output with magnitudes similar in size to our main estimates (see Appendix Tables E3 and E4).

Overall, the results presented in this section provide evidence for a causal effect of the use of the vernacular in printing on knowledge and ideas.

7 Printing in the Vernacular and Longer-Run Development

Up to this point, we have shown that the increased use of the vernaculars was associated with more and varied knowledge and ideas at the city level. This linguistic transformation in printing was likely to have impacted a variety of professions, including craftsmen and tradesmen. Indeed, recent work by Dittmar and Seabold (2022) shows that cities with a greater number of merchants’ manuals, which were typically written in the vernacular, experienced higher city growth. In the remainder of the paper, we therefore study the longer-run implications of this transformation of the printing market on economic growth. Additionally, we provide evidence for one particular channel through which the increased use of the vernaculars likely affected economic growth: an increase in upper-tail human capital.

7.1 Empirical Approach

To study the long-run implications of the increased use of the vernacular in printing on economic growth, we examine the relationship between the number of titles printed at the city level post-1517 (1518-1600) to city population growth in the 17th century. City growth has been shown to be a good proxy for economic development in Malthusian economies (e.g. De Long and Shleifer, 1993; Acemoglu et al., 2005; Cantoni, 2015). We consider the same sample of printing cities as before with the additional requirement that information on population size is available. Our sample therefore comprises a total of 232 cities.

We start by estimating the following model using an Ordinary Least Squares (OLS) regression:

$$Y_{i,1600-1700} = \alpha + \beta_1 Vern_{i,1518-1600} + X_i + \delta_i + \mu_i, \quad (5)$$

where $Y_{i,1600-1700}$ refers to the natural logarithm of population growth in city i over the 1600–1700 period. $Vern_i$ is the natural logarithm of one plus the total number of vernaculars works printed in city i between 1518 and 1600. X_i is a vector of control variables. We control for the natural logarithm of population in 1600 as well as for several geographic and historical controls. As geographic controls, we include the natural log of one plus a city’s distance to a trade route,

for the first two outcomes (columns 1 and 2 of Appendix Table E7). There is no pre-trend, however, for the decade just prior to the omitted period (1498–1507), and we observe no pre-trend when using the more aggregated difference-in-differences model (Appendix Table E8). Against the background that few cities were printing early on and printing output was low in the 15th century, the pre-trend for the decade 1488–1497 may be in part idiosyncrasy.

the natural log of one plus a city’s distance to the coast, as well as a city’s latitude and longitude. As historical controls, we include dummy variables for whether a city had a university in 1450, was Protestant in 1600, and hosted a bishop or archbishop before or in 1517. δ_i stands for nation fixed effects (as of 1500), and standard errors are clustered at the territory level.

Estimating equation 5 using OLS raises the concern that the coefficient estimate on vernacular printing output may be biased because we might omit variables, such as a vibrant merchant class, that are correlated with both a city’s vernacular printing output and city growth. We therefore alternatively use a city’s linguistic distance to Latin as an instrument for the number of vernacular titles printed over the period 1518–1600. In Section 6.2 above we showed that cities’ linguistic distance to Latin created persistent differences in cities’ vernacular printing output after the Reformation. While the analysis was restricted to cities located outside the Holy Roman Empire, we will show in the following that cities’ linguistic distance to Latin is a significant determinant of cities’ vernacular printing output in the full sample of printing cities.

For our IV estimation, we hence estimate the following first-stage equation:

$$Vern_{i,1518-1600} = \alpha + \beta_1 LingDistancetoLatin_i + X_i + \delta_i + \mu_i, \quad (6)$$

where $LingDistancetoLatin_i$ stands for the natural logarithm of a city’s linguistic distance to Latin, normalized to the range [0, 100] (for details on how we measure a city’s linguistic distance to Latin, see Section 6.2). The vector X_i comprises the same control variables as in equation 5.

Based on equation 6, Panel a. of Figure 11 shows the relationship between our instrument and cities’ vernacular printing output. We find a strong negative relationship between a city’s distance to Latin and its vernacular printing output. That is, the more distant a city’s vernacular is to Latin, the lower its vernacular printing output on average. This negative relationship also holds within the subsamples of cities with the same main language, as illustrated in Appendix Figure F1.

While we do not include cities’ Latin printing output as a control in equation 6, Panel b. of Figure 11 shows that the relationship observed in Panel a. is robust to additionally controlling for cities’ Latin printing output.

The exclusion restriction requires that a city’s linguistic distance to Latin affects city population growth only via a city’s vernacular printing output. One concern may be that our instrument affects our outcomes of interest also via a city’s Latin printing output. As a first step, Panel c. of Figure 11 shows the relationship between our instrument and cities’ Latin printing output when controlling for cities’ vernacular printing output. We find no relationship between these two variables. This may seem surprising at first. Yet in 16th-century Europe, Latin printing output was mainly determined by non-market forces, in contrast to vernacular printing output, as it largely depended on the demand from universities and religious entities (Chartier, 1989; Mathis and Mathis, 2015). This is consistent with the correlates of vernacular and Latin printing output we observe

in our data. Appendix Table F1 reports the results from the regressions underlying Panels b. and c. of Figure 11. Column (1) suggests that a city’s vernacular printing output was, among other things, determined by its population size and its distance to a trade route. By contrast, the presence of a university as well as the presence of a bishop were important determinants of cities’ Latin printing output (column 2). We will return to this in Section 7.3 when we discuss the validity of our instrumental variable strategy.

7.2 Vernacular Printing Output and City Growth

The results from estimating equation 5 using both OLS and IV are reported in columns (1) to (3) of Table 1. Column (1) controls for country fixed effects and city population in 1600 while columns (2) and (3) add geographic and historical controls, respectively. In Panel A, we find a statistically significant relationship between a city’s vernacular printing output and city growth. The estimate on vernacular printing output is 0.066, which suggests that a 10% increase in a city’s vernacular printing output was associated with a 0.7% increase in city growth in the subsequent century.

Panels B and C report IV estimation results using a city’s linguistic distance to Latin as an instrument for a city’s vernacular printing output between 1518 and 1600. The first stage results show that even when controlling for country fixed effects and geographic and historical conditions, a city’s linguistic distance to Latin is a strong predictor of vernacular printing output in this period with an F-statistic of 30 (column 3, Panel C). The coefficient estimate on vernacular printing output is now 0.129 (column 3 of Panel B). That is, a 10% increase in cities’ vernacular printing output resulted in 1.3% higher city growth. Alternatively, we may consider a one standard deviation increase in vernacular printing output, which would be equal to going from Bamberg with 100 works to Valencia with 905 works or from Rotterdam with 163 works to Milan with 1,436 works. If taken at face value, our point estimate implies that a one standard deviation increase in vernacular printing output is associated with an increase in city growth by 17%. This effect is similar to going from the 50th to the 75th percentile, which corresponds to going from Aix-en-Provence with 90 works to Toledo with 474 works.

7.3 Robustness Analysis

To assess the validity of our instrument and the robustness of our results, we conduct several exercises, the results of which are presented in Table 2.

First, we exclude from our analysis cities with low printing output – that is, cities with the bottom 25% of vernacular printing output – as measurement error in our instrument might be larger for cities with only few works (column 1). Second, we exclude cities with a presence of Greek migrants in the second half of the 15th century (column 2). These cities saw an increase in upper-tail human capital and city growth in the 16th century (Link, 2023). Also, they are concentrated in the south-eastern parts of Western Europe (Link, 2023), where cities also tend to have a lower

linguistic distance to Latin.²² Third, we exclude cities with at least one plague between 1500 and 1600 (column 3). Dittmar and Meisenzahl (2020) show that plague shocks induced a subset of Protestant cities in the Holy Roman Empire to adopt city laws, which resulted in greater public goods provision and, in turn, higher city growth (see Section 6.1). Finally, we drop cities located in the Holy Roman Empire, the birthplace of the Reformation, in order to rule out that other changes that took place in the Holy Roman Empire in the aftermath of the Reformation are driving our results. The results for these various exercises are reported in columns (1)–(4) and confirm the robustness of our main results.

Next, we address the concern that our instrument might be working via Latin printing output rather than solely via vernacular printing output. In column (5), we therefore additionally control for cities’ Latin printing output. In spite of the very high correlation between vernacular and Latin printing output (see Appendix Table F1), our estimate remains statistically significant, although the standard error increases substantially. In column (6), we alternatively use a city’s linguistic distance to Latin as an instrument for Latin printing output while controlling for vernacular printing output. As suggested by Panel c. of Figure 11, conditional on vernacular printing output, our instrument has no predictive power for Latin printing output. Hence, the F-statistic on the instrument is close to zero. Consistent with this, Latin printing output has no explanatory power for city growth.

Finally, in column (7) we conduct a placebo test. The idea is that variation in vernacular printing output induced by our instrument should not matter for economic development *before* the rise of the vernacular languages in printing. We therefore now consider city population growth in the 15th century as the dependent variable. The coefficient estimate on vernacular printing output is close to zero and statistically insignificant, suggesting that previous to the rise of the vernacular languages in printing, linguistic distance to Latin did not matter for economic development in European cities.

Overall, the results lend credibility to our identification strategy.

7.4 Vernacular Printing Output and Upper-Tail Human Capital

We now examine whether cities that saw larger increases in vernacular printing output after the Reformation subsequently experienced increased levels of upper-tail human capital. We follow the recent literature and use the number of births of famous individuals in a given city as proxy for its upper-tail human capital (De la Croix and Licandro, 2015; Squicciarini and Voigtländer, 2015; Serafinelli and Tabellini, 2022; Dittmar and Meisenzahl, 2020). Our data comes from Laouenan et al. (2022) (see Section 3). We are able to identify 8,811 notable individuals born in one of the cities in our sample in the 17th century.

We reestimate equation 5 where the dependent variable now is the natural logarithm of one

²² Spanish and Italian cities’ average linguistic distance to Latin is 0.44 and 0.46, respectively, while, for example, Dutch and German cities’ average linguistic distance to Latin is 0.98 and 1.28, respectively.

plus the total number of births of famous individuals in the 17th century. In all specifications, we additionally control for the natural logarithm of the number of famous births in the 15th century. Columns (4) to (6) of Table 1 show the results. As before, the first column (column 4) includes only basic controls, while further controls are added in columns (5) and (6). The OLS estimates reported in Panel A document a statistically significant relationship between a city’s vernacular printing output and future upper-tail human capital. According to the estimate in column (6), an increase in vernacular printing output by 10% was associated with an increase in upper-tail human capital by 1.3%.

Panel B reports the results when instrumenting cities’ vernacular printing output by cities’ linguistic distance to Latin. They suggest that, once again, our OLS estimates are biased downward. According to the IV estimate reported in column (6), a 10% increase in vernacular printing output raised the future births of famous individuals by 3.5%.

We run the same robustness exercises as those described in Section 7.3 above. The results are reported in Appendix Table F2. Overall, they support our instrumental variable strategy to identify the causal effect of cities’ vernacular printing output on upper-tail human capital.

8 Conclusion

Drawing on the Universal Short Title Catalogue, we first document that in early modern Europe, the main language of printing changed from Latin, a language not understood by large segments of society, to the vernaculars in the immediate aftermath of the Protestant Reformation. Exploiting within-city differences in printing in the vernacular and in Latin, we then show that with the increased use of the vernaculars in printing, both the availability and diversity of knowledge and ideas increased. The number of authors from whom works became available increased, more titles from authors with a low socioeconomic background became available, and titles were printed in a greater number of fields. It appears, therefore, that by spurring book production on non-religious subjects, vernacularization contributed to the secularization of European societies in the aftermath of the Protestant Reformation (compare Cantoni et al., 2018).

While the documented change may in part be related to broader societal and institutional changes that took place in many cities after the Reformation, we also provide evidence of a causal effect of language. First, we observe changes in printing even for the subset of Catholic cities located further away from Protestant cities, i.e. cities that did not see institutional and societal changes following the Reformation and that were less likely to experience immediate spillover effects from Protestant cities due to changes in the market for religion. Second, we exploit variation in cities’ linguistic distance to Latin and show that among cities located outside the Holy Roman Empire, those with a low linguistic distance to Latin saw a stronger increase in vernacular printing output and in the diversity of vernacular titles after the Reformation compared with those with a high linguistic distance to Latin.

Finally, we turn to longer-run consequences of the increased use of the vernaculars in writing and printing. We document that cities with a higher number of vernacular titles over the period 1518–1600 later saw significantly higher levels of city population growth. This relationship holds in an instrumental variable setup where we use cities’ linguistic distance to Latin as an instrument for the number of titles printed in cities. We also provide evidence for one particular mechanism through which more and varied knowledge and ideas affected city growth: a rise in upper-tail human capital. In this way, we argue that the change from Latin to the vernaculars as the principal language of writing and printing was an important driver of European dynamism in the early modern period.

In early modern Europe, Latin continued to play a role in certain domains, such as in academia (compare de la Croix et al., 2022) and in the context of formal institutions and administration. We hope that future research will study the determinants and consequences of diglossia in these domains.

To what extent are our findings generalizable to other contexts? In the Arabic-speaking world, for example, it is difficult to assess the potential costs of diglossia given that Arabic diglossia continues to exist. Similar to Latin, Standard Arabic – the language of the Quran – may have acted as a barrier to knowledge production and consumption (Maamouri, 1998). This would be consistent with the fact that book production has remained low; some estimates suggest that the entire Arab World produces fewer books than individual European countries, such as Belgium (The Economist, 2016). Books in certain fields, such as economics, but also children’s books, are available only in Standard Arabic. Moreover, foreign books are typically often sold in English or French rather than in Arabic (The Economist, 2016). The dearth of written material in spoken tongues likely represents a major barrier to knowledge acquisition, and might be an important reason why the Arabic-speaking world scores low in terms of technological innovation (UNDP, 2003; WIPO, 2016). However, we do not wish to suggest that this is solely attributable to diglossia. Further research into this and other contexts, both past and present, will hopefully provide new insights into the role of language barriers for economic development.

References

- Acemoglu, D., S. Johnson, and J. Robinson (2005). The Rise of Europe: Atlantic Trade, Institutional Change, and Economic Growth. *American Economic Review* 95(3), 546–579.
- Agarwal, R. and P. Gaule (2020). Invisible Geniuses: Could the Knowledge Frontier Advance Faster? *American Economic Review: Insights* 2(4), 409–24.
- Aghion, P., U. Akcigit, A. Hyytinen, and O. Toivanen (2018). The Social Origins and IQ of Inventors. *mimeo*.
- Anderson, B. (2006). *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso Books.
- Bairoch, P., J. Batou, and P. Chèvre (1988). *La Population des Villes Européennes, 800–1850*. Librairie Droz.
- Bayer, A. and C. E. Rouse (2016). Diversity in the Economics Profession: A New Attack on an Old Problem. *Journal of Economic Perspectives* 30(4), 221–42.
- Becker, S. O. and L. Pascali (2019). Religion, Division of Labor, and Conflict: Anti-Semitism in Germany over 600 Years. *American Economic Review* 109(5), 1764–1804.
- Becker, S. O., S. Pfaff, and J. Rubin (2016). Causes and Consequences of the Protestant Reformation. *Explorations in Economic History* 62, 1–25.
- Becker, S. O., F. J. Pino, and J. Vidal-Robert (2021). Freedom of the Press? Catholic Censorship during the Counter-Reformation. *CEPR Discussion Paper 16092*.
- Becker, S. O., J. Rubin, and L. Woessmann (2021). Religion in Economic History: A Survey. In A. Bisin and G. Federico (Eds.), *Handbook of Historical Economics*, pp. 585–639. New York: Elsevier.
- Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019a). Do Tax Cuts Produce More Einsteins? The Impacts of Financial Incentives vs. Exposure to Innovation on the Supply of Inventors. *Journal of the European Economic Association* 17(3), 651–677.
- Bell, A., R. Chetty, X. Jaravel, N. Petkova, and J. Van Reenen (2019b). Who Becomes an Inventor in America? The Importance of Exposure to Innovation. *Quarterly Journal of Economics* 134(2), 647–713.
- Binzel, C., A. Link, and R. Ramachandran (2023). The Protestant Reformation and Language Choice in the Holy Roman Empire. *CEPR Discussion Paper 17818*.
- Blaydes, L. and E. Chaney (2013). The Feudal Revolution and Europe’s Rise: Political Divergence of the Christian West and the Muslim World before 1500 CE. *American Political Science Review* 107(1), 16–34.
- Büntgen, U., C. Ginzler, J. Esper, W. Tegel, and A. J. McMichael (2012). Digitizing Historical Plague. *Clinical Infectious Diseases* 55(11), 1586–1588.
- Burke, P. (2004). *Languages and Communities in Early Modern Europe*. Cambridge University Press.
- Cantoni, D. (2012). Adopting a New Religion: The Case of Protestantism in 16th Century Germany. *The Economic Journal* 122(560), 502–531.
- Cantoni, D. (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, D., J. Dittmar, and N. Yuchtman (2018). Religious Competition and Reallocation: The Political Economy of Secularization in the Protestant Reformation. *Quarterly Journal of Economics* 133(4), 2037–2096.
- Carter, D. A., B. J. Simkins, and W. G. Simpson (2003). Corporate Governance, Board Diversity, and Firm Value. *Financial Review* 38(1), 33–53.
- Chaney, E. (2016). Religion and the Rise and Fall of Islamic Science. *mimeo*.
- Chartier, R. (1989). The Hanged Woman Miraculously Saved: An occasionnel. In *The Culture of Print Power and Uses of Print in Early Modern Europe*, pp. 59–91. Polity Press.
- Ciolek, T. M. (2004). Old World Trade Routes (owtrad) Project. Technical Report, Asia Pacific Research Online, Canberra: www.ciolek.com.
- Colin, S. (1999). The Vernacular. In D. Abulafia (Ed.), *The New Cambridge Medieval History*, pp. 71–83. Cambridge University Press.
- de la Croix, D., F. Docquier, A. Fabre, and R. Stelter (2022). The Academic Market and the Rise of Universities in Medieval and Early Modern Europe (1000-1800). *CEPR Discussion Paper 14509*.
- de la Croix, D., M. Doepke, and J. Mokyr (2018). Clans, Guilds, and Markets: Apprenticeship Institutions and Growth in the Pre-Industrial Economy. *Quarterly Journal of Economics* 133, 1–70.
- De la Croix, D. and O. Licandro (2015). The Longevity of Famous People from Hammurabi to Einstein. *Journal of Economic Growth* 20(3), 263–303.
- De Long, B. J. and A. Shleifer (1993). Political Losers as a Barrier to Economic Development. *Journal of Law & Economics* 36(2), 671–702.
- Dittmar, J. (2019). The Economic Origins of Modern Science: Technology, Institutions, and Markets. *mimeo*.
- Dittmar, J. and R. R. Meisenzahl (2020). Public Goods, Human Capital, and Economic Growth: Evidence from German History. *Review of Economic Studies* 87(2), 959–996.
- Dittmar, J. and S. Seabold (2022). New Media and Competition: Printing and Europe’s Transformation after Gutenberg. *mimeo*.
- Dittmar, J. E. (2011). Information Technology and Economic Change: The Impact of the Printing Press. *Quarterly Journal of Economics* 126(3), 1133–1172.
- Edwards, Mark U, J. (2004). *Printing, Propaganda, and Martin Luther*. Fortress Press.
- Eisenstein, E. L. (1980). *The Printing Press as an Agent of Change (Volumes 1 and 2 in One)* (First Paperback Edition ed.). Cambridge University Press.
- Ekelund, Jr, R. B., R. F. Hébert, and R. D. Tollison (2002). An Economic Analysis of the Protestant Reformation. *Journal of Political Economy* 110(3), 646–671.
- Febvre, L. and H.-J. Martin (1976). *The Coming of the Book: The Impact of Printing 1450-1800*. Verso Books.
- Ferguson, C. A. (1959). Diglossia. *Word* 15(2), 325–340.
- Ginsburgh, V. and S. Weber (2020). The Economics of Language. *Journal of Economic Literature* 58(2), 348–404.

- Graff, H. J. (1991). *The Legacies of Literacy: Continuities and Contradictions in Western Culture and Society*, Volume 598. Indiana University Press.
- Grendler, P. F. (1991). *Schooling in Renaissance Italy: Literacy and Learning, 1300-1600*. JHU Press.
- Houston, R. A. (2011). Literacy. *European History Online*, <http://iegeu.eu/en/threads/backgrounds/literacy>.
- Houston, R. A. (2014). *Literacy in Early Modern Europe: Culture and Education 1500-1800, Second Edition*. Routledge, London and New York.
- Hsieh, C.-T., E. Hurst, C. I. Jones, and P. J. Klenow (2019). The Allocation of Talent and U.S. Economic Growth. *Econometrica* 87(5), 1439–1474.
- Kadens, E. (2001). The Vernacular in a Latin World: Changing the Language of Record in Thirteenth-Century Flanders, PhD Dissertation.
- Kuran, T. (2011). *The Long Divergence: How Islamic Law Held Back the Middle East*. Princeton University Press.
- Laitin, D. D. and R. Ramachandran (2016). Language Policy and Human Development. *American Political Science Review* 110(3), 457–480.
- Laitin, D. D. and R. Ramachandran (2022). Linguistic Diversity, Official Language Choice and Human Capital. *Journal of Development Economics* 156, 102811.
- Laitin, D. D., R. Ramachandran, and S. L. Walter (2019). The Legacy of Colonial Language Policies and their Impact on Student Learning: Evidence from an Experimental Program in Cameroon. *Economic Development and Cultural Change* 68(1), 239–272.
- Laouenan, M., P. Bhargava, J.-B. Eyméoud, O. Gergaud, G. Plique, and E. Wasmer (2022). A Cross-verified Database of Notable People, 3500BC-2018AD, <https://medialab.github.io/bhht-datascape>. *Sci Data* 9(290), 1–19.
- Levine, S. S., E. P. Apfelbaum, M. Bernard, V. L. Bartelt, E. J. Zajac, and D. Stark (2014). Ethnic Diversity Deflates Price Bubbles. *Proceedings of the National Academy of Sciences* 111(52), 18524–18529.
- Link, A. (2023). The Fall of Constantinople and the Rise of the West. *mimeo*.
- Logan, R. K. (1986). *The Alphabet Effect: The Impact of the Phonetic Alphabet on the Development of Western Civilization*. William Morrow & Company.
- Maamouri, M. (1998). *Language Education and Human Development: Arabic Diglossia and Its Impact on the Quality of Education in the Arab Region*. ERIC.
- Marshall, R. K. (1999). *The Local Merchants of Prato: Small Entrepreneurs in the Late Medieval Economy*, Volume 113. Johns Hopkins University Press Baltimore, MD.
- Mathis, R. and M.-A. Mathis (2015). Books in Foreign Languages: Publishing in the Netherlands, 1500-1800. In *Specialist Markets in the Early Modern Book World*, Volume 40, pp. 294–311. Brill.
- McNeil, W. H. (1990). *The Rise of the West: A History of the Human Community*. University of Chicago Press.
- Meisenzahl, R. and J. Mokyr (2012). The Rate and Direction of Invention in the British Industrial Revolution: Incentives and Institutions. In J. Lerner and S. Stern (Eds.), *The Rate and Direction of Inventive Activity Revisited*, pp. 443–479. University of Chicago Press.

- Mokyr, J. (2005). The Intellectual Origins of Modern Economic Growth. *Journal of Economic History* 65(2), 285–351.
- Mokyr, J. et al. (2002). *The Gifts of Athena: Historical Origins of the Knowledge Economy*. Princeton University Press.
- Nüssli, C. (2008). Periodical Historical Atlas of Europe, www.euratlas.com.
- Pettegree, A. (2010). *The Book in the Renaissance*. Yale University Press, New Haven and London.
- Pettegree, A. (2015). *Brand Luther: 1517, Printing, and the Making of the Reformation*. Penguin, New York.
- Platteau, J.-P. (2017). *Islam Instrumentalised: Religion and Politics in Historical Perspective*. Cambridge University Press.
- Pollock, S. (2006). *The Language of the Gods in the World of Men: Sanskrit, Culture, and Power in Premodern India*. University of California Press.
- Rasul, I. and D. Rogger (2015). The Impact of Ethnic Diversity in Bureaucracies: Evidence from the Nigerian Civil Service. *American Economic Review* 105(5), 457–61.
- Roberson, Q. M. and H. J. Park (2007). Examining the Link between Diversity and Firm Performance: The Effects of Diversity Reputation and Leader Racial Diversity. *Group & Organization Management* 32(5), 548–568.
- Rubin, J. (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, J. (2017). *Rulers, Religion, and Riches: Why the West Got Rich and the Middle East Did Not*. Cambridge University Press.
- Sasaki, Y. (2017). Publishing Nations: Technology Acquisition and Language Standardization for European Ethnic Groups. *Journal of Economic History* 77(4), 1007–1047.
- Schich, M., C. Song, Y. Y. Ahn, A. Mirsky, M. Martino, A. B. Barabasi, and D. Helbing (2014). A Network Framework of Cultural History. *Science* 345(6196), 558–562.
- Schilling, H. (2017). *Martin Luther: Rebel in einer Zeit des Umbruchs*. C.H. Beck.
- Schulz, J. F. (2022). Kin Networks and Institutional Development. *Economic Journal* 132(647), 2578–2613.
- Serafinelli, M. and G. Tabellini (2022). Creativity over Time and Space: A Historical Analysis of European Cities. *Journal of Economic Growth* (27), 1–43.
- Sonderegger, S. (1998). Geschichte Deutschsprachiger Bibelübersetzungen in Grundzügen. In v. W. Besch, A. Betten, O. Reichmann, and S. Sonderegger (Eds.), *Sprachgeschichte. Ein Handbuch zur Geschichte der deutschen Sprache und ihrer Erforschung, Volume 1, 2nd Edition, Handbooks of Linguistics and Communication Science*. de Gruyter, Berlin/New York.
- Squicciarini, M. P. and N. Voigtländer (2015). Human Capital and Industrialization: Evidence from the Age of Enlightenment. *Quarterly Journal of Economics* 130(4), 1825–1883.
- Stasavage, D. (2016). Representation and Consent: Why They Arose in Europe and Not Elsewhere. *Annual Review of Political Science* 19(1), 145–162.

- Struppa, D. C. and A. Kruslin (2016). An Introduction to the Datini Archives: Early Capitalism in Renaissance Tuscany. *Journal of Business and Management* 22(2), 125–131.
- Tambovtsev, Y. (2007). How Can Typological Distances between Latin and Some Indo-European Language Taxa Improve Its Classification? *The Prague Bulletin of Mathematical Linguistics* (88), 73–90.
- The Economist (2016). *Arabic Publishing: Plus de kutub, please*. <https://www.economist.com/middle-east-and-africa/2016/06/16/plus-de-kutub-please>.
- UNDP (2003). Arab Human Development Report 2003: Building a Knowledge Society.
- Van Bavel, B. J. (2010). *Manors and Markets: Economy and Society in the Low Countries 500–1600*. Oxford University Press.
- Versteegh, K. (2014). *The Arabic Language, 2nd Edition*. Edingburgh University Press.
- Voigtländer, N. and H.-J. Voth (2013). Gifts of Mars: Warfare and Europe’s Early Rise to Riches. *Journal of Economic Perspectives* 27(4), 165–186.
- von Polenz, P. (2000). *Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart: Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart, Band I: Einführung, Grundbegriffe, 14.-16. Jahrhundert*. de Gruyter.
- von Polenz, P. (2013). *Deutsche Sprachgeschichte vom Spätmittelalter bis zur Gegenwart: 17. und 18. Jahrhundert*. de Gruyter.
- WIPO (2016). World Intellectual Property Organization (WIPO) Statistics Database.
- Wright, R. (1982). *Late Latin and Early Romance in Spain and Carolingian France*, Volume 8. Francis Cairns.
- Wright, R. (1991). *Latin and the Romance Languages in the Early Middle Ages*. Routledge London.

Figures

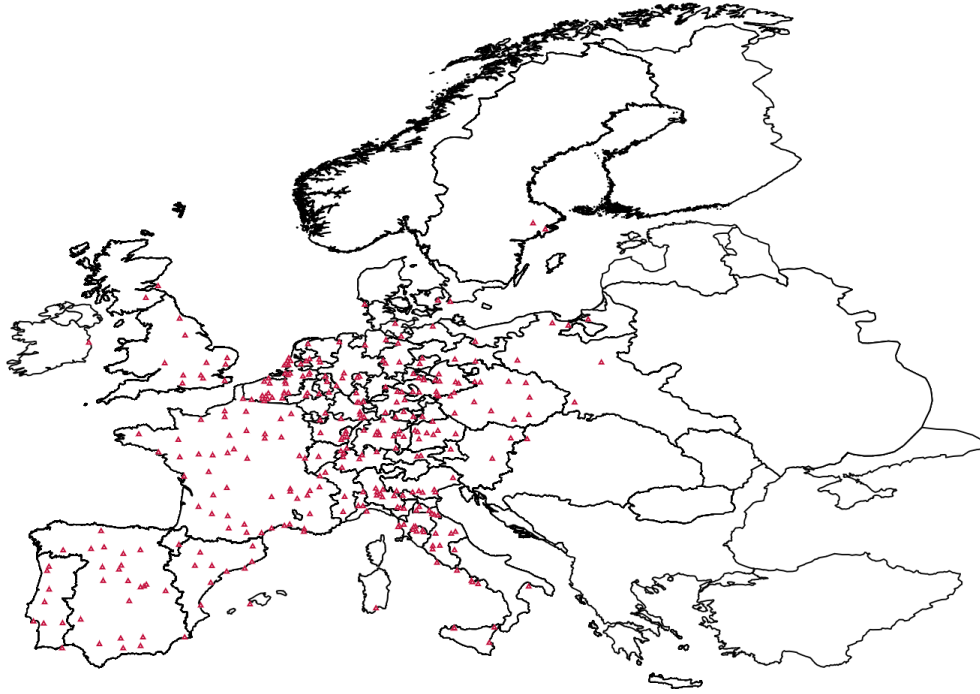


Figure 2: Printing Cities in Early Modern Europe

Notes: Cities with some printing output over the period 1451–1600 based on the Universal Short Title Catalogue (USTC).

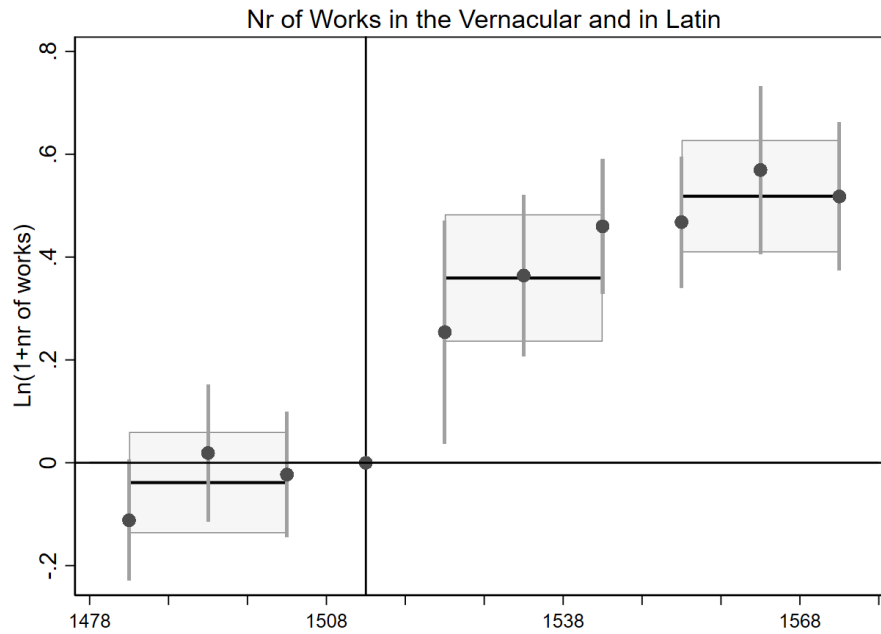


Figure 3: The Rise of the Vernaculars in Printing in Early Modern Europe: Within-City Differences in Titles Printed in the Vernacular and in Latin

Notes: The dependent variable is $\ln(1+\text{total number of vernacular/Latin works printed in city } i \text{ and decade } t)$. The figure shows the coefficient estimates on the interactions between a dummy for vernacular works and decade fixed effects as dots, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, with their 90% confidence intervals indicated as boxes. The regression includes city fixed effects and is restricted to cities with some printing output over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517. The regressions results are reported in column (1) of Appendix Table E1.

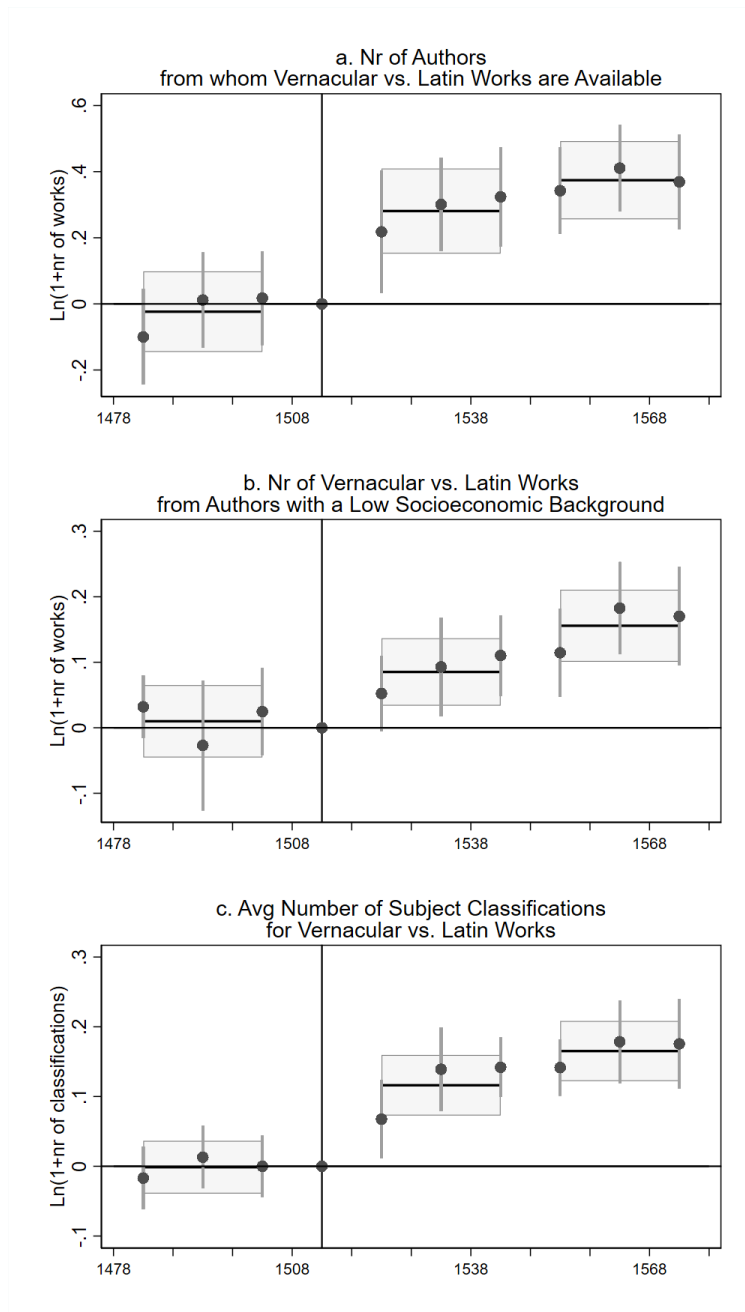


Figure 4: Increase in the Diversity of Printed Works: Within-City Differences

Notes: The dependent variable is $\ln(1+\text{total number of authors from whom vernacular/Latin works were available in city } i \text{ and decade } t)$ in Panel a., $\ln(1+\text{total number of vernacular/Latin works from authors with 50 works or more with a low socioeconomic background in city } i \text{ and decade } t)$ in Panel b., and $\ln(1+\text{average number of subject classifications for vernacular/Latin works in city } i \text{ and decade } t)$ in Panel c. All three panels show the difference-in-differences estimates on the interactions between a dummy for vernacular and decade fixed effects as dots, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, with their 90% confidence intervals indicated as boxes. All regressions include city fixed effects and are restricted to cities with some printing output over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517. The regression results are reported in columns (2)–(4) of Appendix Table E1.

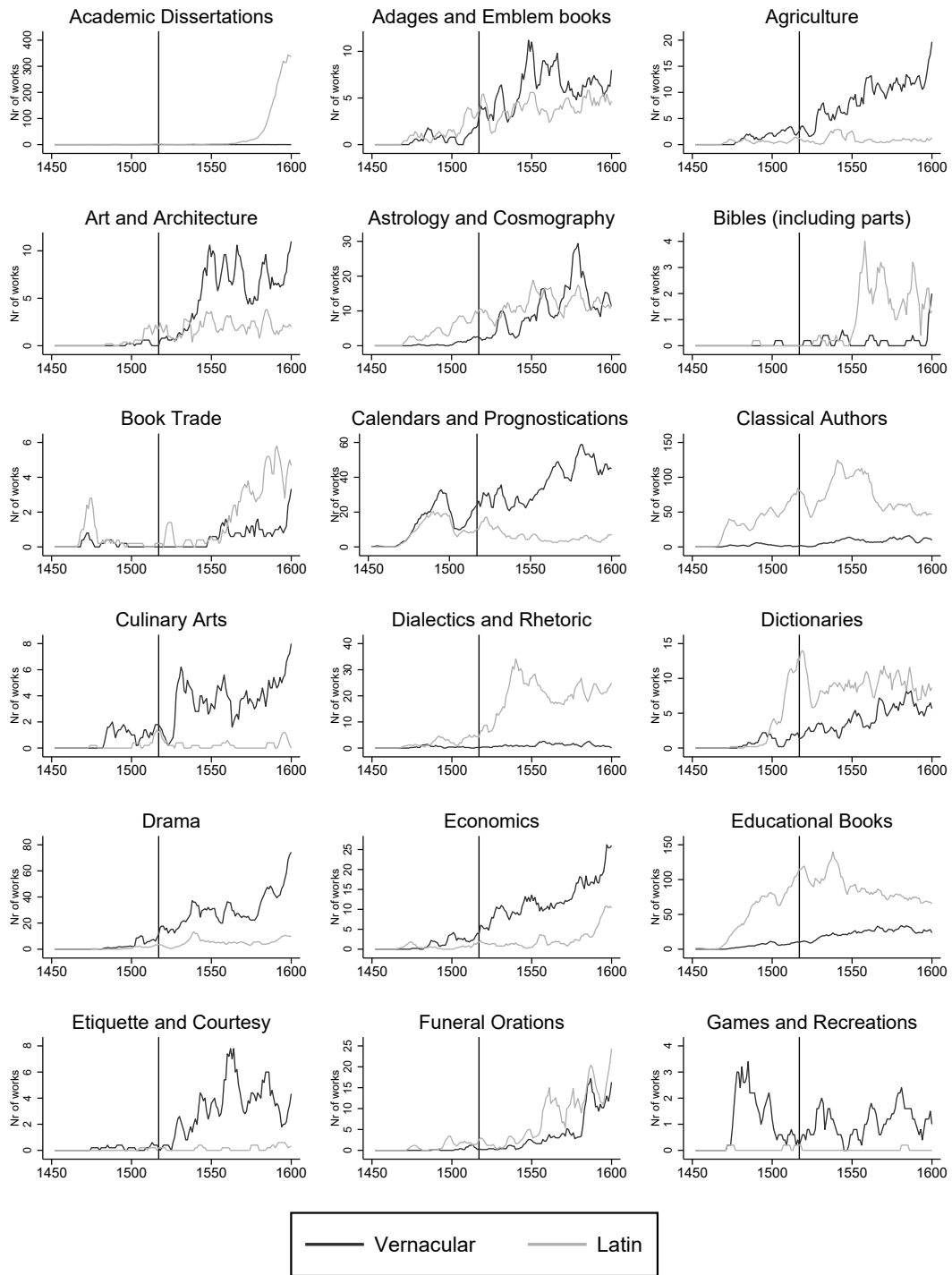


Figure 5: Total Number of Works Printed in the Vernacular and in Latin, by Subject Classification (A–Ga)

Notes: Universal Short Title Catalogue (USTC). 5-year moving average. The vertical line indicates the onset of the Protestant Reformation in 1517.

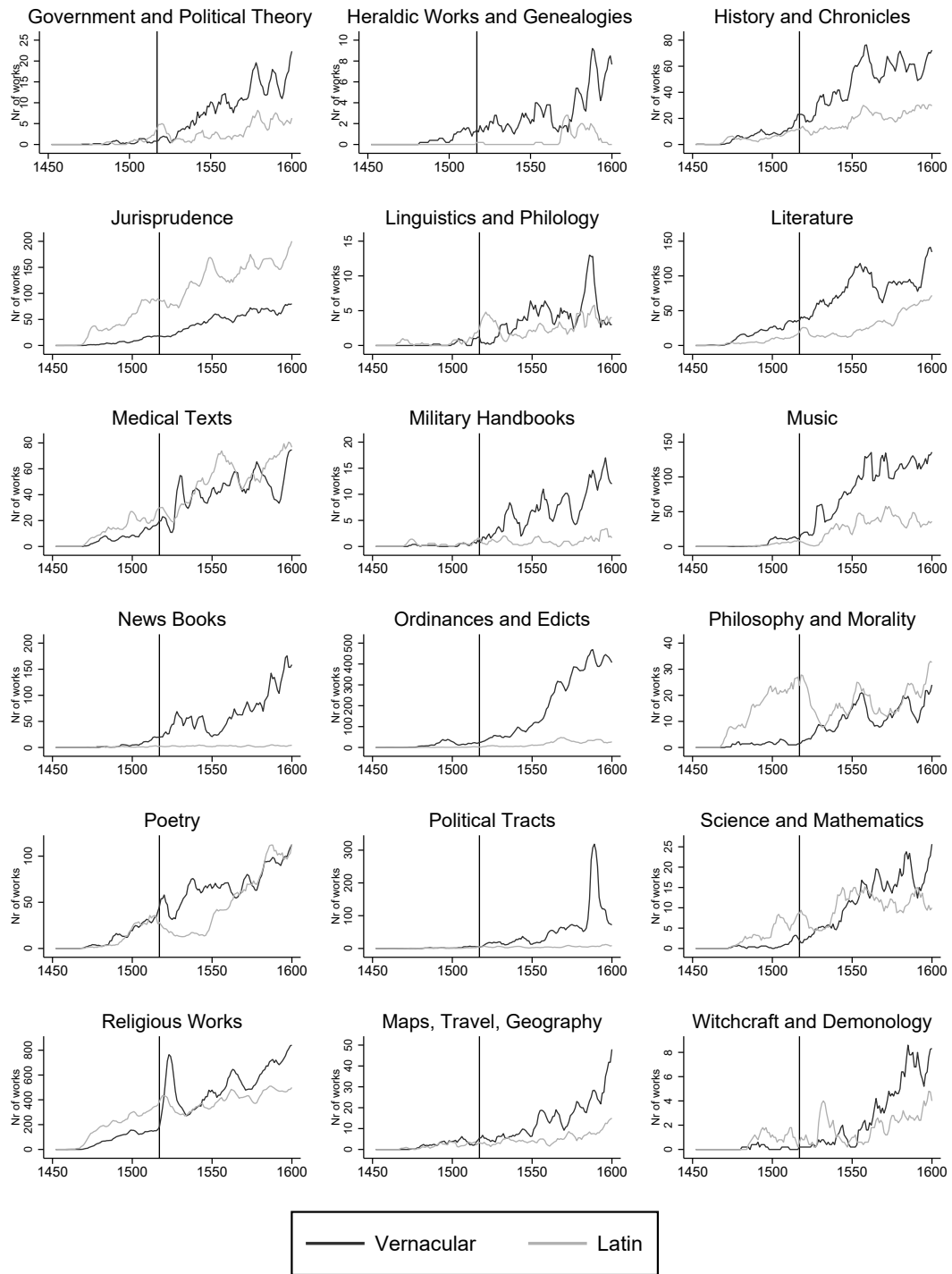


Figure 6: Total Number of Works Printed in the Vernacular and in Latin, by Subject Classification (G-Z)

Notes: Universal Short Title Catalogue (USTC). 5-year moving average. The vertical line indicates the onset of the Protestant Reformation in 1517.

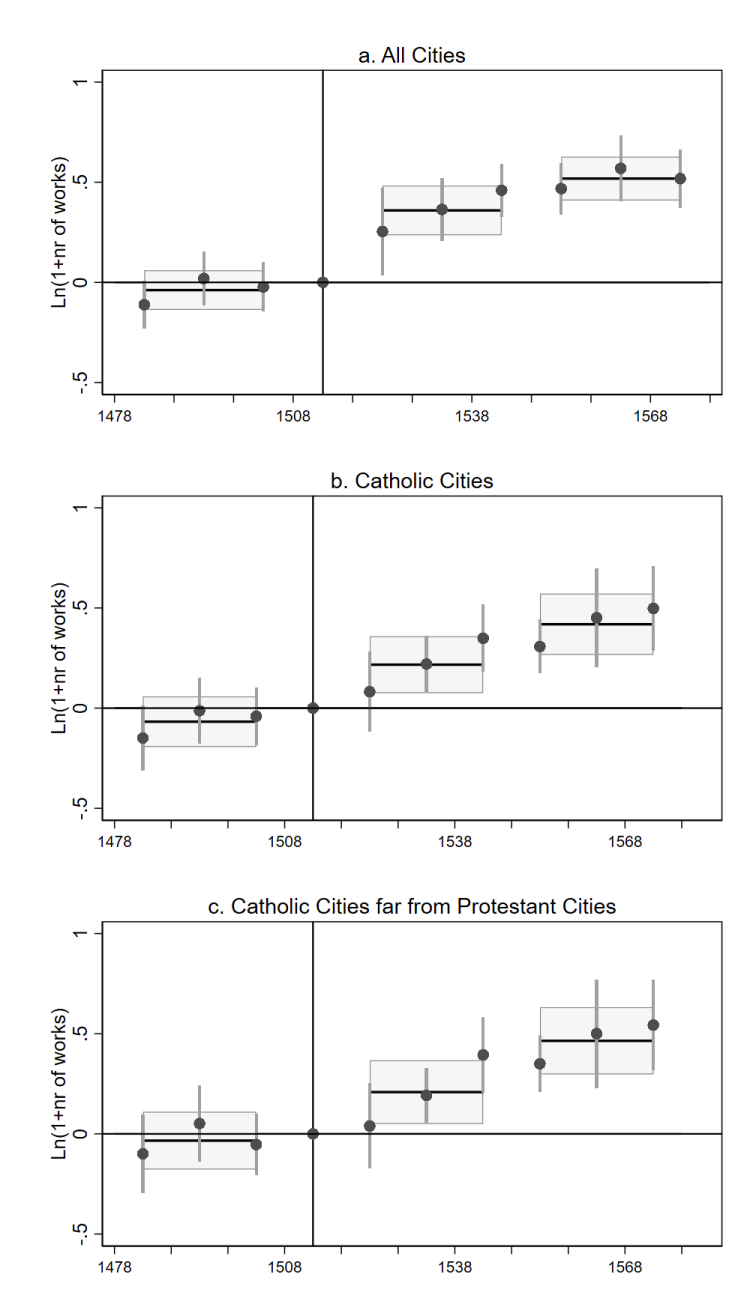


Figure 7: The Rise of the Vernaculars in Printing Across Samples (Within-City Differences)

Notes: Throughout, the dependent variable is $\ln(1+\text{total number of works in the vernacular/Latin printed in city } i \text{ and decade } t)$, as in Figure 3. We first show again our main results, based on the full sample of (Protestant and Catholic) printing cities (“All Cities”), we then drop all Protestant printing cities (“Catholic Cities”) and, finally, we additionally drop Catholic cities located close (within 100km) to a Protestant city (“Cath. City far from Prot. City”). Each figure shows the coefficient estimates on the interactions between a dummy for vernacular works and decade fixed effects as dots, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, with their 90% confidence intervals indicated as boxes. All regressions include city fixed effects and are restricted to cities with some printing output over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517.

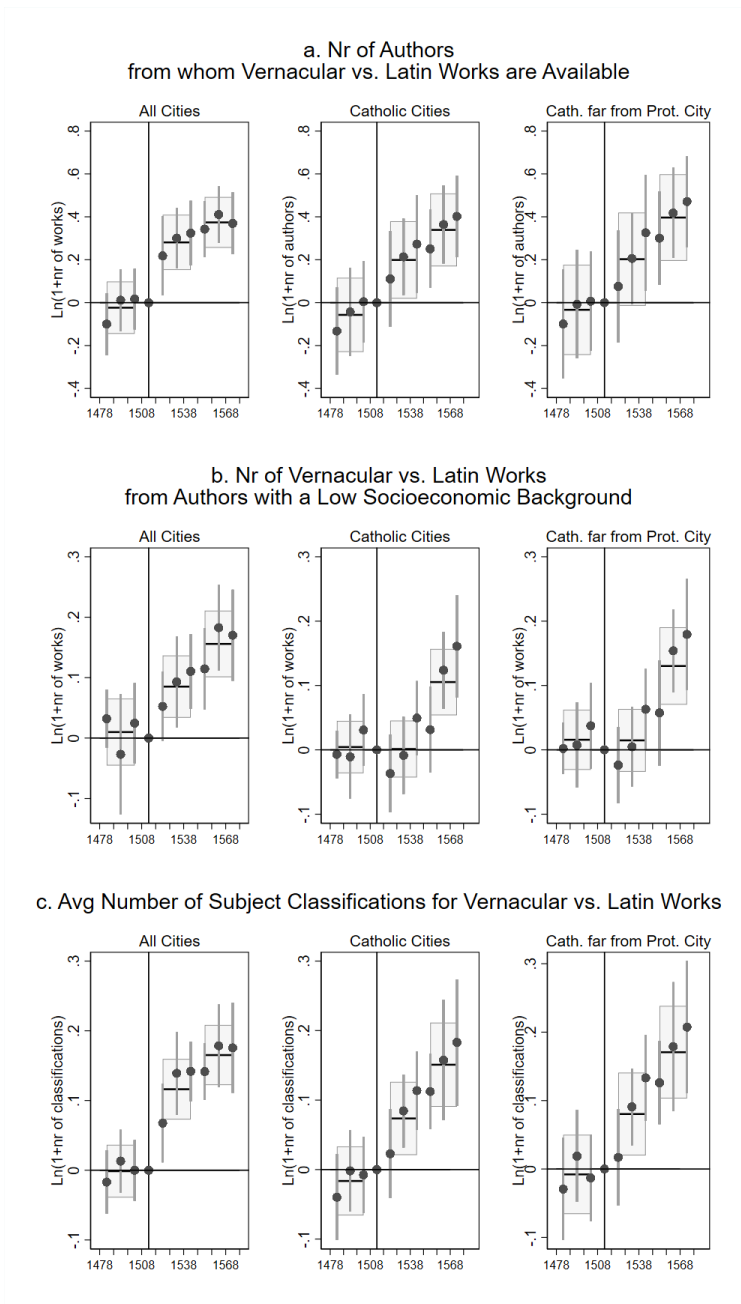


Figure 8: Increase in the Diversity of Printed Works Across Samples (Within-City Differences)

Notes: As in Figure 4 before, the dependent variable is $\ln(1+\text{total number of authors from whom vernacular/Latin works were available in city } i \text{ and decade } t)$ in Panel a., $\ln(1+\text{total number of vernacular/Latin works from authors with 50 works or more with a low socioeconomic background in city } i \text{ and decade } t)$ in Panel b., and $\ln(1+\text{average number of subject classifications for vernacular/Latin works in city } i \text{ and decade } t)$ in Panel c. We first show again our main results, based on the full sample of (Protestant and Catholic) printing cities (“All Cities”), we then drop all Protestant printing cities (“Catholic Cities”) and, finally, we additionally drop Catholic cities located close (within 100km) to a Protestant city (“Cath. City far from Prot. City”). Each figure plots the coefficient estimates on the interactions between a dummy for vernacular and decade fixed effects as dots, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, with their 90% confidence intervals indicated as boxes. All regressions include city fixed effects and are restricted to cities with some printing output over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517.

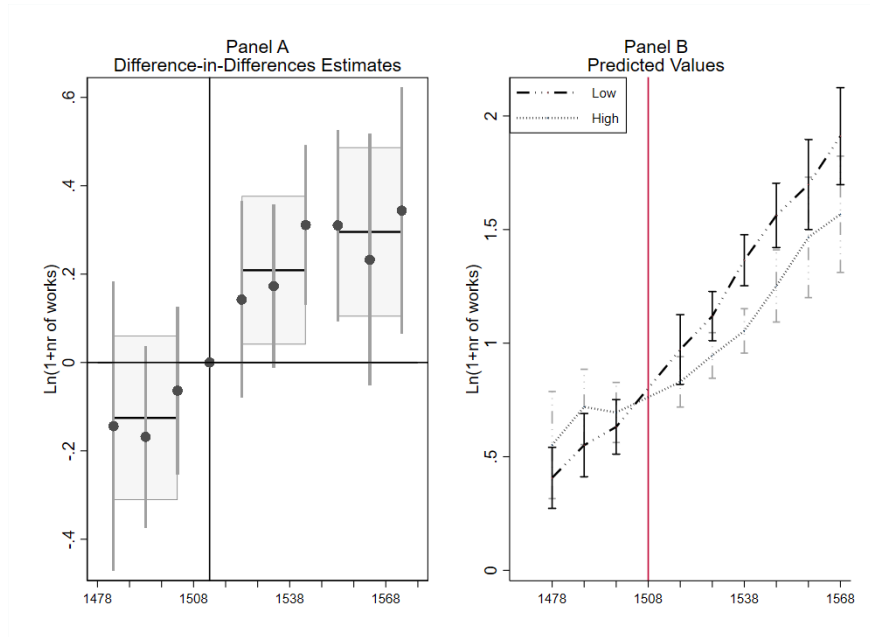


Figure 9: The Rise of Works Printed in the Vernacular: Heterogeneous Effects by Cities' Linguistic Distance to Latin

Notes: The regression is restricted to cities located outside the Holy Roman Empire. Panel A shows the coefficient estimates on the interactions between a dummy for vernacular works and decade fixed effects as dots, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, with their 90% confidence intervals are indicated as boxes. Panel B shows the predicted values of the dependent variable for cities with a low and a high linguistic distance to Latin with their 90% confidence intervals. The dependent variable is $\ln(1+\text{total number of vernacular works printed in city } i \text{ and decade } t)$. Cities with a linguistic distance to Latin at the 50th percentile or below (above) are classified as low (high) linguistic distance cities (for details see Section 6.2.1). The regression includes city fixed effects and is restricted to cities located outside the Holy Roman Empire with some printing output over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517. The regression results are reported in column (1) of Appendix Table E3.

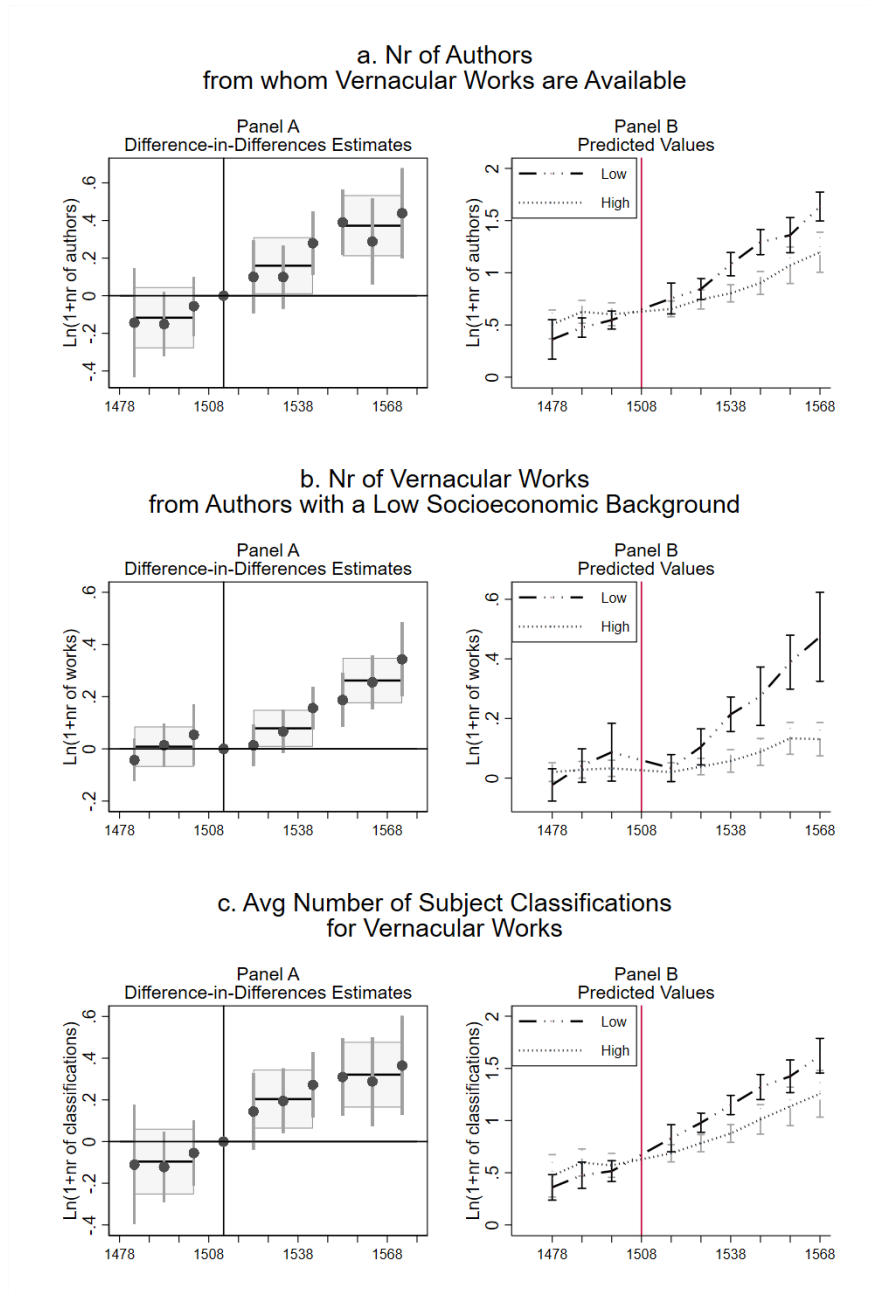


Figure 10: Increase in the Diversity of Vernacular Works: Heterogeneous Effects by Cities' Linguistic Distance to Latin

Notes: All regressions are restricted to cities located outside the Holy Roman Empire. Panel A plots the difference-in-differences estimates on the interactions between a dummy for low linguistic distance city and decade fixed effects as dots, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, and their 90% confidence intervals are indicated as boxes. Panel B plots the predicted values of the dependent variable for cities with a low and a high linguistic distance to Latin with their 90% confidence intervals. The dependent variable is $\ln(1+\text{total number of authors from whom vernacular/Latin works were available in city } i \text{ and decade } t)$ in Panel a., $\ln(1+\text{total number of vernacular works from authors with 50 works or more with a low socioeconomic background in city } i \text{ and decade } t)$ in Panel b., and $\ln(1+\text{average number of subject classifications for vernacular works})$ in city i and decade t in Panel c. Cities with a linguistic distance to Latin at the 50th percentile or below (above) are classified as low (high) linguistic distance cities (for details see Section 6.2.1). All regressions include city fixed effects and are restricted to cities with some printing output over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517. The regression results are reported in columns (2)–(4) of Appendix Table E3.

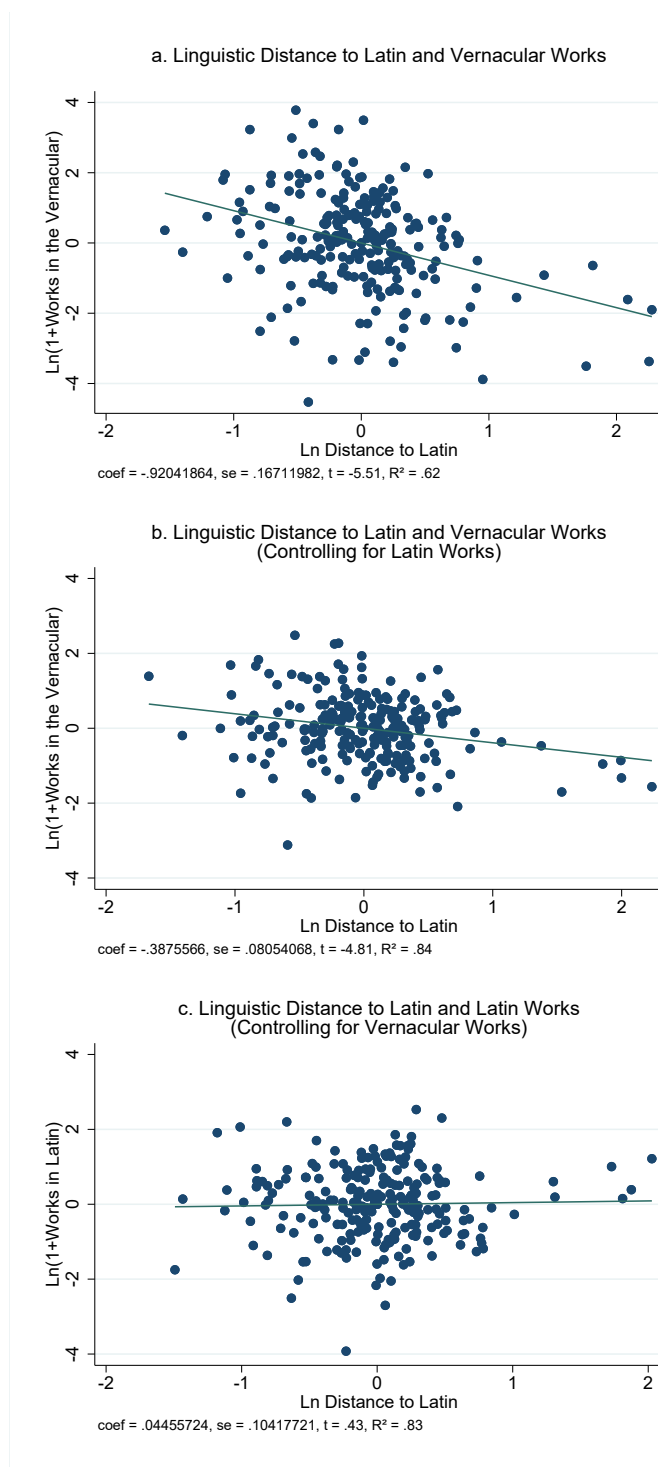


Figure 11: Cities' Linguistic Distance to Latin and Vernacular and Latin Printing Output

Notes: Based on equation 6, Panel a. shows the relationship between cities' linguistic distance to Latin and cities' vernacular printing output over the period 1518–1600. Panel b. is based on the same equation with cities' Latin printing output over the period 1518–1600 as additional control. Panel c. shows the relationship between cities' linguistic distance to Latin and cities' Latin printing output when controlling for cities' vernacular printing output.

Tables

Table 1: Cities' Vernacular Printing Output and Longer-Run Development

<i>Panel A: OLS</i>						
	<i>City Growth 1600–1700</i>			<i>Famous People Born 1600–1700</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(1 + Vernacular Works 1518-1600)	0.048*** (0.017)	0.063*** (0.021)	0.066*** (0.018)	0.138*** (0.045)	0.134*** (0.051)	0.133*** (0.051)
R^2	0.191	0.240	0.255	0.620	0.648	0.648
N	232	232	232	232	232	232
<i>Panel B: IV (Second Stage)</i>						
	<i>City Growth 1600-1700</i>			<i>Famous People Born 1600-1700</i>		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(1 + Vernacular Works 1518-1600)	0.087** (0.038)	0.133*** (0.046)	0.129*** (0.044)	0.285*** (0.057)	0.340*** (0.072)	0.351*** (0.077)
N	232	232	232	232	232	232
<i>Panel C: IV (First Stage)</i>						
	<i>Ln(1 + Vernacular Works 1518–1600)</i>					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln Distance to Latin	-1.063*** (0.136)	-0.938*** (0.155)	-0.920*** (0.167)	-1.001*** (0.139)	-0.929*** (0.163)	-0.914*** (0.178)
F-Statistic on IV	61.000	36.704	30.333	53.032	32.642	26.309
R^2	0.531	0.598	0.618	0.570	0.610	0.628
Country Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Ln Population 1600	Yes	Yes	Yes	Yes	Yes	Yes
Geographic Controls	No	Yes	Yes	No	Yes	Yes
Historic Controls	No	No	Yes	No	No	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Panel A reports OLS estimation results while Panel B reports IV estimation results with the first stage results reported in Panel C. The dependent variable in columns (1)-(3) of Panel A and B is $\ln(1 +$ the number of births of famous people) in city i over the period 1600 to 1700. The dependent variable in columns (4)-(6) of Panel A and B is \ln population growth in city i over the period 1600 to 1700. The dependent variable in Panel C is $\ln(1 +$ the number of vernacular works) printed in city i over the period 1518 to 1600. Geographic controls include latitude, longitude, log distance to the coast, and log distance to a trade route. Historical controls include dummies for hosting a university in 1600, being Protestant in 1600, and hosting a bishop in 1517. In columns (1) to (3), we additionally control for lagged upper-tail human capital (\ln famous people born in the period 1400–1500). Throughout, standard errors are clustered at the territory level and shown in parentheses.

Table 2: Cities' Vernacular Printing Output and Future City Growth: Robustness Analysis

	<i>City Population Growth (Second Stage, IV Estimation)</i>						
	<i>1600–1700</i>						<i>1400–1500</i>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(1 + Vernacular Works 1518–1600)	0.168** (0.072)	0.174*** (0.052)	0.129** (0.064)	0.115** (0.053)	0.235* (0.125)		0.049 (0.046)
Ln(1 + Latin Works 1518–1600)						-1.500 (3.669)	
F-Statistic on IV	11.554	35.710	17.309	27.581	23.155	0.183	9.085
N	174	195	175	162	232	232	155
	<i>Panel B: IV (First Stage)</i>						
	<i>Ln(1 + Vernacular Works 1518–1600)</i>						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln Distance to Latin	-0.807*** (0.237)	-0.903*** (0.145)	-0.844*** (0.203)	-0.890*** (0.169)	-0.388*** (0.081)	0.045 (0.104)	-1.324*** (0.439)
Observations	174	195	175	162	232	232	155
R^2	0.596	0.574	0.612	0.685	0.852	0.838	0.517
Excl. Cities w/ Low Printing Output	Yes	No	No	No	No	No	No
Excl. Cities w/ Greek Presence in 1500	No	Yes	No	No	No	No	No
Excl. Cities w/ Plague in 1500–1600	No	No	Yes	No	No	No	No
Excl. Cities in the HRE	No	No	No	Yes	No	No	No
Ln(1 + Latin Works 1518–1600)	No	No	No	No	Yes	-	No
Ln(1 + Vernacular Works 1518–1600)	-	-	-	-	-	Yes	-

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table examines the robustness of the IV results reported in column (3) of Table 1. In column (1) we exclude cities whose vernacular printing output is within the bottom 25% of the sample. In column (2) we exclude cities with a Greek presence around the year 1500. In column (3) we exclude cities with at least one plague between 1500 and 1600. In column (4) we exclude cities located in the Holy Roman Empire. In column (5), we add a city's Latin printing output as a control variable. In column (6), we use cities' linguistic distance to Latin as instrument for cities' Latin printing output while controlling for vernacular printing output. In column (7), we consider city population growth in the preceding century (1400–1500) as the dependent variable. In this specification, we control for the natural logarithm of the population in 1400 (instead of 1600).

Appendix

Table of Contents

A	Literacy Rates in the Vernacular in Early Modern Europe	48
B	Vernacularization in Early Modern Europe	49
C	Universal Short Title Catalogue (USTC)	50
D	Summary Statistics	54
E	Vernacularization in Printing and the Knowledge and Ideas Available at the City Level	55
E.1	Main Results	55
E.2	Vernacular Works Printed in Cities Located Outside the Holy Roman Empire: Heterogeneous Effects by Cities' Linguistic Distance to Latin	59
F	Vernacular Printing Output and Longer-Run Development	65
F.1	Linguistic Distance to Latin as Instrumental Variable	65
F.2	Robustness Analysis: Vernacular Printing Output and Upper-Tail Human Capital	67

A Literacy Rates in the Vernacular in Early Modern Europe

In the context of Flanders, for example, Kadens (2001) notes that the first legal document in Flemish was written in 1249, and in the following decades Flemish increasingly became the language of legal documents.²³ Graff (1991) and Houston (2014) document that merchants and other professional groups became increasingly literate over the course of the 14th and 15th centuries. For example, in the context of England, Graff (1991, p.105) notes that “administrative needs and commercial requirements insured that the larger merchant and the burgher class would be virtually universally literate at least in English [...] Little Latin was required for trade, banking or international commerce.” An important archive from the merchant Francesco Datini (1335–1410) of Prato, Italy, with over 150,000 letters, shows that the vernacular was the preferred mode of communication for the merchant classes (Marshall, 1999; Struppa and Kruslin, 2016). Graff (1991) and other sources document the rise of the vernacular across Europe in schools and other institutions for formal learning, albeit at different speeds. For example, based on data from the year 1480 from the *catasto*, “a comprehensive survey of the persons and their status, the assets, and the liabilities of every household in Florence,” Grendler (1991, Table 3.1, p.75) reports there were a total of 1,031 boys enrolled in school. Of these, only 23 were in grammar schools where boys studied in Latin and only 52 were in *chierico* – that is, when a boy lives “at home while he studied the Latin curriculum to become a clergyman.” Hence, a total of just 75 boys (7.2%) were learning Latin. On the other hand, 300 boys were enrolled in *leggere*, which “usually meant learning to read and/or write, that is, early elementary schooling,” and 253 boys were enrolled in *abbaco* schools, where they were taught basic arithmetic, geometry, bookkeeping, and reading and writing in the vernacular. Thus, a total of 553 students (or 53.6%) attained literacy in the vernacular as compared with only 7.2% in Latin. (The remaining 403 boys are classified as attending *alla scuola*. These were boys who attended school without indicating what kind of school.) Finally, for Germany, one estimate suggests that 1–4% of the entire population and well over 5% of the city population was literate in the vernacular in 1500 (von Polenz (2013, p. 128) based on Engelsing (1973)).

²³ Interestingly, those regions that were some of the early ones to accelerate vernacularization – parts of Flanders such as Bruges, Ghent, Ypres, and the Italian city states – were to become some of the most dynamic parts of Europe in the 13th and 14th centuries. Van Bavel (2010, p. 314–319) discusses the increasing use of the vernacular (both French and Dutch) in the medieval Low Countries from the 13th century onward.

B Vernacularization in Early Modern Europe

Table B1: Selected Works in Defense of the Vernacular

Year	Author	Title
1529	Dante	<i>De vulgari eloquentia</i>
1540	J. de Barros	<i>Louvor de nossa linguagem</i>
1542	S. Speroni	<i>Dialogo della lingua</i>
1549	J. Du Bellay	<i>Deffense et illustration de la langue francaise</i>
1574	M. Viziana	<i>Alabancas de las lenguas castellana y valenciana</i>
1586	S. Stevin	<i>Weerdigheyt der duytsche tael</i>
1615	R. Carew	<i>On the excellence of the English Tongue</i>

Source: Burke (2004, p. 65).

C Universal Short Title Catalogue (USTC)

Table C1: Books and Pamphlets by Country (N=354,354).

Country	Freq.	Percent
Balkans	211	0.06
Bohemia and Moravia	1,748	0.49
Denmark	1,078	0.30
England	14,385	4.06
France	77,461	21.86
Holy Roman Empire	99,995	28.22
Hungary	686	0.19
Italian States	77,709	21.93
Low Countries	29,352	8.28
Poland	4,215	1.19
Portugal	1,510	0.43
Scotland	375	0.11
Spain	15,195	4.29
Sweden	440	0.12
Swiss Confederation	14,964	4.22
<i>Missing information</i>	15,030	4.24
Total	354,354	100.00

Table C2: Books and Pamphlets by Language (N=354,354).

Language	Freq.	Percent
Arabic	11	0.00
Armenian	2	0.00
Catalan	756	0.21
Czech	1,886	0.53
Danish	824	0.23
Dutch	12,001	3.39
English	11,612	3.28
French	46,514	13.13
German	52,071	14.69
Greek	3,427	0.97
Hebrew	409	0.12
Hungarian	493	0.14
Italian	39,769	11.22
Latin	165,943	46.83
Polish	1,019	0.29
Portuguese	848	0.24
Spanish	10,865	3.07
Swedish	393	0.11
<i>Missing information</i>	5,511	1.56
Total	354,354	100.00

Table C3: Books and Pamphlets by Subject Classification (N=354,354).

Classification	Freq.	Percent
Academic Dissertations	6,456	1.82
Adages and Emblem Books	1,542	0.44
Agriculture	1,026	0.29
Art and Architecture	828	0.23
Astrology and Cosmography	2,434	0.69
Bibles (including parts)	4,358	1.23
Book Trade	335	0.09
Calendars and Prognostications	5,432	1.53
Classical Authors	14,996	4.23
Culinary Arts	431	0.12
Dialectics and Rhetoric	2,184	0.62
Dictionaries	2,238	0.63
Drama	4,377	1.24
Economics	2,219	0.63
Educational Books	14,497	4.09
Etiquette and Courtesy	373	0.11
Funeral Orations	2,997	0.85
Games and Recreations	180	0.05
Government and Political Theory	1,298	0.37
Heraldic Works and Genealogies	410	0.12
History and Chronologies	8,165	2.30
Jurisprudence	20,985	5.92
Linguistics and Philology	695	0.20
Literature	12,360	3.49
Medical Texts	10,662	3.01
Military Handbooks	893	0.25
Music	12,755	3.60
News Books	8,692	2.45
Ordinances and Edicts	23,351	6.59
Philosophy and Morality	3,896	1.10
Poetry	13,890	3.92
Political Tracts	7,494	2.11
Religious Works	106,673	30.10
Science and Mathematics	2,454	0.69
Maps, Travel, Geography	2,172	0.61
Wedding Pamphlets	2	0.00
Witchcraft and Demonology	517	0.15
<i>Missing information</i>	50,087	14.13
Total	354,354	100.00

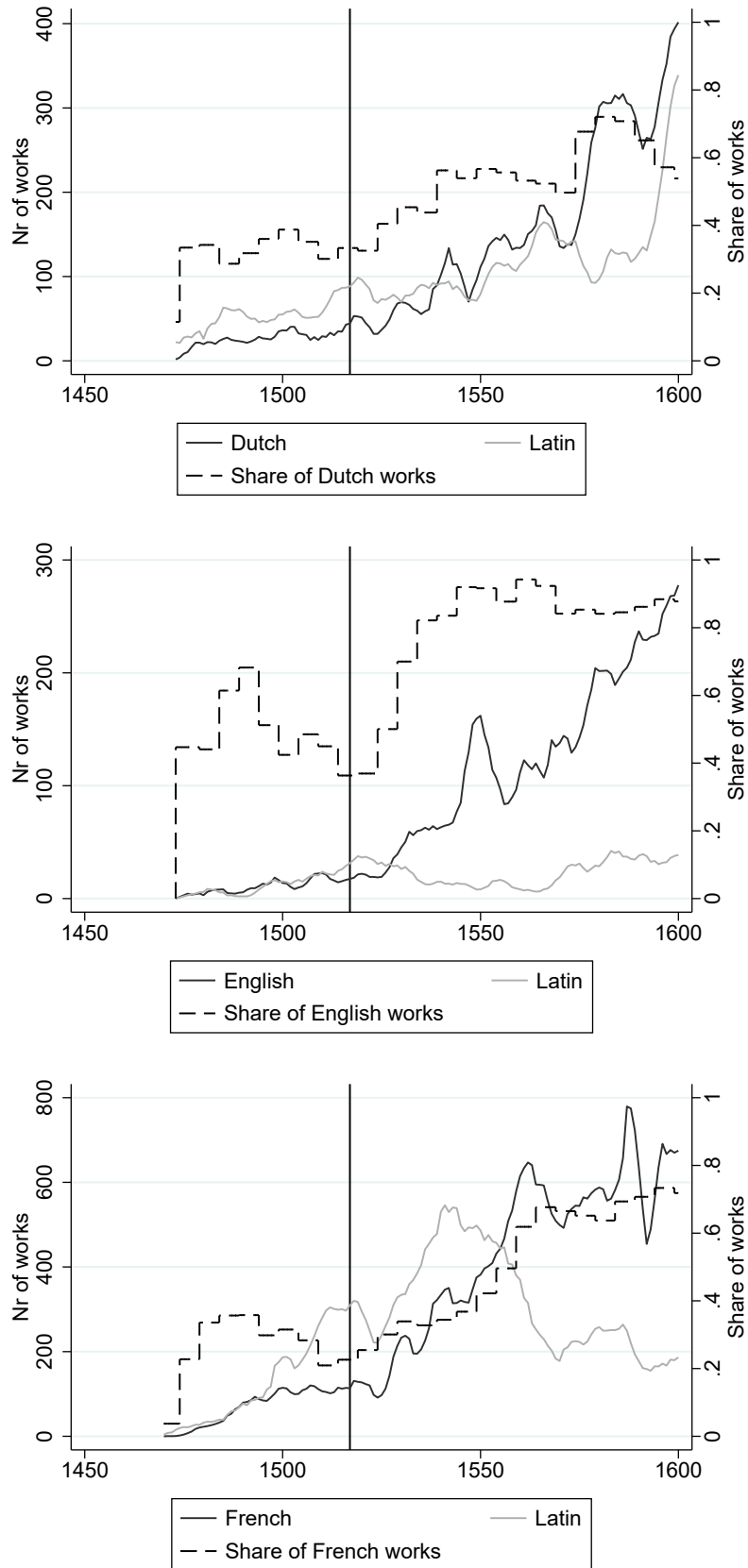


Figure C1: Total Number of Titles Printed in Europe, By Language Area

Notes: Universal Short Title Catalogue (USTC). 5-year moving average. The vertical line indicates the onset of the Protestant Reformation in 1517.

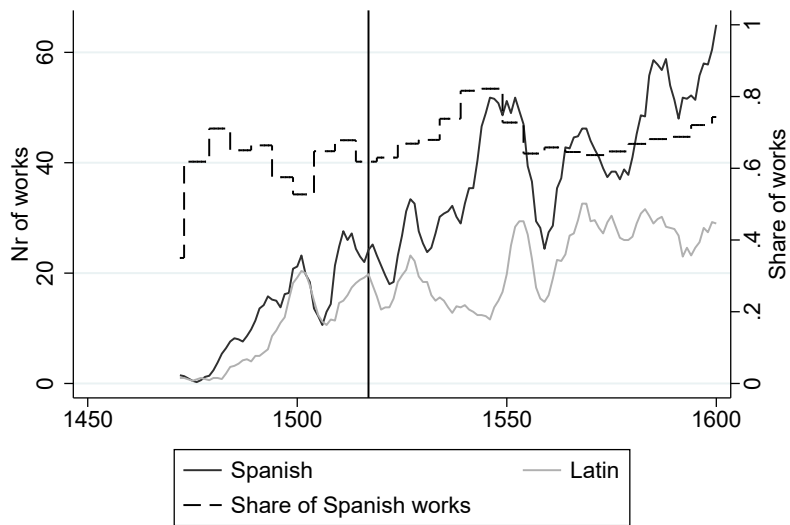
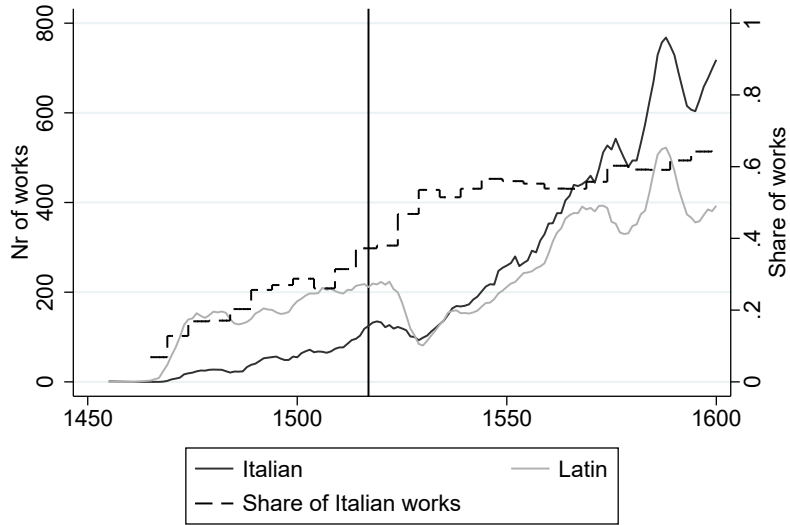
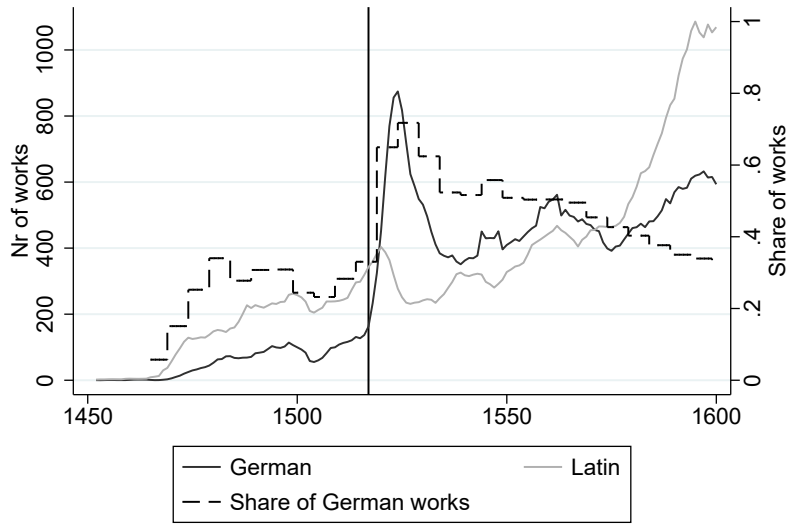


Figure C2: Total Number of Titles Printed in Europe, By Language Area (cont'd)

Notes: Universal Short Title Catalogue (USTC). 5-year moving average. The vertical line indicates the onset of the Protestant Reformation in 1517.

D Summary Statistics

Table D1: Summary Statistics (Sections 4–6)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Number of Works (Vernacular)	3,230	28.34	147.14	0	2,936
Number of Works (Latin)	3,230	27.95	145.67	0	2,893
Number of Authors (Vernacular)	3,230	9.71	40.43	0	673
Number of Authors (Latin)	3,230	11.92	48.34	0	671
Number of Works from Authors w/ a Low SE Background (Vernacular)	3,230	0.97	6.68	0	176
Number of Works from Authors w/ a Low SE Background (Latin)	3,230	1.00	6.89	0	167
Average Number of USTC Subject Classifications (Vernacular)	3,230	0.82	2.37	0	25.5
Average Number of USTC Subject Classifications (Latin)	3,230	0.75	2.19	0	20.5
Distance to the Coast (in km) ¹	3,230	152.15	135.18	0.03	480.31
Distance to a Trade Route (in km) ¹	3,230	90.33	123.34	0.00	989.73
Latitude ¹	3,230	47.36	4.48	37.01	59.52
Longitude ¹	3,230	6.54	6.23	-9.08	21.00
=1 if Bishop in 1517 ¹	3,230	0.48	0.50	0	1
=1 if University in 1450 ¹	3,230	0.17	0.38	0	1
=1 if Protestant in 1600 ¹	3,230	0.36	0.48	0	1

¹ Time-invariant.

Notes: The sample is comprised of European cities with some printing output over the considered period (1478–1577), for which location and additional city-level information is available (for details, see Section 3). There are in total 323 such cities. We aggregate outcomes over 10 decades, from 1478–1487 to 1568–1577, resulting in 3,230 observations in total. “SE” stands for socioeconomic.

Table D2: Summary Statistics (Section 7)

Variable	Obs.	Mean	Std. Dev.	Min	Max
Number of Vernacular Works 1518–1600	232	592.06	1912.13	0	19,879
Number of Latin Works 1518–1600	232	463.34	1700.67	0	16,121
Log City Growth 1600–1700	232	0.03	0.56	-1.61	2.48
Number of Famous People Born 1600–1700	232	37.98	86.44	0	1,045
Linguistic Distance to Latin	232	0.82	0.70	0.07	6.11
Population in 1600	232	25.26	36.00	2	300
Number of Famous People Born 1400–1500	232	10.06	24.27	0	290
Distance to the Coast (in km)	232	128.79	124.35	0.02	426.47
Distance to a Trade Route (in km)	232	95.02	137.25	0.00	989.74
Latitude	232	47.27	4.76	36.43	59.52
Longitude	232	6.34	6.39	-9.08	21
=1 if Bishop in 1517	232	0.53	0.50	0	1
=1 if Protestant in 1600	232	0.37	0.48	0	1
=1 if University in 1450	232	0.22	0.41	0	1

Notes: The sample is comprised of European printing cities (see the notes of Appendix Table D1) for which city population data is available for 1500 and 1600 (N=232).

E Vernacularization in Printing and the Knowledge and Ideas Available at the City Level

E.1 Main Results

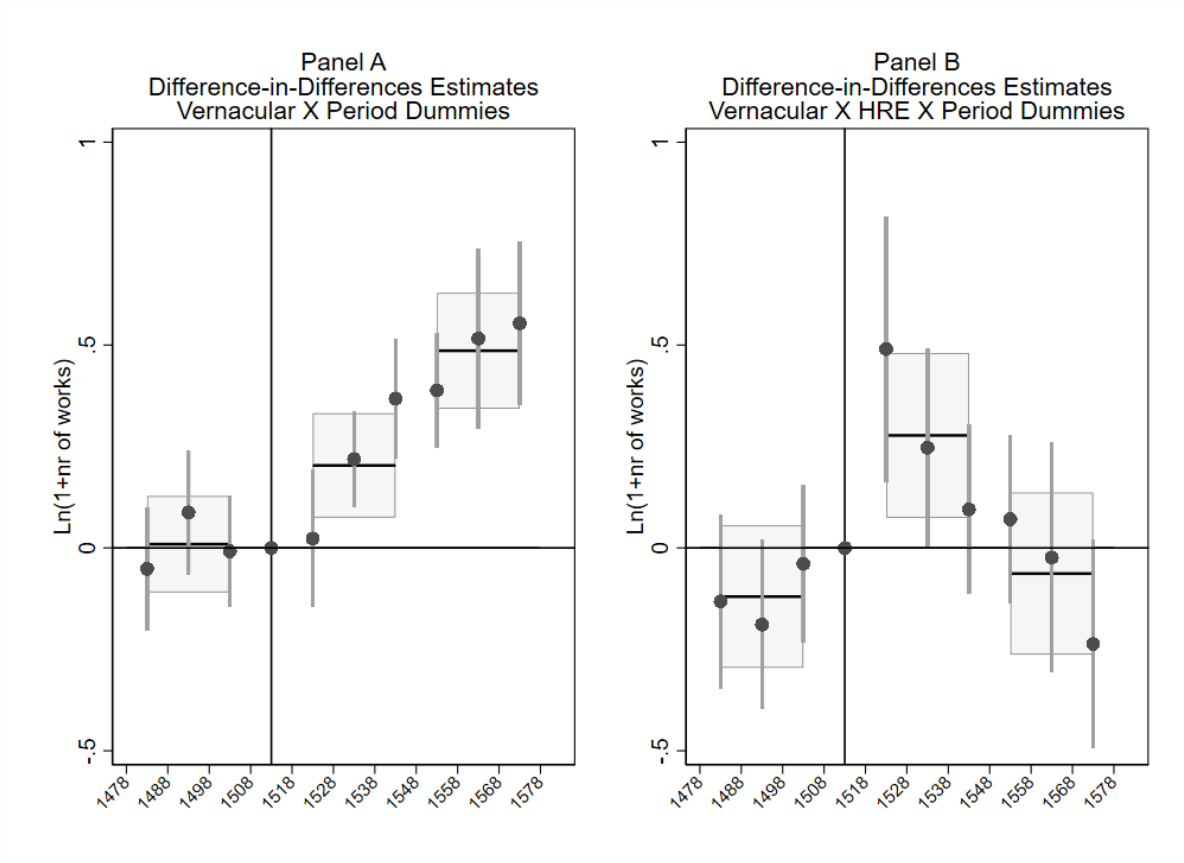


Figure E1: The Rise of the Vernaculars in Printing: Take-Off in the Holy Roman Empire (HRE)

Notes: This figure shows within-city differences in vernacular and Latin printing output over time (relative to 1508–1517), allowing for a differential change for cities located in the Holy Roman Empire. The dependent variable is $\ln(1+\text{total number of vernacular/Latin works printed in city } i \text{ and decade } t)$. Panel A shows the coefficient estimates on the interactions between a dummy for vernacular works and decade fixed effects, with the omitted decade being 1508–1517 and with their 90% confidence intervals indicated with vertical lines. The coefficient estimates in Panel B give the differential change in vernacular relative to Latin printing output for cities located in the Holy Roman Empire (compared with 1508–1517), with their 90% confidence intervals. Coefficient estimates on the interactions for 30-year periods are shown with horizontal lines, with their 90% confidence intervals are indicated as boxes. The regression includes city fixed effects and is restricted to cities with some printing over the period 1478–1577. Standard errors are clustered at the territory \times decade level. The vertical line indicates the onset of the Protestant Reformation in 1517.

Table E1: Regression Results: The Rise of the Vernaculars in Printing and the Diversity of Printed Works

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Vernacular	-0.128*** (0.049)	-0.174*** (0.065)	-0.058** (0.025)	-0.053** (0.021)
Vernacular X 1478-1487	-0.112 (0.071)	-0.100 (0.087)	0.032 (0.029)	-0.017 (0.027)
Vernacular X 1488-1497	0.019 (0.081)	0.012 (0.087)	-0.027 (0.060)	0.013 (0.027)
Vernacular X 1498-1507	-0.023 (0.074)	0.018 (0.086)	0.025 (0.040)	-0.000 (0.026)
Vernacular X 1518-1527	0.254* (0.132)	0.218* (0.112)	0.052 (0.035)	0.068** (0.034)
Vernacular X 1528-1537	0.364*** (0.095)	0.301*** (0.086)	0.093** (0.046)	0.139*** (0.036)
Vernacular X 1538-1547	0.460*** (0.079)	0.324*** (0.091)	0.110*** (0.037)	0.142*** (0.026)
Vernacular X 1548-1557	0.468*** (0.077)	0.342*** (0.079)	0.115*** (0.041)	0.142*** (0.024)
Vernacular X 1558-1567	0.569*** (0.099)	0.411*** (0.079)	0.183*** (0.043)	0.179*** (0.036)
Vernacular X 1568-1577	0.518*** (0.087)	0.369*** (0.087)	0.170*** (0.046)	0.176*** (0.039)
1478-1487	-0.105 (0.087)	-0.040 (0.082)	-0.051 (0.038)	-0.062** (0.025)
1488-1497	-0.130* (0.075)	-0.058 (0.064)	0.032 (0.032)	-0.051** (0.023)
1598-1507	-0.046 (0.056)	-0.024 (0.053)	0.002 (0.020)	-0.032* (0.018)
1518-1527	0.120 (0.074)	0.085 (0.068)	0.116*** (0.030)	0.017 (0.023)
1528-1537	-0.027 (0.065)	-0.029 (0.053)	0.088*** (0.028)	-0.021 (0.021)
1538-1547	0.057 (0.059)	0.044 (0.056)	0.132*** (0.028)	0.009 (0.017)
1548-1557	0.191*** (0.067)	0.178*** (0.054)	0.130*** (0.028)	0.054*** (0.019)
1558-1567	0.308*** (0.077)	0.275*** (0.061)	0.151*** (0.033)	0.096*** (0.026)
1568-1577	0.440*** (0.079)	0.403*** (0.065)	0.147*** (0.030)	0.142*** (0.027)
City Fixed Effects	Y	Y	Y	Y
R^2	0.721	0.730	0.514	0.761
Observations	6,460	6,460	6,460	6,460

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 1 with standard errors clustered at the territory \times decade level in parentheses. Restricted to cities with some printing output over the period 1478–1577. The dependent variable is $\ln(1+\text{total number of vernacular/Latin works in city } i \text{ and decade } t)$ in column (1), $\ln(1+\text{total number of authors from whom vernacular/Latin works were available in city } i \text{ and decade } t)$ in column (2), $\ln(1+\text{total number of vernacular/Latin works from authors with 50 works or more with a low socioeconomic background in city } i \text{ and decade } t)$ in column (3), and $\ln(1+\text{average number of subject classifications for vernacular/Latin works in city } i \text{ and decade } t)$ in column (4).

Table E2: Regression Results: Regression Results: The Rise of the Vernaculars in Printing and the Diversity of Printed Works (30-Years Periods)

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Vernacular	-0.128*** (0.049)	-0.174*** (0.065)	-0.058** (0.025)	-0.053** (0.021)
Vernacular X Pre	-0.038 (0.060)	-0.024 (0.074)	0.010 (0.034)	-0.001 (0.023)
Vernacular X Post1	0.359*** (0.075)	0.281*** (0.078)	0.085*** (0.031)	0.116*** (0.026)
Vernacular X Post2	0.518*** (0.066)	0.374*** (0.072)	0.156*** (0.033)	0.165*** (0.026)
1478-1487	-0.142* (0.081)	-0.078 (0.074)	-0.040 (0.038)	-0.070*** (0.023)
1488-1497	-0.101 (0.066)	-0.041 (0.058)	0.013 (0.024)	-0.044** (0.021)
1598-1507	-0.038 (0.054)	-0.004 (0.052)	0.009 (0.025)	-0.032* (0.017)
1518-1527	0.068 (0.066)	0.054 (0.062)	0.100*** (0.032)	-0.007 (0.020)
1528-1537	-0.025 (0.060)	-0.019 (0.054)	0.092*** (0.024)	-0.009 (0.019)
1538-1547	0.107* (0.057)	0.066 (0.054)	0.145*** (0.029)	0.022 (0.018)
1548-1557	0.166*** (0.063)	0.162*** (0.053)	0.109*** (0.027)	0.043** (0.020)
1558-1567	0.333*** (0.064)	0.293*** (0.058)	0.164*** (0.030)	0.103*** (0.022)
1568-1577	0.440*** (0.073)	0.401*** (0.063)	0.155*** (0.027)	0.147*** (0.024)
City Fixed Effects	Y	Y	Y	Y
R^2	0.720	0.730	0.513	0.760
Observations	6,460	6,460	6,460	6,460

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 2 with standard errors clustered at the territory \times decade level in parentheses. "Pre" refers to the pre-Reformation period 1478–1507, "Post1" refers to the post-Reformation period 1518–47, and "Post2" refers to the post-Reformation period 1548–1577. Restricted to cities with some printing output over the period 1478–1577. The dependent variable is $\ln(1+\text{total number of vernacular/Latin works in city } i \text{ and decade } t)$ in column (1), $\ln(1+\text{total number of authors from whom vernacular/Latin works were available in city } i \text{ and decade } t)$ in column (2), $\ln(1+\text{total number of vernacular/Latin works from authors with 50 works or more with a low socioeconomic background in city } i \text{ and decade } t)$ in column (3), and $\ln(1+\text{average number of subject classifications for vernacular/Latin works in city } i \text{ and decade } t)$ in column (4).

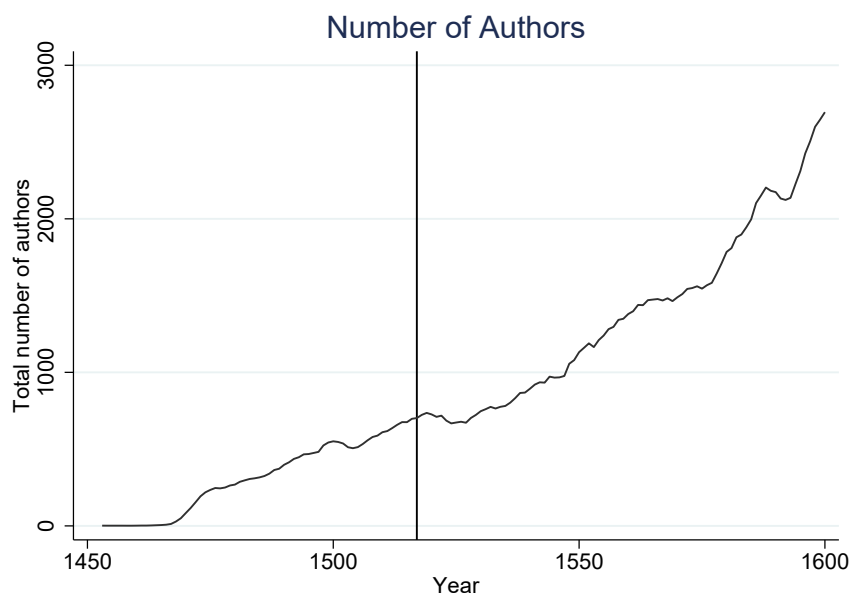


Figure E2: The Rise in the Total Number of Authors over Time

Notes: 5-year moving averages. The figure shows the total number of authors over time in early modern Europe. The vertical line indicates the onset of the Protestant Reformation in 1517. Source: Universal Short Title Catalogue (USTC).

E.2 Vernacular Works Printed in Cities Located Outside the Holy Roman Empire: Heterogeneous Effects by Cities' Linguistic Distance to Latin

Table E3: Heterogeneous Effects by Cities' Linguistic Distance to Latin: Main Results

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Low Distance to Latin X 1478-1487	-0.159 (0.198)	-0.144 (0.175)	-0.043 (0.049)	-0.110 (0.173)
Low Distance to Latin X 1488-1497	-0.172 (0.122)	-0.151 (0.103)	0.014 (0.050)	-0.122 (0.101)
Low Distance to Latin X 1498-1507	-0.069 (0.113)	-0.056 (0.095)	0.054 (0.070)	-0.055 (0.095)
Low Distance to Latin X 1518-1527	0.146 (0.133)	0.100 (0.118)	0.013 (0.047)	0.144 (0.111)
Low Distance to Latin X 1528-1537	0.185* (0.108)	0.100 (0.101)	0.066 (0.050)	0.195** (0.094)
Low Distance to Latin X 1538-1547	0.331*** (0.106)	0.280*** (0.101)	0.156*** (0.048)	0.272*** (0.094)
Low Distance to Latin X 1548-1557	0.343*** (0.130)	0.391*** (0.104)	0.187*** (0.062)	0.310*** (0.112)
Low Distance to Latin X 1558-1567	0.280* (0.168)	0.289** (0.137)	0.255*** (0.062)	0.289** (0.128)
Low Distance to Latin X 1568-1577	0.399** (0.175)	0.438*** (0.144)	0.343*** (0.086)	0.365** (0.143)
1478-1487	-0.205 (0.155)	-0.068 (0.091)	-0.008 (0.014)	-0.180 (0.130)
1488-1497	-0.048 (0.111)	0.053 (0.074)	-0.000 (0.011)	-0.052 (0.086)
1498-1507	-0.071 (0.094)	0.029 (0.075)	0.005 (0.010)	-0.080 (0.078)
1518-1527	0.054 (0.076)	0.080 (0.052)	-0.008 (0.014)	0.036 (0.055)
1528-1537	0.162** (0.064)	0.171*** (0.059)	0.011 (0.011)	0.133*** (0.051)
1538-1547	0.263*** (0.065)	0.230*** (0.054)	0.029 (0.019)	0.226*** (0.055)
1548-1557	0.447*** (0.099)	0.330*** (0.068)	0.060** (0.024)	0.361*** (0.083)
1558-1567	0.646*** (0.158)	0.498*** (0.105)	0.105*** (0.029)	0.485*** (0.109)
1568-1577	0.741*** (0.160)	0.623*** (0.115)	0.103*** (0.031)	0.606*** (0.132)
City Fixed Effects	Y	Y	Y	Y
R^2	0.809	0.814	0.529	0.805
Observations	1,630	1,630	1,630	1,630

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 3 with standard errors clustered at the territory \times decade level in parentheses. Restricted to cities located outside the Holy Roman Empire with some printing output over the period 1478–1577. Cities with a linguistic distance to Latin at the 50th percentile or below (above) are classified as low (high) linguistic distance cities (for details see Section 6.2.1). Throughout, only vernacular works are considered.

Table E4: Heterogeneous Effects by Cities' Linguistic Distance to Latin: Main Results (30-Years Periods)

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Low Distance to Latin X Pre	-0.133 (0.111)	-0.117 (0.098)	0.008 (0.047)	-0.096 (0.095)
Low Distance to Latin X Post1	0.221** (0.100)	0.160* (0.091)	0.079* (0.043)	0.204** (0.085)
Low Distance to Latin X Post2	0.341*** (0.115)	0.373*** (0.098)	0.262*** (0.052)	0.321*** (0.095)
1478-1487	-0.217** (0.109)	-0.081 (0.067)	-0.032* (0.019)	-0.187** (0.090)
1488-1497	-0.066 (0.100)	0.037 (0.066)	0.003 (0.017)	-0.064 (0.081)
1498-1507	-0.041 (0.094)	0.058 (0.073)	0.026 (0.026)	-0.061 (0.076)
1518-1527	0.019 (0.068)	0.052 (0.055)	-0.038** (0.015)	0.008 (0.051)
1528-1537	0.146** (0.065)	0.143** (0.057)	0.005 (0.017)	0.129** (0.049)
1538-1547	0.314*** (0.063)	0.286*** (0.056)	0.066*** (0.017)	0.258*** (0.050)
1548-1557	0.448*** (0.090)	0.338*** (0.067)	0.025 (0.030)	0.355*** (0.074)
1558-1567	0.618*** (0.138)	0.459*** (0.097)	0.102*** (0.033)	0.470*** (0.097)
1568-1577	0.768*** (0.133)	0.654*** (0.089)	0.141*** (0.043)	0.626*** (0.108)
City Fixed Effects	Y	Y	Y	Y
R^2	0.809	0.814	0.526	0.804
Observations	1,630	1,630	1,630	1,630

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 4 with standard errors clustered at the territory \times decade level in parentheses. "Pre" refers to the pre-Reformation period 1478–1507, "Post1" refers to the post-Reformation period 1518–1547, and "Post2" refers to the post-Reformation period 1548–1577. Restricted to cities with some printing output over the period 1478–1577. Cities with a linguistic distance to Latin at the 50th percentile or below (above) are classified as low (high) linguistic distance cities (for details see Section 6.2.1). Throughout, only vernacular works are considered.

Table E5: Heterogeneous Effects by Cities' Linguistic Distance to Latin: Dropping Cities Close to the Median Linguistic Distance to Latin

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Low Distance to Latin X 1478-1487	-0.141 (0.225)	-0.128 (0.217)	-0.060 (0.044)	-0.106 (0.193)
Low Distance to Latin X 1488-1497	-0.137 (0.140)	-0.093 (0.127)	-0.025 (0.039)	-0.100 (0.121)
Low Distance to Latin X 1498-1507	-0.022 (0.129)	0.058 (0.109)	0.048 (0.064)	-0.030 (0.111)
Low Distance to Latin X 1518-1527	0.131 (0.160)	0.086 (0.138)	-0.039 (0.040)	0.127 (0.136)
Low Distance to Latin X 1528-1537	0.171 (0.134)	0.090 (0.118)	0.014 (0.044)	0.189 (0.117)
Low Distance to Latin X 1538-1547	0.315** (0.130)	0.291** (0.121)	0.120** (0.048)	0.257** (0.116)
Low Distance to Latin X 1548-1557	0.368** (0.146)	0.387*** (0.115)	0.147** (0.065)	0.331** (0.128)
Low Distance to Latin X 1558-1567	0.395** (0.181)	0.356** (0.157)	0.226*** (0.058)	0.370** (0.149)
Low Distance to Latin X 1568-1577	0.474*** (0.177)	0.485*** (0.160)	0.312*** (0.085)	0.432*** (0.147)
1478-1487	-0.201 (0.127)	-0.074 (0.087)	-0.000 (0.014)	-0.175 (0.110)
1488-1497	-0.074 (0.097)	-0.005 (0.087)	-0.000 (0.014)	-0.064 (0.079)
1498-1507	-0.112 (0.087)	-0.074 (0.066)	-0.000 (0.014)	-0.101 (0.073)
1518-1527	0.060 (0.088)	0.082 (0.055)	-0.000 (0.014)	0.040 (0.068)
1528-1537	0.154** (0.072)	0.152** (0.067)	-0.000 (0.014)	0.126** (0.058)
1538-1547	0.273*** (0.076)	0.211*** (0.070)	0.011 (0.018)	0.245*** (0.075)
1548-1557	0.386*** (0.086)	0.277*** (0.063)	0.044* (0.026)	0.317*** (0.072)
1558-1567	0.569*** (0.129)	0.418*** (0.092)	0.078*** (0.026)	0.451*** (0.101)
1568-1577	0.679*** (0.159)	0.554*** (0.117)	0.083* (0.049)	0.556*** (0.137)
City Fixed Effects	Y	Y	Y	Y
R^2	0.769	0.791	0.440	0.770
Observations	1,290	1,290	1,290	1,290

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 3 with standard errors clustered at the territory \times decade level in parentheses. Restricted to cities with some printing output over the period 1478–1577. Cities with a linguistic distance to Latin between the 40th and the 60th percentiles are dropped from the sample. Throughout, only vernacular works are considered.

Table E6: Heterogeneous Effects by Cities' Linguistic Distance to Latin: Dropping Cities Close to the Median Linguistic Distance to Latin (30-Years Periods)

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Low Distance to Latin X Pre	-0.100 (0.130)	-0.055 (0.119)	-0.012 (0.041)	-0.079 (0.112)
Low Distance to Latin X Post1	0.205* (0.121)	0.156 (0.107)	0.032 (0.038)	0.191* (0.105)
Low Distance to Latin X Post2	0.412*** (0.129)	0.409*** (0.112)	0.228*** (0.049)	0.378*** (0.111)
1478-1487	-0.221** (0.086)	-0.111 (0.070)	-0.024 (0.024)	-0.188** (0.073)
1488-1497	-0.093 (0.093)	-0.024 (0.075)	-0.006 (0.021)	-0.074 (0.078)
1498-1507	-0.073 (0.091)	-0.019 (0.070)	0.030 (0.026)	-0.076 (0.076)
1518-1527	0.023 (0.082)	0.047 (0.062)	-0.035 (0.024)	0.008 (0.066)
1528-1537	0.136* (0.072)	0.119* (0.064)	-0.009 (0.020)	0.125** (0.060)
1538-1547	0.327*** (0.077)	0.278*** (0.077)	0.054** (0.022)	0.278*** (0.068)
1548-1557	0.364*** (0.086)	0.266*** (0.069)	0.004 (0.031)	0.294*** (0.072)
1558-1567	0.560*** (0.113)	0.392*** (0.082)	0.077** (0.033)	0.448*** (0.090)
1568-1577	0.710*** (0.138)	0.591*** (0.091)	0.125** (0.060)	0.583*** (0.117)
City Fixed Effects	Y	Y	Y	Y
R^2	0.769	0.790	0.435	0.770
Observations	1,290	1,290	1,290	1,290

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 4 with standard errors clustered at the territory \times decade level in parentheses. Restricted to cities with some printing output over the period 1478–1577. Cities with a linguistic distance to Latin between the 40th and the 60th percentile are dropped from the sample. "Pre" refers to the pre-Reformation period 1478–1507, "Post1" refers to the post-Reformation period 1518–47, and "Post2" refers to the post-Reformation period 1548–77. The table is restricted to cities with some printing output over the period 1478–1577. Throughout, only vernacular works are considered.

Table E7: Heterogeneous Effects by Cities' Linguistic Distance to Latin: Adding Time-Varying Geographic and Historical Controls

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Low Distance to Latin X 1478-1487	-0.171 (0.210)	-0.123 (0.180)	-0.036 (0.033)	-0.118 (0.180)
Low Distance to Latin X 1488-1497	-0.234* (0.136)	-0.213** (0.093)	0.012 (0.051)	-0.164 (0.105)
Low Distance to Latin X 1498-1507	-0.057 (0.138)	-0.092 (0.130)	0.043 (0.039)	-0.045 (0.113)
Low Distance to Latin X 1518-1527	0.110 (0.119)	0.114 (0.109)	0.023 (0.045)	0.114 (0.099)
Low Distance to Latin X 1528-1537	0.168 (0.115)	0.115 (0.098)	0.086** (0.040)	0.174* (0.101)
Low Distance to Latin X 1538-1547	0.344*** (0.104)	0.308*** (0.097)	0.170*** (0.040)	0.301*** (0.084)
Low Distance to Latin X 1548-1557	0.382*** (0.140)	0.418*** (0.100)	0.184*** (0.045)	0.344*** (0.118)
Low Distance to Latin X 1558-1567	0.434*** (0.148)	0.392*** (0.119)	0.247*** (0.058)	0.410*** (0.110)
Low Distance to Latin X 1568-1577	0.418** (0.193)	0.442*** (0.159)	0.328*** (0.056)	0.381** (0.152)
1478-1487	1.436 (0.933)	0.738 (0.792)	-0.062 (0.312)	1.271* (0.761)
1488-1497	0.843 (0.716)	0.472 (0.786)	-0.391 (0.278)	0.721 (0.651)
1498-1507	-0.048 (0.648)	0.014 (0.449)	-0.249 (0.353)	-0.017 (0.581)
1518-1527	0.291 (0.613)	-0.592 (0.538)	-0.224 (0.238)	0.206 (0.501)
1528-1537	-0.029 (0.505)	-0.595 (0.484)	-0.380 (0.234)	0.011 (0.431)
1538-1547	-1.391** (0.571)	-1.213** (0.515)	-0.590** (0.243)	-1.165** (0.483)
1548-1557	-1.207** (0.603)	-1.101** (0.529)	-0.194 (0.355)	-1.287** (0.540)
1558-1567	-3.408*** (1.141)	-2.011** (0.922)	-0.783** (0.349)	-2.339*** (0.833)
1568-1577	-1.957 (1.432)	-1.578 (1.198)	-0.607 (0.430)	-1.432 (1.190)
City Fixed Effects	Y	Y	Y	Y
R^2	0.825	0.827	0.581	0.820
Observations	1,630	1,630	1,630	1,630

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 3 and additionally including interactions between geographic and historical controls and decade fixed effects. Geographic controls include the natural log of one plus a city's distance to a trade route, the natural log of one plus a city's distance to the coast, as well as a city's latitude and longitude. Historical controls include dummy variables for whether a city had a university in 1450, was Protestant in 1600, and hosted a bishop or archbishop before or in 1517. Standard errors are clustered at the territory \times decade level and are shown in parentheses. Restricted to cities with some printing output over the period 1478–1577. Throughout, only vernacular works are considered.

Table E8: Heterogeneous Effects by Cities' Linguistic Distance to Latin: Adding Time-Varying Geographic and Historical Controls (30-Years Periods)

	Ln(1+ total nr of works)	Ln(1+ total nr of authors)	Ln(1+ total nr of works from authors w/ low socioeco. background)	Ln(1+ avg nr of subject classifications)
	(1)	(2)	(3)	(4)
Low Distance to Latin X Pre	-0.154 (0.118)	-0.150 (0.102)	0.007 (0.033)	-0.109 (0.096)
Low Distance to Latin X Post1	0.207** (0.098)	0.181** (0.089)	0.093*** (0.034)	0.197** (0.080)
Low Distance to Latin X Post2	0.412*** (0.114)	0.427*** (0.096)	0.253*** (0.040)	0.378*** (0.091)
1478-1487	1.398 (1.152)	0.829 (0.933)	-0.154 (0.325)	1.251 (0.939)
1488-1497	0.668 (0.839)	0.399 (0.758)	-0.378 (0.263)	0.602 (0.736)
1498-1507	0.164 (0.725)	0.171 (0.480)	-0.170 (0.355)	0.122 (0.645)
1518-1527	0.078 (0.600)	-0.675 (0.512)	-0.378* (0.226)	0.026 (0.485)
1528-1537	-0.116 (0.485)	-0.748 (0.465)	-0.395* (0.231)	-0.038 (0.419)
1538-1547	-1.092* (0.575)	-1.024** (0.505)	-0.422* (0.247)	-0.936* (0.500)
1548-1557	-1.271* (0.672)	-1.214** (0.562)	-0.344 (0.376)	-1.361** (0.600)
1558-1567	-3.359*** (1.216)	-2.108** (0.909)	-0.797** (0.374)	-2.271*** (0.858)
1568-1577	-1.943 (1.557)	-1.597 (1.253)	-0.443 (0.426)	-1.427 (1.276)
City Fixed Effects	Y	Y	Y	Y
R^2	0.809	0.814	0.526	0.804
Observations	1,630	1,630	1,630	1,630

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: OLS regressions based on equation 4 and additionally including interactions between geographic and historical controls and decade fixed effects. Geographic controls include the natural log of one plus a city's distance to a trade route, the natural log of one plus a city's distance to the coast, as well as a city's latitude and longitude. Historical controls include dummy variables for whether a city had a university in 1450, was Protestant in 1600, and hosted a bishop or archbishop before or in 1517. Standard errors clustered at the territory \times decade level in parentheses. Restricted to cities with some printing output over the period 1478–1577. "Pre" refers to the pre-Reformation period 1478–1507, "Post1" refers to the post-Reformation period 1518–47, and "Post2" refers to the post-Reformation period 1548–77. Restricted to cities with some printing output over the period 1478–1577. Throughout, only vernacular works are considered.

F Vernacular Printing Output and Longer-Run Development

F.1 Linguistic Distance to Latin as Instrumental Variable

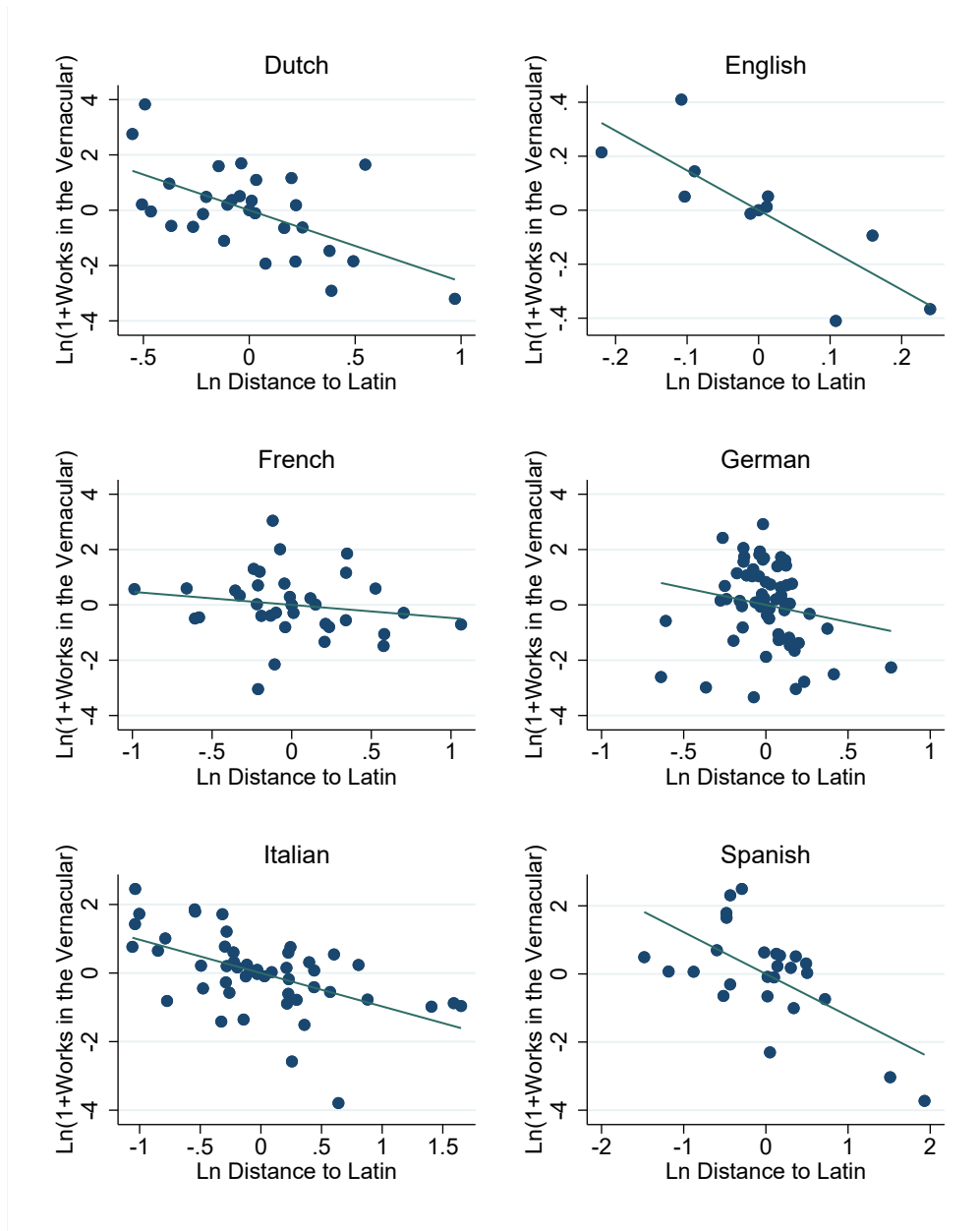


Figure F1: Cities' Linguistic Distance to Latin and Vernacular Printing Output, by Cities' Main Language

Notes: Based on equation 6, the figure shows the relationship between cities' linguistic distance to Latin and vernacular printing output by cities' main European language.

Table F1: Determinants of Cities' Vernacular and Latin Printing Output (1518–1600)

	Vernacular Works (1)	Latin Works (2)
Ln Distance to Latin	-0.388*** (0.081)	0.045 (0.104)
Ln Population in 1600	0.493*** (0.100)	0.052 (0.117)
Ln(1 + Distance to a Trade Route)	-0.059* (0.033)	-0.030 (0.038)
Ln(1 + Distance to the Coast)	0.034 (0.053)	0.059 (0.058)
Latitude	0.062** (0.031)	0.011 (0.033)
Longitude	0.002 (0.025)	0.034 (0.030)
Bishop in 1517	-0.310** (0.136)	0.377*** (0.118)
University in 1450	-0.071 (0.132)	0.494*** (0.181)
Protestant in 1600	0.468 (0.305)	-0.210 (0.323)
Ln(1 + Latin Works 1518-1600)	0.684*** (0.034)	
Ln(1 + Vernacular Works 1518-1600)		0.895*** (0.047)
R^2	0.852	0.838
N	232	232
Country Fixed Effects	Yes	Yes

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table reports OLS regression results with standard errors clustered at the territory level in parentheses. The dependent variable in column (1) is $\ln(1 +$ the number of vernacular works printed in city i over the period 1518 to 1600). The dependent variable in column (2) is $\ln(1 +$ the number of Latin works printed in city i over the period 1518 to 1600). The first control is $\ln(\text{linguistic distance to Latin of city } i)$. We next consider geographic controls, log distance to a trade route, log distance to the coast, latitude, and longitude. We then consider several historical controls: indicator variables for the city hosting a university in 1600, being Protestant in 1600, and hosting a bishop in 1517. We also control for the log of city population in 1600, as well as for $\ln(1 +$ the number of Latin works) printed in city i over the period 1518 to 1600 in column (1) and $\ln(1 +$ the number of vernacular works) printed in city i over the period 1518 to 1600 in column (2).

F.2 Robustness Analysis: Vernacular Printing Output and Upper-Tail Human Capital

Table F2: Cities' Vernacular Printing Output and Future Births of Famous People: Robustness Analysis

	<i>Famous People Born 1600–1700 (Second Stage, IV Estimation)</i>						
	<i>1600–1700</i>						<i>1400–1500</i>
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln(1 + Vernacular Works 1518–1600)	0.445*** (0.123)	0.373*** (0.100)	0.427*** (0.089)	0.309*** (0.085)	0.717*** (0.174)		0.031 (0.080)
Ln(1 + Latin Works 1518–1600)						-4.950 (10.062)	
F-Statistic on IV	10.193	33.313	16.953	20.451	24.310	0.199	12.403
N	174	195	175	162	232	232	155
	<i>Panel B: IV (First Stage)</i>						
	<i>Ln(1 + Vernacular Works 1518–1600)</i>						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Ln Distance to Latin	-0.731*** (0.208)	-0.840*** (0.168)	-0.737*** (0.208)	-0.728*** (0.198)	-0.340*** (0.076)	0.032 (0.099)	-1.447*** (0.411)
Observations	174	195	175	162	232	232	155
R^2	0.607	0.581	0.627	0.724	0.856	0.838	0.429
Excl. Cities w/ Low Printing Output	Yes	No	No	No	No	No	No
Excl. Cities w/ Greek Presence in 1500	No	Yes	No	No	No	No	No
Excl. Cities w/ Plague in 1500–1600	No	No	Yes	No	No	No	No
Excl. Cities in the HRE	No	No	No	Yes	No	No	No
Ln(1 + Latin Works 1518–1600)	No	No	No	No	Yes	-	No
Ln(1 + Vernacular Works 1518–1600)	-	-	-	-	-	Yes	-

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: This table examines the robustness of the IV results reported in column (6) of Table 1. In column (1) we exclude cities whose vernacular printing output is within the bottom 25% of the sample. In column (2) we exclude cities with a Greek presence around the year 1500. In column (3) we exclude cities with at least one plague between 1500 and 1600. In column (4) we exclude cities located in the Holy Roman Empire. In column (5), we add a city's Latin printing output as a control variable. In column (6), we use cities' linguistic distance to Latin as an instrument for cities' Latin printing output while controlling for vernacular printing output. In column (7), we consider the births of famous people in the 15th century as the dependent variable. In this specification, we control for the natural logarithm of famous people born 1200–1300 (instead of 1400–1500).