

Schaff, Felix S. F.

Working Paper

The unequal spirit of the protestant reformation: Particularism and wealth distribution in early modern Germany

EHES Working Paper, No. 239

Provided in Cooperation with:

European Historical Economics Society (EHES)

Suggested Citation: Schaff, Felix S. F. (2023) : The unequal spirit of the protestant reformation: Particularism and wealth distribution in early modern Germany, EHES Working Paper, No. 239, European Historical Economics Society (EHES), s.l.

This Version is available at:

<https://hdl.handle.net/10419/298552>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

EHES Working Paper | No. 239 | September 2023

The Unequal Spirit of the Protestant Reformation:
Particularism and Wealth Distribution in
Early Modern Germany

Felix S.F. Schaff,
European University Institute

The Unequal Spirit of the Protestant Reformation: Particularism
and Wealth Distribution in Early Modern Germany

Felix S.F. Schaff,
European University Institute

Abstract

This paper studies the impact of the Protestant Reformation on wealth distribution and inequality in confessionally divided Germany, between 1400 and 1800. The Reformation expanded social welfare, but provided it in a particularistic way to insiders only. This gave Protestantism an ambiguous character in terms of redistribution and its impact on inequality. I develop a theoretical framework of this trade-off between welfare expansion and particularistic provision, and test its implications empirically, using a Difference-in-Differences and an Instrumental Variable strategy. In line with the theoretical framework, I document that the Reformation exacerbated inequality overall, by making marginal poor people relatively poorer. The result is driven by the introduction of new particularistic poor relief policies in Protestant communities. Protestantism is an underappreciated driver of preindustrial inequality, long before the onset of industrialisation and modern economic growth.

JEL Codes: D31, H23, I38, N33

Keywords: Wealth, Poverty, Inequality, Political Economy, Protestantism, Welfare, Germany

¹Corresponding Author: Felix S.F. Schaff (felix.schaff@eui.eu)

Notice

The material presented in the EHES Working Paper Series is property of the author(s) and should be quoted as such. The views expressed in this Paper are those of the author(s) and do not necessarily represent the views of the EHES or its members

“And Jesus said to his disciples, Truly I say to you, It is hard for a man with much money to go into the kingdom of heaven.” [...] “Give what property you have in exchange for money, and give the money to the poor”

— *New Testament Matthew 19:23, Luke 12:33*

1 Introduction

What is the impact of religious confession on the distribution of wealth and inequality? Since Max Weber’s seminal book (1930) “The Protestant Ethic and The Spirit of Capitalism” scholars have debated the socio-economic impact of the Protestant Reformation. While considerable attention has been given to this event’s effect on economic growth, its impact on inequality has only been studied marginally (see Becker et al. 2016). This lack of knowledge clashes with the key role recent work attributes to ideology and institutions in explaining how high levels of economic inequality came about and persisted over the long run of history (Abramitzky 2008, Piketty 2020, Alfani 2021). Moreover, religion might be a key variable for understanding preferences towards redistributive policies aimed at tackling inequality today (Alesina and Giuliano 2010, Enke et al. 2022).

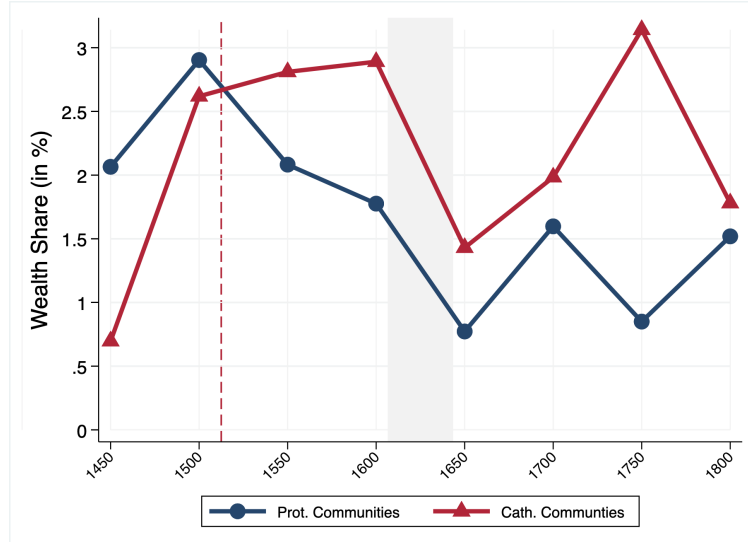
Much discussion about the Reformation’s socio-economic impact centres on the expansion of public goods provision, especially in terms of social welfare including more generous poor relief (Lindert 2004, Kahl 2009, Dittmar and Meisenzahl 2020). This line of argumentation implies more redistribution of economic resources to poor people, and a more egalitarian distribution of income and wealth under Protestantism. And yet a key dimension of redistributive policies is how universal or particularistic the provision of social welfare is. Societies can choose to redistribute only between people that are socially close or also to distant strangers (Enke et al. 2022), with obvious implications for inequality. This generates a trade-off: if Protestant redistribution did not go as far as marginal poor people (see Hartung 1989, Jütte 1994), it probably increased the gap between poor strata and the rest of society, thus increasing economic inequality. Figure 1 suggests that this might have been the case. It illustrates a divergence in the wealth share of poor strata — defined as the bottom 20 percent of the distribution (Dollar and Kraay 2002) — in Protestant and Catholic communities in my dataset. (The vertical red line represents the beginning of the Reformation period, the grey box the Thirty Years’ War.¹)

In this paper I first develop a theoretical framework of this trade-off. To test the predictions deriving from that framework, I then construct a panel database on local religious confession, wealth distribution and inequality in German communities from 1400 to 1800. It is based on c.380,000 household wealth observations collected from archival tax records and secondary sources.² The data

¹The Thirty Years’ War (1618-48) was the most violent and destructive conflict in European history. It started as a conflict over whether Protestantism or Catholicism is the “right” Christian faith. It took place mostly in Germany, where it killed about 40 percent of the population.

²Since household wealth observations have to be aggregated at the community-year level to calculate inequality measures, the final dataset consists of 368 observations. Note that all studies of preindustrial inequality employ relatively small datasets, because of the large amount of household-level data needed to reconstruct one distribution for a given point in time. For example, Milanovic (2018) employs 41 observations. Further note that household-level inequality cannot be studied with any of the existing datasets of preindustrial German towns.

Figure 1: Wealth Share of the Bottom 20% in Protestant and Catholic Communities



Notes: Values were collapsed into 50-year intervals and represent half-century-averages. To avoid that communities with more observations dominate the trend, every community has the same weight in the average. Because of the uneven number of years and the low number of observations in 1400, I collapsed the values for the years 1400 until 1450 into one data point. The vertical red line represents the beginning of the Reformation period, the grey box the Thirty Years' War.

make it possible to observe wealth in a highly disaggregated way across the distribution. They allow me to study inequality — that is, *relative* and not *absolute* wealth distribution — within Protestant and Catholic communities.

With its religious heterogeneity among communities, Germany is the ideal testing ground for shedding light on the redistributive effect of the Protestant Reformation.³ Causal identifications stems from the fact that conversion to Protestantism was superimposed on communities by their local rulers following the Peace of Augsburg (1555). I complement the main analysis with flexible difference-in-differences estimates, which allow me to formally test the assumption of parallel trends. To provide additional evidence that the estimates are not driven by self-selection into Reformation adoption and to address potential parallel trends assumption violations I also employ an instrumental variable strategy. I use the distance to Wittenberg as instrument, as previously employed by Becker and Woessmann (2009), Cantoni (2015) and Becker and Pascali (2019). It exploits the fact that the Reformation was more likely to be adopted the closer a community was to the movements' starting point.

I find that the Reformation increased economic inequality, by making poor people poorer relative to the rest of the population. This result is statistically significant and economically meaningful: the bottom fifth of the population lost about 39 percent of its pre-treatment wealth share. The finding is robust to a rich set of controls, including economic growth and warfare, and I do not find evidence of significantly different pre-trends. Changes in other parts of the wealth distribution do not mechanically drive the changes in the bottom tail, and the main result holds when dropping various parts of the dataset, including regions where observations are geographically clustered.

³The geographical area this paper is interested in, is the inner part of the Holy Roman Empire of the German Nation that was subject to the Imperial constitution. As a shorthand I refer to the Empire as Germany.

The main empirical pattern of poor strata losing significant wealth shares was not short-lived. It endured until the end of the early modern age when industrialisation began. Interestingly, I find no significant evidence that the Reformation had an effect on top wealth shares. This is contrary to what one would expect if inequality were driven by factors like economic growth, capital accumulation or higher upper-tail human capital, as these tend to increase inequality from the top of the distribution. Similarly, I do not find evidence for significant changes to the wealth shares of middling classes. I interpret these results in the following way: there were concentrated losses at the bottom of the wealth distribution, but dispersed gains for the rest of society. The IV estimates are slightly larger than the main difference-in-difference results. However, I document that the instrument was not related to higher inequality before, but only after the beginning of the Reformation. This suggests that, conditional on several covariates, the exclusion restriction most likely holds.

I argue that the findings reflect the Reformation's shift towards a more particularistic provision of poor relief, which outweighed the expansion of social welfare available to some insiders. In Protestant places strangers and able-bodied but non-working poor were excluded from the local Christian community, and declared as non-eligible for receiving poor relief. The consequence was a reduction in the supply of poor relief to the excluded groups and a decrease of transfers from better-off to marginal poor people. These new low-redistribution policies left behind the bottom of the poor in Protestant society, while the Catholic Church stuck to its universal approach to charity (see opening quotes) during the Counter-Reformation. Moreover, the particularistic Protestant poor relief system came with a whole set of policies and implications that were also particularistic in nature: the prohibition of begging, the disincentivising of almsgiving, the stigmatisation of the poor in the labour market. These reduced the share of economic resources held by marginal poor people in a way that was analogous to poor relief in the strict sense and exacerbated its effect. In consequence, the Reformation reshuffled the bottom end of the income and wealth distribution, making some poor even poorer and increasing inequality.

I provide several pieces of evidence for the plausibility of this hypothesised mechanism. First, within the poor strata of society it was the bottom decile, the poorest of the poor, that lost the largest share in Protestant places. This finding is consistent with the notion that Protestant communities excluded people at the very margin of society from poor relief. Second, I collect additional data on the introduction of so-called "church ordinances" in many newly Protestant localities. These laws regulated poor relief, among other welfare tasks, in a particularistic manner. They legally excluded community outsiders, but also restricted begging and introduced poor badges that would stigmatise paupers in the labour market. All these particularistic measures redistributed economic resources away from marginal poor people. Accordingly, church ordinances are found to explain substantial parts of the negative effect of the Reformation on poor strata. The effect is particularly large in places where the laws explicitly restricted poor peoples' opportunities for begging. Third, the Reformation also reduced poor relief by shutting down monasteries, which for centuries had performed universal redistribution to the poor for the Catholic Church. I show evidence of this with an additional dataset on the closure of monasteries during the Reformation.

The paper relates to several strands of literature. First, it improves our understanding of the deep historical causes and persistence of economic inequality and poverty. Recent research has revealed a striking empirical pattern: contrary to a conventional “Kuznetsian view” (see Kuznets 1955), inequality in Europe did not start to grow with the beginning of industrialisation in the eighteenth or nineteenth centuries, but increased almost constantly since about the sixteenth century. Much of the high levels of inequality observed in the early twentieth century might have preindustrial roots (Piketty 2020, Alfani 2021). The leading explanations for this preindustrial inequality rise have stressed the role of economic growth (van Zanden 1995, Puga and Treffer 2014), warfare and increasing fiscal capacity of emerging states (Alfani 2021, Schaff 2023), feudal institutions (Galor et al. 2009) and demographic expansion (Clark 2007, Milanovic 2016).⁴ Instead, my results emphasise the role of ideological and institutional change (see also Abramitzky 2008, Piketty 2020). I investigate the impact of a major, but hitherto overlooked, socio-political shock on inequality in the long run of history, namely the Reformation. My findings suggest that the emergence of a new, more particularistic poor relief system, triggered by Protestantism, influenced redistribution, and consequently increased poverty and inequality. The findings connect to studies employing contemporary outcomes with the aim to show that Protestants have lower preferences for redistribution and experience higher income inequality today (Guiso et al. 2003, Alesina and Giuliano 2010, Basten and Betz 2013). My results suggest that today’s “Protestant inequality” might have its historical origins in the sixteenth century and persisted until at least the nineteenth century.

Second, the paper contributes to the classical debate in the social sciences about the economic impact of the Reformation in the long run of history (see Becker and Woessmann 2009, Cantoni 2015, Cantoni et al. 2018, Kersting et al. 2020). Much emphasis has been put on the Reformation leading to an expansion of public goods provision, especially in terms of social welfare and poor relief (Lindert 2004, Dittmar and Meisenzahl 2020). This implies more redistribution to the poor and less inequality. My paper theoretically shifts this focus, taking into account a second dimension of redistributive policies (see Enke et al. 2022), namely that social welfare was provided in a more particularistic way under Protestantism. I show that the new Protestant trade-off between generosity and universal or particularistic provision of social welfare increased inequality, by making poor people economically worse off, relative to the rest of society. I also show that the closure of monasteries during the Reformation did not only redistribute economic resources between rulers and the Church (Cantoni et al. 2018), but also affected the distribution among individuals in an inegalitarian way. However, I do not criticise work that suggests that Protestantism had positive effects on the economy, for example through education or the provision of public goods (Becker and Woessmann 2009, Dittmar and Meisenzahl 2020). I also do not exclude that Protestantism’s stricter poor relief and lower equity led to other efficiency gains, such as a higher incentives to work (see Spenkuch 2017). Due to data limitations this paper cannot look at absolute wealth, but only relative wellbeing.

The next section presents the theoretical framework of the trade-off between generosity and partic-

⁴This literature is closely related to a larger literature on the drivers of inequality growth and decline in history, and the long-term economic, social and political consequences of inequality (see Persson and Tabellini 1994, Acemoglu et al. 2005, Piketty 2014, Galor and Moav 2004, Scheidel 2017, Ager et al. 2021).

ularistic provision of poor relief under Catholicism and Protestantism, and formulates a hypothesis about the implications of the Reformation for wealth distribution and inequality. Section 3 describes in detail the data. Section 4 presents the empirical strategies to test the main hypothesis, reports the main results of the paper and discusses limitations. Section 5 attempts to disentangle some of the mechanisms at work, and Section 6 concludes.

2 Redistribution under Catholicism and Protestantism: Historical Evidence and Conceptual Framework

To structure my analysis of the implications of Protestantism for the distribution of wealth and inequality, I first provide historical evidence of the Reformation’s implications for social welfare and the treatment of poor strata in society. Based on that evidence, I then model the Reformation as a shift in the trade-off between per capita redistribution and the social distance towards those people that were eligible to receive support: the Reformation led to an expansion of social welfare including poor relief to fellow Christian and community members (Lindert 2004, Kahl 2009, Dittmar and Meisenzahl 2020), but it did so in a particularistic manner to community insiders only (see Enke et al. 2022).⁵ This made the Reformation highly ambiguous in its implications for redistribution to the poor and inequality.

2.1 The Catholic Economy of Salvation

In the Catholic economy of salvation, the poor played a central role. They represented Jesus Christ, and were glorified as “God’s best friends” and the inheritors of the Kingdom of Heaven. Helping the poor was a “good work”, a moral duty that erased sinful behaviour on behalf of the donor. If at death accumulated good works outweighed sins, the good Christian could be saved from purgatory and went to heaven (Kahl 2009: 270). This ideology created powerful incentives for Catholics and church organisations to make taking care of the poor a priority.

Poor relief practice reflected this appreciation of the poor. It was based on private almsgiving and poor relief by a variety of mostly uncoordinated ecclesiastical organisations, such as monasteries, hospitals (*Spital*), churches and confraternities. Poor relief was among their principal tasks. The Catholic Church was arguably the biggest redistributor in human history, channelling money to its own elites (see Schilling 1994: 97), but also to poor people: between one third and one fourth of church income went to support the poor (Kahl 2009: 269-271). A poor person could make ends meet, through a combination of payments and offerings from several charitable organisations, private charity, and begging. Catholic welfare provision was thus relatively inclusive, embracing a variety of needs indiscriminately and providing relief to the poor universally (Ackels 1984: 100, Jütte 1994: 138). The poor continued to have this elevated role in the Catholic worldview also after the Reformation began. The Catholic Church was critical with what it saw as Protestant

⁵For example, evidence from contemporary surveys suggest that communitarian respondents oppose federal welfare and redistributive programs like universal health care. Universalists, instead, do not mind that federal redistribution is impersonal and is extended to socially distant people. Interestingly, religion has been found to be a key determinant of the degree to which a society tends to be more universal or particularistic in its welfare provision and redistribution (Enke et al. 2022).

stigmatisation of the poor and expressly stuck to its universal approach to charity during the Counter-Reformation (Jütte 1994: 125-138, Kahl 2009: 279-280).

Where the Reformation was introduced, the Catholic Church and its welfare provision literally disappeared. For example, the monasteries through which the Old Church had provided poor relief were simply closed. Confiscated assets were freed from the Catholic Church for secular use, but often went into the coffers of local rulers, for instance to build palaces or wage war (Cantoni et al. 2018).⁶ Monasteries and other ecclesiastical organisations could not fulfil their redistributive function anymore. Poor relief had to be replaced by an entirely new system in Protestant places.

2.2 The Reformation: Poor Relief as a Public Responsibility and the Expansion of Social Welfare

Where the Reformation was introduced, the Christian approach to poor relief changed drastically. It involved a new trade-off between per capita redistribution and social distance to the recipient. Luther's and other reformers' reinterpretation of Christianity provided a radically new vision of how to deal with poverty, organisationally and ideologically. First of all, reformers envisioned poor relief to be a secular task now, performed by communities instead of ecclesiastical organisations. This followed directly from Luther's "doctrine of the two kingdoms", which postulated the complete separation of the spiritual from the secular realm (Scribner and Dixon 2003: 36). In 1521 Luther and Karlstadt formulated the first Protestant poor law, for the city of Wittenberg, which was the legal basis for a new poor relief system. This first poor law subsequently served as a blueprint that was adopted throughout Protestant Germany when *secular* rulers decreed the introduction of new poor relief systems in their territories. The centrepiece of the new policy was to create for the first time a budget for welfare tasks in Protestant towns and villages, the so-called "poor chest" (*Armenkasten*), to give systematic support to the poor. Note that it were not just cities that provided public poor relief. Villages fulfilled that task in an analogous way and had "poor chests" too (Blickle 2015: 160). The chest was envisaged to be financed with a mix of endowments of secularised church property, and voluntary contributions from the community. Luther himself was involved in practically setting up the new institutions of poor relief, for example, in the small Saxon community of Leisnig in 1523. Such community-based systems of poor relief remained in place until about the late eighteenth or early nineteenth centuries (Jütte 1994: 106-109, 1996: 392, Kahl 2009: 272). The introduction of the new welfare system was not a gradual separation between church and state, but outright a substitution of the Old Church.

Under this new system, poor relief became for the first time a public responsibility: it became society's duty to take systematically care of fellow Christians, and the poor were entitled to receive a minimum standard of living from the chest. It usually included foodstuffs, fuel, clothing and other goods, and up to several florin of cash payment (see Jütte 1994). It was also common that communities gave more generous alms (*reiche Almosen*) to certain poor groups that had a higher social status within the community.⁷ In this sense, the new Protestant poor relief system was a

⁶For instance, the Duke of Saxony closed almost all monasteries in his territory, through which he obtained the immense sum of about 150,000 florin (Cantoni et al. 2018: 2059).

⁷For example in the merchant-city Nuremberg, those that received more generous alms were impoverished mer-

generous expansion of social welfare (Lindert 2004, Kahl 2009, Dittmar and Meisenzahl 2020).

2.3 Particularistic Provision of Poor Relief under Protestantism

Generously expanding social welfare was not the only novelty introduced by Luther and his fellow reformers. Protestantism also provided poor relief in a less universal way. It introduced two mechanism that placed *part of* the poor population outside the local community of “good Christians”, therewith excluding them from receiving social welfare. The first mechanism was to exclude all people that were not legally part of the local community — the non-residents — as it was now the community’s responsibility to deliver poor relief. Poor people without citizenship were denied access to communal poor relief and were turned away (Ogilvie 1997: 45-57, Battenberg 1991: 51-55).⁸ This exclusion was made legally quite explicit. For example, in the chest law that introduced the new Protestant poor relief system in Leisnig in 1523 — under the guidance of Luther himself — a whole chapter was dedicated to “Rejection of the Burden from Strangers” (my translation) (Sehling 1902: 601). This particularistic approach meshed well with the material interests of communities. The exclusion of poor strangers from secular poor relief enabled communities to keep low the number of recipients that burdened the communal budget (Ogilvie 1997: 50-51).

The second mechanism by which Protestantism excluded a certain sub-group of the poor from poor relief was by declaring them as undeserving or “bad Christians”, and therefore non-eligible for social welfare.⁹ This sub-group were the non-working but able-bodied poor, such as those that could not find a job.¹⁰ Their poverty was “not a misfortune to be pitied and relieved, but a moral failing to be condemned” (Tawney 1926: 230). Only poor individuals that were strictly unable to work, such as invalids, children or old people, were considered “deserving” of community relief under the new Protestant welfare system. “Who does not work shall not eat” was the guiding principle, so able-bodied “fake beggars” were excluded from communal help (Jütte 1996: 392, Kahl 2009: 271). This particularistic approach resulted directly from the reformed Christian ideology. Luther wanted good Christians to work, a contribution to the community that was pleasing to God. Instead idleness and begging were to be disdained.¹¹ That logic implied sympathy for the deserving poor, but placed the able-bodied non-working undeserving poor outside the the local community of “good Christians”.

It is not possible to quantify the number of poor people excluded from public relief in Protestant chants (Hartung 1989: 170).

⁸Communities in preindustrial times were often hostile to outsiders in general, and did not allow them to participate politically, practice an occupation, or use other public goods. Becoming member was difficult, especially for poor people who did not fulfil the minimum wealth criterion (Ogilvie 1997: 45-57).

⁹See Hudson and Coukos (2005) for an account on how the exclusionary attitude of Protestantism towards poor continued to shape peoples’ preferences for redistribution and policy debates in the US in the 19th and 20th centuries.

¹⁰The exclusion of able-bodied unemployed from poor relief was likely based on a flawed understanding of the labour market. To presume that every able-bodied non-working poor person was unwilling to work amounts to assuming an unlimited supply of jobs. Yet the economy in sixteenth-century Germany was declining in terms of per capita output (Alfani et al. 2022: 107), and burdened with all kinds of entry barriers that made it hard to enter the labour market, such as having connections to guild masters (see Ogilvie 1997). In consequence, unemployment was in actuality widespread (Schilling 1994: 378).

¹¹Religious reformers like Luther, Zwingli, Karlstadt Bugenhagen, Hyperius and Bucer differed on many points in their theology, but they were surprisingly united in their negative views on poverty (Jütte 1996: 392).

communities systematically, but historians have long conjectured that their numbers were substantial (Hartung 1989: 168). Scattered evidence comes from a few cases where community-level estimates of the share of poor relief recipients in relation to the total population are available (see Table 1). The evidence is patchy, but across four centuries after the beginning of the Reformation, the share of public poor relief recipients was lower in almost all Protestant cases compared to Catholic ones. The difference in the average share of recipients is considerable, suggesting a substantial divergence in how exclusionary the two confessions were.

Table 1: Population Share of Recipients of Public Poor Relief

Community	Year	Confession	Recipients as % of Population
Trier	1591	Catholic	8.3%
Trier	1600	Catholic	34.6%
Trier	1625	Catholic	27.1%
Trier	1649	Catholic	16.9%
Solothurn (villages)	1768	Catholic	22.5%
Würzburg	1794	Catholic	4.3%
Cologne	1799	Catholic	8.2%
Cologne	1816	Catholic	38.0%
<i>Average Share in Catholic Communities:</i>			20.0%
Nuremberg	1531	Protestant	1.2%
Ulm	1531	Protestant	2.8%
Frankenberg	1533-42	Protestant	4.2%
Frankfurt a.M.	1539	Protestant	3.6%
Augsburg	1550	Protestant/Mixed	5.5%
Augsburg	1574	Protestant/Mixed	4.3%
Zürich (villages)	1590	Protestant	4.5%
Augsburg	1610	Protestant/Mixed	7.2%
Berlin	1665	Protestant	2.0%
Berlin	1799	Protestant	7.2%
Berlin	1860	Protestant	6.0%
<i>Average Share in Protestant Communities:</i>			4.4%

Notes: Data on poor relief recipients for Cologne and Berlin from Fischer (1982: 58, 83) and Jütte (1994: 54), for Trier from Ackels (1984: 94), for Augsburg from Clasen (1984: 89) and Röck (1989: 169), for Nuremberg and Ulm from Hartung (1989: 168, 172), for Würzburg and Frankenberg from Jütte (1996: 388) and for the villages of Solothurn and Zürich, and Frankfurt a.M. from Jütte (1994: 54). For Solothurn and Zürich the data refer to the villages surrounding the towns. I include these Swiss localities in the table because Switzerland was still officially part of the Holy Roman Empire at the time of the Reformation.

2.4 Beyond Poor Relief: Second Order Effects of Protestant Particularism

The particularistic Protestant approach to poor relief came with a whole set of policies and implications that were also particularistic in nature, or direct consequences of the particularistic poor relief system, and that had similar distributive effects. One of these policies was the restriction and often outright prohibition of begging. Since communities took adequate care of “their” poor, the reasoning went, those that were still begging must have been undeserving outsiders. Moreover, it was assumed that beggars could just take up a job if they wanted and simply needed a strong enough incentive to do so (Kahl 2009: 274-278). Prohibiting begging was therefore not illogical. Yet it closed an important income channel for many poor people, reducing their share of economic

resources.

Additionally, restrictions to begging went hand in hand with the elimination of the incentives for donors to engage in private almsgiving. Protestantism denied private almsgiving its theological function of avoiding purgatory and declared it as outright wasteful. Instead, better-off individuals were urged to donate only to the common chest, in order to redistribute money more “efficiently” to deserving community members (Jütte 1996: 396-397, Scribner and Dixon 2003: 58). In other words, there was a shift in norms about almsgiving that was associated with lower private charity, which closed an income channel for poor people.

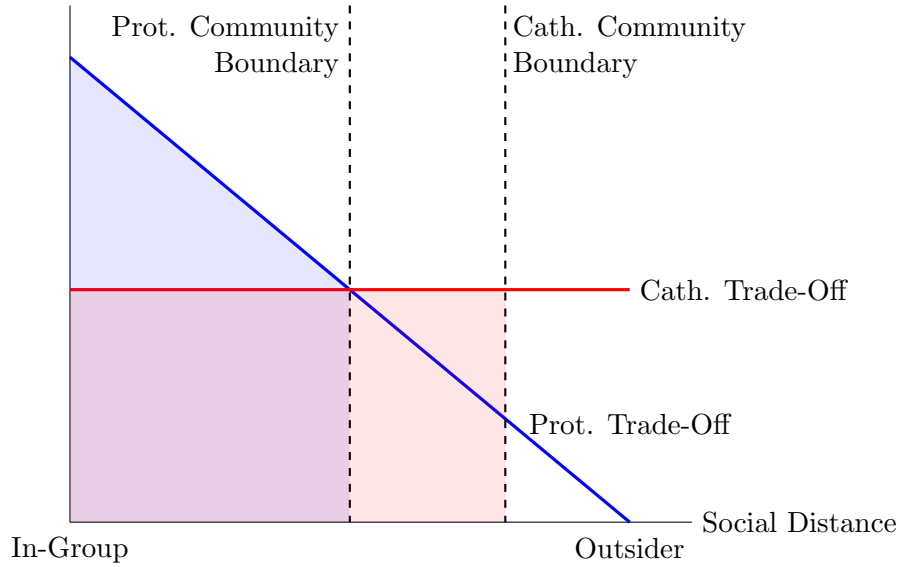
Figure 2: Poor Badge from Sixteenth-Century Nuremberg



Source: Germanic National Museum Nuremberg.

Ultimately, the particularistic Protestant approach implied a stigmatisation of poverty. In many Protestant communities poor relief recipients had to identify themselves wearing a poor badge, to differentiate between insider and outsider. This put poor people in the same social category of non-trustworthy people like prostitutes or lepers (Jütte 1994: 161-161; see Figure 2). Stigma deriving from discriminatory symbolic policies can prevent the poor from entering the labour market and improving their economic situation through work. This redistributes income and wealth away from the poor (Dewan and Wolton 2019). For example, in Nuremberg and Augsburg the poor badges reportedly discredited the poor with potential employers, while no beggar emblem was necessary in Catholic Trier. It was enough that the priest attested one's need to receive poor relief (Hartung 1989: 169, 171, 174, Ackels 1984: 80), thus avoiding discrediting stigmatisation, giving the poor a better chance in the labour market and channelling a larger share of income and wealth to them.

Figure 3: Framework: Redistribution under Catholicism and Protestantism
p.c. Redistribution



Notes: In this framework the redistributive impact and thus inequality, either under Catholicism or Protestantism, was the result of a trade-off: between, the generosity of redistribution (y-axis) and, the social distance between the providers of poor relief and those needy people that were eligible to receive support, that is the universalism-particularism dimension (x-axis).

2.5 Conceptual Framework

In Figure 3 I develop a simple theoretical framework that summarises the historical evidence. I loosely follow arguments made by Dittmar and Meisenzahl (2020) on welfare expansion due to the Reformation, and by Enke et al. (2022) on the universalism-particularism dimension of redistributive policies. The framework refers to all poor or needy people living in a community, but not all inhabitants had membership, that is, citizenship. The redistributive impact and thus inequality, either under Catholicism or Protestantism, was the result of a trade-off: between, first, the generosity of redistribution (y-axis). And, second, the social distance between the providers of poor relief (the Church under Catholicism, and secular communities under Protestantism) and those needy people that were eligible to receive support, that is, the universalism-particularism dimension (x-axis). Under the assumption that more socially distant people were also poorer, then the more the correlation between these two factors tends towards zero, or even a positive value, the more economic resources are redistributed towards poor people. The result is a higher share of economic resources in the hands of poor people, and lower inequality. Conversely, the more negative the correlation is, the lower is redistribution to needy people, lowering their share of economic resources and increasing inequality. It is important to recognise that this trade-off refers to the entire poor relief system coming with one or the other confession, including the provisions from ecclesiastical and secular authorities.

Catholicism provided a medium amount of per capita redistribution through poor relief. Yet it did so in a universal manner, as indicated by the flat slope of the red line: poor people from in- or outside the local Christian community received support. The boundary of the Catholic community was relatively wide and not limited to local community insiders. The total amount of redistribution,

indicated by the red and purple areas, was distributed over a large group of needy people.

The advent of the Reformation shifted this trade-off in newly Protestant communities, as indicated by the declining blue line. Per capita redistribution through poor relief may have been more generous for some people under Protestantism than under Catholicism, and as mentioned even particularly generous for community members with higher social status.¹² But redistribution was provided in a particularistic manner, as represented by the tighter Protestant community boundary: only members of the local Christian community, to whom the social distance was short, received support. Outsiders, such as strangers or those able-bodied non working poor that were considered religiously undeserving “bad Christians” did not receive poor relief. This does not mean that outsiders were non-Christians. Outsiders were most likely Christians too, even Protestant. They were just not considered good enough Christians to deserve support, for instance because they did not work. The total amount of redistribution under Protestantism, indicated by the blue and purple areas, was thus more concentrated compared to Catholicism. It left out a fairly large portion of the poor population.

It is important to recognise that this redistributive effect was not just the result of poor relief in the narrow sense. As mentioned, Protestantism also restricted begging, disincentivised private almsgiving and stigmatised poor people, making it harder for them to enter the labour market. All these factors were part of or were implications of the particularistic Protestant poor relief system and had analogous distributive implications.

Note that the framework is not making a strong claim about the total amount of resources employed for poor relief either under Protestantism or Catholicism. We lack systematic information about the details of local welfare budgets for the time that could credibly support one or the other claim. In other words, we do not know the size of the areas under the coloured lines. I therefore assume that the total amount of redistribution was approximately equal under both systems. But even if Protestantism redistributed a larger amount of resources, making poor strata richer in *absolute* terms compared to Catholicism, the particularistic character of this redistribution would still reduce poor peoples’ *relative* share, thus increasing inequality.¹³

2.6 Hypothesis: Implications of the Reformation for the Distribution of Economic Resources and Inequality

The shift in the trade-off between generosity of redistribution and the universalism-particularism dimension of who is eligible to receive social welfare suggests the following hypothesis regarding

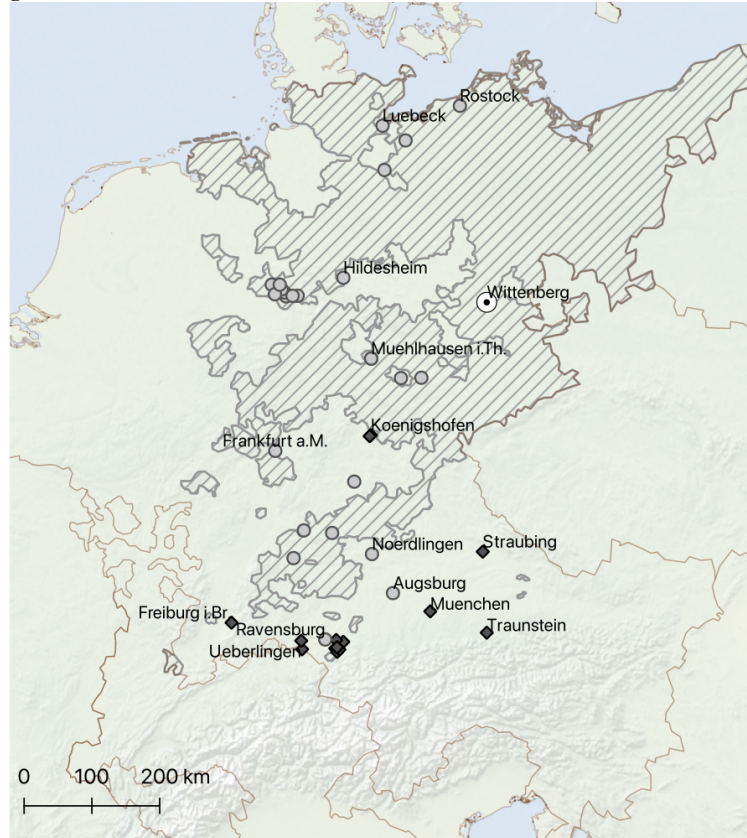
¹²In reality, the Protestant line consisted probably of several downward steps, but for clarity I represent it as linearly declining.

¹³If anything, it would be more reasonable to assume that the total amount of redistribution was higher under Catholicism. First, because the Catholic Church redistributed between one third and one fourth of its income to the poor (Kahl 2009: 269-271), while the assets that generated this income were usually confiscated by local rulers in Protestant territories, and employed to build palaces or wage war instead (Cantoni et al. 2018). Second, in the conception of reformers, the lack of incoming-generating assets to finance poor relief was supposed to be compensated with private charity. But the little available evidence suggests that the propensity to engage in private charity in Protestant communities was low. For example, the chest keeper in sixteenth century Protestant Nuremberg received private donations of only about 10 to 17 gulden per month, a ridiculously small amount for such a large and rich merchant city (Hartung 1989: 170).

redistribution and inequality following the Reformation. Protestantism implied a reduction in the supply of poor relief to the excluded groups and a decrease of transfers from better-off to poor strata. The new low-redistribution policy came with a whole set of policies and implications that were also particularistic in nature and left the bottom of the poor behind in Protestant society. In consequence, the Reformation reshuffled the bottom end of the income and wealth distribution, making some poor even poorer relative to the rest of society. The inequality data employed in the empirical analysis come from property taxes,¹⁴ which implies that any income above subsistence is covered. Given these data, I would expect to observe a reduction in the wealth shares of poor strata, particularly among the poorest of the poor in Protestant places. This reshaping of the bottom end of the wealth distribution increased the overall gap between the poor and the rest of the population, thus increasing inequality.

3 Data

Figure 4: Protestant and Catholic Communities in the Dataset



Notes: Dots (light grey) correspond to Protestant communities and squares (dark grey) to Catholic communities, by 1600. Wittenberg is not part of the dataset. Borders of the Holy Roman Empire around 1545 and Protestant areas around 1559 from Alfani et al. (2022). Selected cities labelled. Not all communities are visible because of visual overlap on the map, and not all, sometimes tiny, Protestant areas are visible because of the scale of the map.

Figure 4 provides an overview of the 18 Catholic and 25 (eventually) Protestant communities

¹⁴The tax base usually included real estate, crops, animals, cash money, loans, household goods and other property. Property is a stock, but it also gives information about income flows in so far as not all income has been used for subsistence.

included in the analysis, and the major Protestant areas in the mid sixteenth century. As can be seen from examining the geographical distribution of communities, there was a concentration of Protestant places closer to the town of Wittenberg in the north-east and centre of Germany, while Catholic communities tended to lie in the south. There were Catholic places in the north and north-west, such as Cologne or Münster, but they are less frequent and considering the north of Königshofen, none is part of the dataset. I will show that the clustering of some of the communities does not drive the results (see the Appendix for results that take concerns about spatial dependence into account).

I construct data on wealth distribution and inequality at the community-year level (in steps of 25 years).¹⁵ The Appendix provides a detailed discussion of the dataset and how it has been constructed from the original sources. Economic inequality is expressed in different ways: the wealth percentiles of upper, middle and lower segments of the population and Gini coefficients. For the 43 rural and urban communities within the borders of the inner part of the Holy Roman Empire that are covered, at least one observation before and one after the beginning of the Reformation is available, resulting in an unbalanced panel. Not all regions are equally well covered and especially the centre-west is underrepresented. The selection of a place into the dataset simply depended on whether or not the relevant archival documents stood the test of time since their creation and are still available today. Attrition is therefore most likely to be random and not correlated with the error term of my estimation model. The Difference-in-differences and IV research designs account fairly well for selection bias.

Overall, a total of 368 observations of wealth distribution and inequality are available between 1400 and 1800. Calculating distributional measures requires lots of information. Every locality-year observation corresponds to a distribution, and each distribution requires information about each household in a locality in a given year. The distributions used in this analysis are based on information about more than 380,000 households (see the Appendix for an illustration of how the distributions needed to calculate inequality indicators have been obtained from the underlying household-level data).¹⁶ Limitations of the analysis deriving from the relatively small dataset are discussed below.

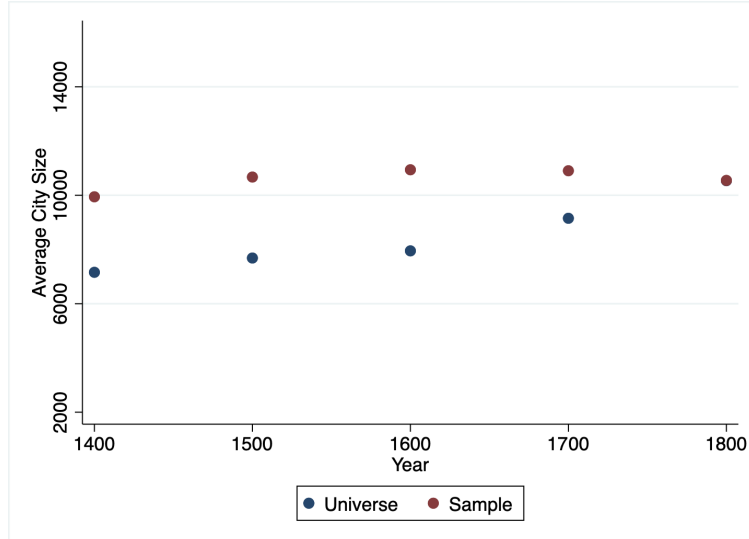
Unfortunately, for the pre-census age there exist only very few statistics of the universe of German towns to which I could compare my sample, and to my best knowledge no statistics for the universe of villages. Figure 5 compares the average population size of the urban part of my sample to all cities and towns included in an upgraded version of the “Bairoch-dataset” (Bairoch et al. 1988).

¹⁵The dataset is a substantially extended version of the one built by Alfani, Gierok and Schaff (2022). It also differs from Schaff (2023) in that the main outcome variable has been newly constructed from the household-level wealth information, namely the wealth share of the bottom 20 percent of the population, but also other outcomes such as the Top 1 percent, the middling 40 percent and the Top 90 percent wealth shares.

¹⁶Unfortunately, it is not possible to use the household-level data to construct a large individual-level panel. The main problem is that we cannot convert wealth estimates across time into a common currency for all the communities. The lack of convertibility was a consequence of preindustrial Germany’s political fragmentation, which led to a plethora of local currencies. Even the most comprehensive accounts of currencies and conversion rates do not even come close to making it possible to make wealth estimates comparable across communities and time (Alfani et al. 2022). This is not a problem for calculating the distribution of wealth, but it makes the a comparison of wealth levels impossible. Moreover, the sources give information about wealth, but rarely about other household-level characteristics. One would lack individually varying covariates in a large panel.

My sample is somewhat larger on average, but not much, and this difference gradually disappears towards 1800. Moreover, the share of Protestant places in my dataset (58.1 percent) is only slightly below the average of places recorded in the *Deutsches Städtebuch* (68.4 percent). To address potential concerns about the small sample size causing uncertainty of statistical inference, the Appendix reports p-values and confidence intervals obtained with Wild Cluster Bootstrapping. I also report results obtained from an extended sample, including localities that were observed only before or after the beginning of the Reformation.

Figure 5: Population in German towns in “Bairoch dataset”: Sample vs. Universe



Notes: Century-averages of the population in towns in an extended version of the “Bairoch-dataset” over the period of analysis. The “Universe” are all towns included in the extended Bairoch-dataset. The “Sample” includes only those towns that are included in my dataset. For details of the calculations see the Appendix.

The data are based on tax registers or wealth lists prepared to levy taxes. These were locally collected property taxes levied on households of citizens and non-citizens. The tax registers give a fairly accurate picture of the distribution of property among households in communities, subject to some limitations discussed below. The tax base was the totality of mobile and immobile assets, such as real estate, crops, animals, cash money, loans, household goods and other property. The variety of taxed items make sure that all social classes are represented in the registers. Since premodern tax systems did not systematically distinguish between wealth and income, but simply aimed at capturing the overall economic capacity of a taxpayer, also elements of income (above subsistence) were effectively taxed. Yet real estate was the most important asset class everywhere. Premodern tax systems did not distinguish between personal and business wealth, and subsumed all assets at the household-level. The use of these tax sources to study the distribution of wealth over the long-term and in comparative perspective is an established tradition in German historiography (Isenmann 2014, Alfani et al. 2022). Analogous sources have been used in studies of other regions in preindustrial Europe (van Zanden 1995, Alfani 2021). The panel-regression setup makes it possible to limit the analysis to within-community variation, and account as good as possible for the possibility that differences in taxation practices between localities could systematically influence the measured wealth distribution.

The data mostly refer to the distribution of wealth (or capitalised income). Studying the distribution of wealth is interesting in itself, even in the industrial period (see Piketty 2014). Yet wealth distribution is also the best, and usually the only proxy for trends in income distribution in preindustrial, predominantly agricultural economies, like Germany, where income mainly derived from land (Alfani 2021). The use of wealth data implies that interventions affecting specifically income, such as taxation or redistribution through poor relief, likely took time until they affected the wealth distribution significantly. Potentially, the wealth distributions have a bias towards the middle, as, and not dissimilar to today, some privileged individuals could obtain tax exemptions. These were for example clergymen, officials or members of the nobility, in so far as they resided in urban or rural communities. It is not possible to systematically measure the prevalence of these groups. In a pre-census age, one can only approximate roughly the groups' size. Between 0.5 and 1.5 percent of the population were part of the nobility or clergy in preindustrial Germany (Saalfeld 1980: 480). It is likely that most tax-exempt households belonged to these groups, but not all noblemen or clergymen were exempt. For instance, lower-level clerics were usually included in the tax registers.

These minor possible omissions could lead to an underestimation of wealth inequality. However, the potential omission of small groups should not conceal that practically the entirety of civilian households, and therefore the economically most interesting categories like craftsmen, peasants, and merchants, is covered by the wealth distributions. Note that German tax registers have been found to cover poor households very well (Alfani et al. 2022). Potentially missing parts of the distribution are likely to bias my estimates of the impact of the Reformation downwards. The Appendix shows that the *potential* exclusion of parts of the poor from the wealth distributions does not change the trend of poor peoples' wealth shares, which is reassuring given that my regressions include unit- and time-fixed effects. (Additionally, the Appendix reproduces examples of the archival sources, showing also how households with zero wealth were typically recorded.)

A potential concern could be the presence of "hidden wealth", perhaps differentially among confessional groups. Authorities implemented several measures to fight tax evasion, such as obliging individuals to swear an oath on the correctness of their tax payment, checks of tax estimates and payments, and severe penalties for evaders, such as confiscation of one's property and public announcement of evaders' names. These measures were likely imperfect, again very similar to today, but it is reasonable to assume that they increased an individual's cost of tax evasion substantially (Isenmann 2014: 539–541). To my best knowledge there is no study providing evidence of differentially hidden wealth among confessional groups, and a potential bias would most likely work against the result I find.¹⁷ One might also be concerned about Protestants applying different taxation principles, which could influence the recorded distribution of wealth. Again, to my best knowledge, no study demonstrating differential taxation record-keeping among confessional groups in preindustrial Germany exists. The reason is most likely that taxation in the localities of the Holy Roman Empire had a common legal basis, which contained the possibility of local variations

¹⁷If anything, one might conjecture that top wealth holders in Catholic communities were more inclined to hide wealth, because richness was morally sanctioned by the Old Church, while it was implicitly encouraged by Protestantism (Weber 1930). This could bias estimates of wealth concentration in Catholic places negatively. However, the results reported below provide no evidence that Catholics had significantly lower wealth concentration.

(in the Appendix I show archival tax registers from Catholic and Protestant places, which visually exhibit a high degree of homogeneity).¹⁸ Similarly, one might wonder whether local bureaucracies in Protestant places recorded people more rigorously in the tax registers. This would be particularly relevant if poor strata were recorded more thoroughly. If that were the case one would expect, for example, a higher number of taxpayers in Protestant compared to Catholic places. To address this concern I check whether there are any discontinuities in the number of taxpayers during the early phase of the Reformation, and I do not find any significant differences between the two confessions.

Information about whether the Reformation was introduced in a community has been obtained from standard secondary sources, mainly from the *Deutsches Stdtbuch* and the *Historisches Lexikon der deutschen Lnder* (see Cantoni 2012, 2015, Dittmar and Meisenzahl 2020 for a similar approach; see the Appendix for details about coding of variables). Few communities in the dataset reverted to Catholicism after introducing the Reformation. I have coded those few places depending on whether the Reformation was partially or completely reversed.¹⁹ I will show that the results do not depend on these switching cases.

Table 2: Community Characteristics before the Reformation

	(1)	(2)	(3)
	β Protestant	SE	Mean
Bottom 50% Wealth Share pre Reformation	-0.91	(1.84)	12.55
Bottom 20% Wealth Share pre Reformation	0.37	(0.48)	2.18
Middling 40% Wealth Share pre Reformation	-0.11	(2.33)	21.76
Top 10% Wealth Share pre Reformation	-3.24	(3.78)	45.92
Gini Coefficient pre Reformation	-0.01	(0.04)	0.59
Population Size (Log) pre Reformation	0.31	(0.36)	7.51
Occurrence of Epidemics pre Reformation	-0.07	(0.10)	0.38
Urban Community	-0.07	(0.11)	0.61
Soil Quality	0.00	(0.03)	0.72
Longitude	-0.38	(0.25)	10.04
Latitude	3.20***	(0.34)	49.50
Distance to Wittenberg	-185.74***	(18.86)	346.45

Notes: Column 1 shows the estimates on an indicator for a community being Protestant before the Reformation began in bivariate regressions. Columns 2 displays standard errors in parentheses. Columns 3 provides the mean of the dependent variable in the whole sample. The sample size for the the pre-Reformation period is 99 observations. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

What were the characteristics of those communities that eventually became Protestant? As shown in Table 2, there were some differences (estimated with bivariate regressions) between them and Catholic places in the pre-Reformation period. Column 1 shows by how much communities that became Protestant after 1517 differed from the mean values of all communities (Column 3). Column 2 shows the standard errors. All differences, in terms of wealth shares of different parts of

¹⁸The basic procedure for levying wealth taxes goes back to King Rudolf I, who decreed in 1287 that local taxes should be levied by applying a uniform tax rate to the entire property owned by every household (Isenmann 2014: 522).

¹⁹To exemplify my coding approach, consider the cases of Augsburg and Konstanz. In Augsburg Catholicism co-existed with Protestantism after the Schmalkaldic War (1546-48), but Reformation legislation was not genuinely taken back. Konstanz, instead, was entirely re-catholicised and Reformation legislation was reversed. Consequently, Augsburg was coded as Protestant and Konstanz as Catholic.

the population (Bottom 50 percent, Bottom 20 percent, Middling 40 percent, Top 10 percent), overall inequality (Gini), economic and demographic characteristics (population size, occurrence of epidemics) and time-invariant community characteristics (city-status, soil quality) were not statistically significant in my sample. The only significant differences are geographical. Eventually Protestant communities were more likely to be in the North, and closer to the Reformation’s starting point Wittenberg, a fact that will be exploited in the instrumental variable analysis (see Becker and Woessmann 2009, Cantoni 2012; see also Figure 4). However, in the robustness checks I will show that neither Northern communities nor the changing composition of the unbalanced panel drive the impact on inequality.

An important question is how much property the poor actually owned. Table 2 indicates that the bottom fifth of the population owned on average 2.18 percent of the total wealth in my sample before the Reformation. But how much is this really worth? We do not have information about poor peoples’ wealth at the macro-level at the time, but we can combine information about taxable wealth, the number of households, local currencies and real wages in an individual community to get a rough estimate. For instance, in Augsburg, just before the onset of the Reformation, 2.18 percent of total wealth would have implied that a household in the poorest fifth of the population owned property corresponding approximately to 17.14 daily wages of an unskilled worker (see the Appendix for details of the calculation).

One might also wonder what the composition of this property was, and who the poor actually were. These questions are even harder, if not impossible, to answer systematically, because tax registers do not usually record the different types of household property, but only indicate a total sum. Information about occupation and other individual characteristics, too, were rarely recorded systematically. The available evidence suggests, however, that poor people owned a variety of assets. For example, they typically possessed household goods, such as textiles and furniture, agricultural produce, such as foodstuffs, small sums of cash, and sometimes little amounts of savings and real estate. Their individual characteristics were diverse too. Among the poor we find small peasants, shepherds, craftsmen (for example, bakers, shoemakers, carpenters or weavers), small retailers, journeymen, day labourers, construction workers, apprentices, servants, but also the so-called “dishonest occupations”, such as hangmen. Many poor people were also just old, sick, or female, especially after their breadwinning husband had died (Jütte 1994: 71-72).

4 Empirical Analysis

4.1 Econometric Methodology

In order to identify the effect of the Protestant Reformation on economic inequality, I employ the difference-in-differences (DD) setup in Equation 1:

$$I_{i,t} = \alpha_i + \pi_t + \theta Prot_i \times PostReform_t + \gamma' \mathbf{X}_{i,t} + \epsilon_{i,t} \quad (1)$$

$I_{i,t}$ is some measure of wealth inequality in locality i in year t ($t = 1400, 1425, \dots$ until 1800).²⁰ I am mostly interested in wealth percentiles of poor strata, but also consider other inequality measures, such as the Gini coefficient. The difference-in-differences estimate is given by the interaction of post-treatment indicator ($PostReform_t$) and treatment status ($Prot_i$), a standard approach in the literature (see among others Cantoni et al. 2018, Becker and Pascali 2019).²¹ The post-indicator is coded as “one” after 1517,²² which follows the coding of Cantoni (2015: 562) and Dittmar and Meisenzahl (2020). The reason for this coding is that, although there was variation in the official legal adoption of the Reformation (see the Appendix for historical background information), many communities *de facto* implemented reformist ideas before the legal introduction. They did so, for example, through the substitution of the Catholic priest with a Protestant preacher. In other words, the legal introduction might not give an accurate picture of when a community started to live as Protestant. It therefore seems prudent to place the beginning of treatment when the Reformation movement began, and Lutheran ideas started to spread throughout Germany. In the Appendix I repeat the baseline specification using the official legal introduction date, which leads to almost identical results. I assigned communities treatment status based on whether they had become Protestant by 1600.²³ While the year 1600 is arguably arbitrary, it is historically a reasonable choice. The Peace of Augsburg in 1555 established for the first time the right of authorities to convert to Protestantism and impose the new religion on subjects. Several polities did officially introduce the Reformation in the following years, but it is sensible to assume that the conversion process had reached a steady state until about 1600. This coding, too, is a standard approach in the published literature (Rubin 2014, Cantoni 2015, Dittmar and Meisenzahl 2020; similarly Becker and Pascali 2019), and alternative coding of the treatment variable does not alter the results (see Appendix). Theta (θ) is the main coefficient of interest, providing an estimate of how the introduction of the Reformation affected economic inequality.

In order to address concerns about endogeneity, I account for potential omitted variable bias and reverse causality through several strategies. The modelling approach controls for unobserved factors that might confound the relationship between the dependent and the independent variable of interest. Locality fixed effects (α_i) account for characteristics that are time-invariant and locality-specific, such as local geographic or regional factors. They further account for certain constitutional characteristics, for instance, whether a locality was a city, which might have had higher inequality but might have also been more likely to adopt the Reformation, was member of the Hanse, or belonged to a bishopric. Time fixed effects (π_t) account for shocks that might have had an impact

²⁰Inequality measures have been clustered around their closest reference year.

²¹The treatment indicator does not differentiate between various forms of Protestantism. Also this is common practice in the published literature on the topic (see Cantoni 2015, Spenkuch 2017), for two reasons: first, differences existed between the branches of the reformed faith and within them. However, taken together the different Protestant confessions as inspired by Martin Luther, Huldreich Zwingli, Johannes Calvin and others were more similar to each other than to the theology and religious practices of the Old Church (see Blickle 2015: 71-74). Second, Lutheranism, by far the most dominant branch of Protestantism in Germany, is the only reformed faith in the dataset. It was probably less “Weberian” than Calvinism.

²²This coding implies that treatment is not staggered, so recent critiques of difference-in-differences research designs with staggered treatment are not a major issue (see for example Goodman-Bacon 2021).

²³Another historically sensible cut-off year would be 1624, the reference date established by the Peace of Westphalia. Only few territories changed their religious denomination between 1600 and 1624, and none of the places in the dataset did so.

on inequality in all localities, such as macroeconomic trends.

To further mitigate the possibility of the main variable of interest being correlated with the error term, I account for several locality-level *time-variant* observable demographic, economic and institutional characteristics included in vector $\mathbf{X}_{i,t}$. These controls are included in most but not all specifications, since they could be “bad controls”. I include the log population size of a locality because demographic expansion could have played a role in determining the preindustrial wealth distribution (Clark 2007, Milanovic 2016). As population size is a sensible marker of productivity growth, it is also a frequently used proxy for economic development, a potential driver of inequality (Kuznets 1955, van Zanden 1995: 649, 656-658, Deaton 2015: 1-5). But growth could have also driven the introduction of the Reformation: in more entrepreneurial environments, with vibrant markets, rich merchants and high overall prosperity, Protestantism, with its less hostile view on profit-seeking, was particularly appealing (Ekelund et al. 2002). I also include the local occurrence of epidemics, which could have impacted inequality but also the probability of the adoption of the Reformation (Alfani et al. 2022, Dittmar and Meisenzahl 2020). One might be concerned that wars and in particular the German Peasants’ War (1524-25) might have been an alternative treatment impacting Protestant and Catholic places differentially. This war just happened few years after the Reformation began and might have impoverished peasants in the affected areas, thus influencing inequality. To account for such a possibility, I control for the occurrence of battle action nearby. Ultimately a variable is included that indicates the log distance of a community to its nearest university. Universities might have had an impact on inequality (see van Zanden 1995: 658-661), but their role in the distribution of ideas and provision of trained theologians could have influenced the introduction of reformist ideas as well (Kim and Pfaff 2012). The Appendix provides further details about the coding of variables, and I test the robustness of my results to the inclusion of several time-invariant controls interacted with time-dummies. I consider variables that could have had an impact on inequality and Reformation adoption: agricultural potential, seaside location, whether a locality was an urban community, an Imperial city, a Hanse city or belonged to a bishopric.

Unobserved factors are captured with the random error term $\epsilon_{i,t}$. The standard errors are robust and clustered at the locality level to account for the possibility of serial correlation in the error term. The Appendix reports the baseline results with spatial autocorrelation-adjusted standard errors, and Morans’ I test for spatial dependence. They indicate that spatial dependence is a negligible issue for this study. I also report results with standard errors clustered at the territorial level, which does not substantially change the results either.

4.2 Difference-in-Differences Estimates

4.2.1 Motivation

How did the Reformation come about? Some historians argue that the Reformation was above all a response to widespread dissatisfaction with theological and church-matters at the time. People were frustrated about the growing contradictions between preached ideals and lived reality of clerics, corruption, and the sale of indulgences. This, it is claimed, fuelled anti-clericalism, and the Reformation picked up that sentiment (Blickle 2015: 29-40). But for the overwhelming majority of

people the decision to adopt the Reformation was actually taken by their territorial lords, namely the competence conferred on them in the Peace of Augsburg of 1555. The religious confession was practically *imposed* on the population living in the communities of the sixteenth century, not to speak of the generations born into a confessionally settled environment in later centuries. Only about 10 percent of the population in sixteenth-century Germany ever showed an active interest in the ideas of the Reformation, but up to about 80 percent of the population was Protestant during that time. The difference must have been due to the confessional choices of magistrates and princes (Scribner and Dixon 2003: 34).

There exists also a rich literature in the social sciences on the causes of the Reformation, especially why rulers decided to introduce it, and an obvious concern when studying its socio-economic effects is the potentially endogenous choice of religious confession by political elites: selecting into the Reformation could have been correlated with determinants of economic inequality, hereby inducing bias. For example, Ekelund et al. (2002) discuss the possibility that things that could drive inequality, such as trade and growth, were also related to the adoption of the Reformation, although they explicitly aim to explain the diffusion of the Reformation at the country- and not the community-level. Other potential causes were the diffusion of printing, ideological influence by spatial diffusion or political risk (Rubin 2014, Becker et al. 2020, Cantoni 2012; for a review see Becker et al. 2016).

It is not unthinkable that some of these factors also influenced local economic outcomes and inequality. The main motivation for the difference-in-differences analysis comes, therefore, from Figure 1, which plots the raw data. It shows a divergence in the wealth shares of poor strata in newly Protestant places, but similar trends compared to Catholic localities before the Reformation began (for a similar approach to identification, see for example Cantoni 2015, Cantoni et al. 2018, Becker and Pascali 2019, Dittmar and Meisenzahl 2020). Moreover, Table 2 shows that there were no statistically significant differences in my sample between eventually Protestant and Catholic communities, in terms of poor peoples' wealth shares, overall inequality and several other covariates. Flexible DD estimates will provide further evidence for the common-trends assumption before the advent of Protestantism, making it unlikely that treatment and outcomes were both driven by unobservables and suggesting that differences in outcomes are indeed attributable to the Reformation. I will also show that the main results are confirmed when exploiting exogenous variation in treatment allocation in an instrumental variable strategy, which accounts for the possibility that the Reformation was adopted, for example, because of local economic factors.

4.2.2 Baseline Results

Following the theoretical framework (Section 2) one would expect to observe a reduction in the wealth shares of poor strata, particularly among the poorest of the poor in Protestant places. Table 3 reports DD results for the effect of the Reformation on the distribution of wealth in communities. The coefficients represent an average post-treatment difference in wealth shares or inequality. In order to evaluate the hypothesis derived from the theoretical framework, I consider a wide definition of the poor, that is, the wealth shares of the bottom 50 percent (Piketty 2020), but

also narrower definitions, namely the wealth shares of the bottom 20 percent (Dollar and Kraay 2002) and of the bottom 10 percent. While the effect on the bottom 50 percent is insignificant (Column 1), the other two estimates are negative and highly significant (Columns 2 and 3). They are also economically sizeable if one considers that lower deciles of the population owned minuscule shares of overall wealth in communities: the bottom 20 percent in Protestant places lost about one wealth share percentage point, which corresponds to 39.4 percent of their pre-treatment wealth share of about 2.6 percentage points (compare regression coefficients with pre-treatment wealth shares, reported in the summary statistics in the Appendix). This suggests that the Reformation had a sizeable impact on inequality if measured as the wealth percentiles of lower classes. It did not have a significant impact on the bottom half but made specifically people in the poorest fifth of society comparatively poorer.

Table 3: Difference-in-Differences Estimates: Wealth Share Changes of Poor Strata

	(1) Bot. 50%	(2) Bot. 20%	(3) Bot. 10%	(4) Bot. 50%	(5) Bot. 20%	(6) Bot. 10%	(7) 2. Decile	(8) Gini	(9) Bot. 10%/Top 90%
Protestant \times Post-Reform.	-1.094 (1.369)	-1.004** (0.383)	-0.431*** (0.144)	-1.330 (1.292)	-1.023*** (0.360)	-0.428*** (0.137)	-0.595** (0.247)	0.031 (0.023)	-0.004*** (0.001)
Controls	NO	NO	NO	YES	YES	YES	YES	YES	YES
Locality FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	368	368	368	368	368	368	368	368	368
Communities	43	43	43	43	43	43	43	43	43
R^2	0.255	0.165	0.147	0.306	0.228	0.235	0.194	0.318	0.235

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Columns 4 to 9 introduce all demographic and economic, and institutional time-variant controls, to make the comparison between Protestant and Catholic communities as close to *ceteris paribus* as possible. The results do not change. Since the controls contain a proxy for economic development, a community's population size, the results indicate that the distributional, *relative* losses of the poor strata hold even if changes in *absolute* prosperity are accounted for. In the Appendix I test the robustness of my results to the inclusion of several time-invariant controls interacted with time-dummies, which does not change the main result.

The results also hold for the second decile of the distribution (Column 7). Comparing Columns 6 and 7, the coefficients indicate that the bottom 10 percent (first decile of the distribution), the poorest of the poor, lost the largest wealth share among the poor: 44.0 percent, compared to the 37.8 percent loss of the second decile (see summary statistics in the Appendix for pre-treatment wealth shares). This is indicative of Protestantism hurting economically most severely people at the margins of society, as the theoretical framework predicts. I also consider indicators of wealth inequality in the whole population. The Gini coefficient (Column 8) points towards inequality increase, but it is not significantly different from zero. This does not mean that there were no changes to overall inequality in Protestant places. It only means that the Gini coefficient does not capture the changes, possibly because the indicator is more sensible to changes in the middle of a distribution rather than its extremes. The share of the bottom 10 percent in relation to the rest of the population (Column 9), another indicator of wealth inequality, suggests a highly statistically

significant increase in overall inequality.

One might be wondering whether the wealth share reductions of poor strata are simply the mechanical consequence of a higher wealth concentration at the top or middle of the wealth distribution, which would be plausible if inequality were driven by things like economic growth, capital accumulation or higher upper-tail human capital (see Weber 1930, Becker and Woessmann 2009, Dittmar and Meisenzahl 2020). To address this concern I also estimate whether the Reformation had an effect on top and middling wealth shares, reported in Table 4. Columns 1 and 2 show the changes for top wealth shares. The coefficients are positive but small in magnitude. The coefficient in Column 1 corresponds to 6.8 percent of the top 10 percent wealth share and the coefficient in Column 2 to 1.1 percent of the top 1 percent wealth share in the pre-treatment period in eventually Protestant places (see summary statistics in the Appendix for pre-treatment wealth shares). Yet none of the estimates indicates wealth concentration at the top significantly different from zero. In other words, these estimates provide no evidence for significant wealth concentration in the upper strata of society in Protestant communities.²⁴ However, since the coefficients are not precisely zero, one cannot rule out with certainty that the Reformation led to limited redistribution to the rich.

Table 4: Difference-in-Differences Estimates: Wealth Share Changes of Top and Middling Strata

	(1) Top 10%	(2) Top 1%	(3) Mid. 40%	(4) Top 10%	(5) Top 1%	(6) Mid. 40%	(7) Bot. 20%	(8) Bot. 20%
Protestant \times Post-Reform.	3.072 (2.334)	0.123 (1.829)	-0.161 (1.608)	3.493 (2.504)	0.340 (1.849)	-0.514 (1.580)	-0.795** (0.321)	-0.987*** (0.354)
Top 10% wealth share	NO	NO	NO	NO	NO	NO	YES	NO
Middling 40% wealth share	NO	NO	NO	NO	NO	NO	NO	YES
Controls	NO	NO	NO	YES	YES	YES	YES	YES
Locality FE	YES	YES	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES	YES	YES
Observations	368	368	368	368	368	368	368	368
Communities	43	43	43	43	43	43	43	43
R^2	0.196	0.071	0.286	0.237	0.109	0.304	0.292	0.265

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

To check whether the Reformation had any impact on the wealth share of the “middle class”, Column 3 reports the effect of the Reformation on the arithmetic middle (deciles four to seven) of the population. This coefficient, too, is statistically not distinguishable from zero. Columns 4 to 6 indicate that all these results hold when controls are added to the specifications. Could it be that the insignificant results are only the result of a too low statistical power of a relatively small dataset? While this possibility cannot be entirely excluded, Table 3 suggests that the dataset is not too small to capture significant distributional changes caused by the Reformation on lower strata of the population.

One might also ask how it is arithmetically possible that the lower strata of the population lost wealth shares, while the middle and upper parts did not gain. Note that the difference in the wealth

²⁴This result is in line with the recent findings of Kersting et al. (2020), who do not find evidence for Protestant capital accumulation in modern Prussia.

share of middle and upper classes as indicated in Columns 1 to 6 in Table 4 is just not *statistically significant*, which does not rule out that there was any redistribution. Importantly, however, these possible limited gains in other parts of the wealth distribution do not mechanically drive the loss of the poor strata, as Columns 7 and 8 of Table 4 show. There I use the top 10 percent and middling 40 percent wealth shares as controls, and the bottom 20 percent share as outcome. If losses of the poor were just a mechanical result of other strata's gains, the coefficient of interest should be close to zero, but it is not. This suggests that Protestantism led to a highly significant and sizeable reduction of poor peoples' shares, mostly independent of the wealth shares of other parts of the population. Overall, I interpret these results as follows: the wealth share lost in a concentrated fashion at the bottom, was distributed in a dispersed way, not benefiting any other part of the population in particular.

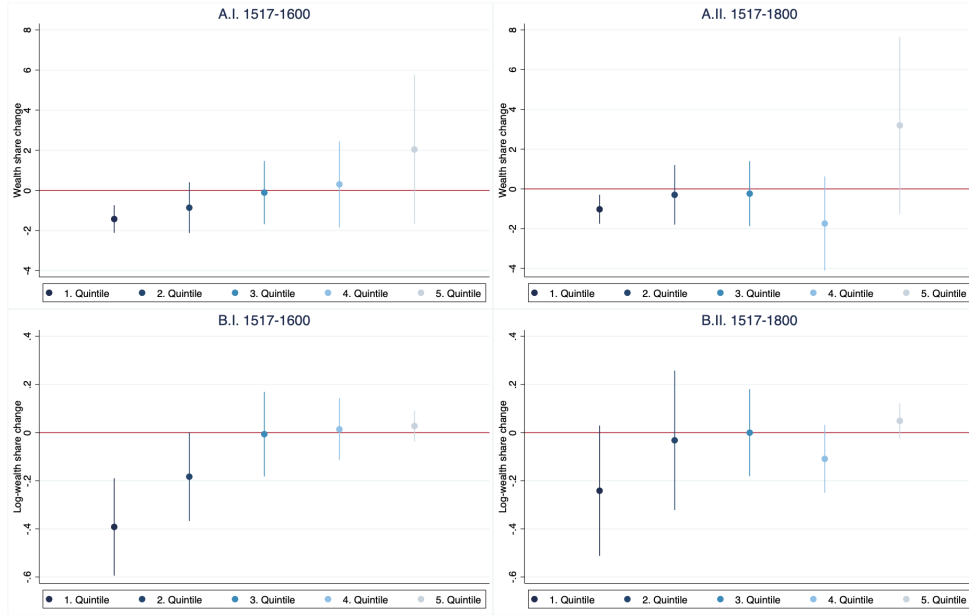
This interpretation of concentrated losses but dispersed gains is supported by Figure 6, where I estimate the saturated specification and plot the coefficients of the effect of Protestantism on the wealth shares of all quintiles of the wealth distribution (Panel A) and, to highlight the relative size of the change, on the log-wealth shares of all quintiles (Panel B).²⁵ The left frame covers the Reformation period until just before the Thirty Years' War, the right frame the whole period of analysis. Only the fifth quintile, that is the bottom 20 percent of the distribution, is negative and significant in both periods, and in absolute as well as in relative terms (p-value 0.79 in Panel B Frame II.). All the other parts of the distribution experienced mostly insignificant and much smaller relative changes.

Another concern might be whether the Reformation led to migration between Protestant and Catholic places, which might imply geographical spillover effects and could have an impact on the results. This type of externality would amount to a violation of the stable unit treatment value assumption (SUTVA). While theoretically possible,²⁶ it is historically unlikely to be a major issue. The Holy Roman Empire was the most fragmented political entity in early modern Europe consisting of more than 300 polities, and no general right to migrate between the territories existed. To the contrary, subjects had to ask permission from their authorities, had to be free from feudal bonds, often had to pay high fees for the right to leave, needed to obtain a passport for moving around, and host communities could simply refuse to take in immigrants they did not want, such as poor people. These economic and institutional barriers restricted mobility considerably given that migrants had to cross many different jurisdictions when covering even short distances. The barriers forced in particular the undesired poor to stay put, because police ordinances outright criminalised their migration, and they probably could not afford the high costs of leaving (Gerteis 1998, Blickle 2015: 189-190). Additionally, because the religious divide was such an incisive cultural dividing line at the time, individuals from the other confession were usually met with hostility and, for example, excluded from receiving welfare benefits such as poor relief from hosting communities (Battenberg 1991: 60). Ultimately, spillover effects would bias my estimates against finding a significant effect

²⁵In order to keep the zeroes, I have added the value one to all wealth quintiles before the log-transformation.

²⁶In theory, people could leave their territories for religious reasons after 1555. In practice, the legal and economic hurdles made emigration at the very least a ruinous endeavour, thwarting the economic motives for migration (Blickle 2015: 189-190).

Figure 6: Difference-in-Differences Estimates: Log-Wealth Share Change of Quintiles of the Population



Notes: Regression estimates of the wealth share change (Panel A) and log-wealth share change (Panel B) of all quintiles of the population in Protestant communities, with respect to Catholic communities (horizontal red line). The treatment is Protestant \times Post-Reformation. The pre-treatment period is 1400 to 1500. Estimation method is OLS. All regressions include a full set of locality and time fixed effects, and all controls. Standard errors clustered at locality level. Confidence intervals indicate significance at the 95-percent level.

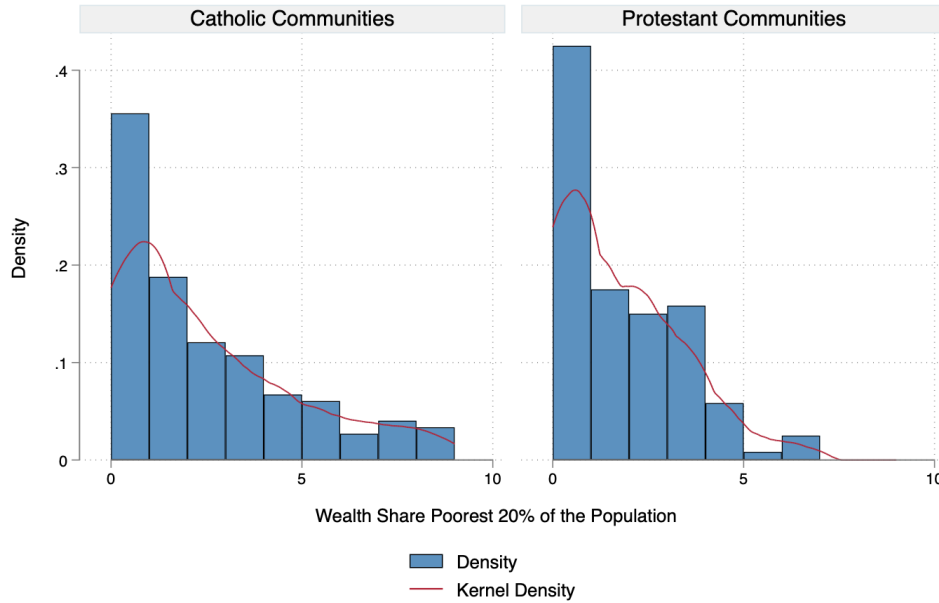
of Protestantism. If Protestantism made poor people worse off, one would expect that the poor tried to leave these places, thus reducing their number and measured inequality. Then the actual effect would be even larger than what my estimates suggest.

4.2.3 Robustness: Wealth Share Distribution of Poor Strata and Alternative Samples

Given the limited nature of the dataset, one might be wondering whether the results are driven by idiosyncratic characteristics of groups of communities in the dataset, that is, outliers. In Figure 7 I plot the raw distributions of wealth shares of the bottom 20 percent in Catholic and Protestant communities, after the beginning of the Reformation. The distributions do not show a concentration of unusual values, which would be indicative of outliers. But comparing the two distributions reveals another interesting fact: the right tale of the distribution, with wealth shares above five percent, is almost entirely missing in the sample of Protestant places. Instead, zero-wealth shares are more common with respect to Catholic places. In other words, the Protestant distribution of wealth shares is more skewed towards the bottom.

Figure 5 shows that the results hold when certain localities are dropped from the analysis. A first concern could be that Northern communities were peripheral and not well integrated in the governance structures of the Holy Roman Empire. They therefore differed from other places in the centre and south of Germany and it might be that the Reformation-variable just picks up this otherness. In Column 1 the four communities high in the north (see Figure 4) are dropped, which does not change the results. I also drop the five largest cities in the dataset in 1500 (Lübeck,

Figure 7: Distribution of Wealth Shares of the Bottom 20% in Protestant and Catholic Communities after 1517



Notes: Distribution of wealth shares of the Bottom 20% of the population in Protestant and Catholic communities after the beginning of the Reformation (1525-1800). The red line is the kernel density estimate. The sample size is 269 observations.

Augsburg, Erfurt, Munich and Frankfurt a.M.), to examine whether these exceptionally large and economically dynamic places drive the results. Column 2 indicates that this is not the case. Another concern could be that the geographical clustering of places, evident in Figure 4, around Wangen (10 communities) in the south and Lippe (8 communities) in the centre-west, could drive the results. Columns 3 and 4 show that although losing this relatively large number of places reduces the power of the dataset considerably, the main results qualitatively hold.

Table 5: Difference-in-Differences Estimates: Alternative Samples

	(1) Bot. 20%	(2) Bot. 20%	(3) Bot. 20%	(4) Bot. 20%	(5) Bot. 20%	(6) Bot. 20%	(7) Bot. 20%	(8) Bot. 20%
Protestant×Post-Reform.	-0.995** (0.392)	-1.137*** (0.410)	-0.604* (0.339)	-0.825** (0.380)	-1.130*** (0.398)	-0.979** (0.384)	-1.045*** (0.376)	-1.581*** (0.434)
Northern localities excl.	YES	NO	NO	NO	NO	NO	NO	NO
5 largest cities in 1500 excl.	NO	YES	NO	NO	NO	NO	NO	NO
Lippe localities excl.	NO	NO	YES	NO	NO	NO	NO	NO
Wangen localities excl.	NO	NO	NO	YES	NO	NO	NO	NO
Confessional reversal localities excl.	NO	NO	NO	NO	YES	NO	NO	NO
Bishopric localities excl.	NO	NO	NO	NO	NO	YES	NO	NO
Second serfdom localities excl.	NO	NO	NO	NO	NO	NO	YES	NO
Balanced Panel 1475-1600	NO	NO	NO	NO	NO	NO	NO	YES
Observations	340	317	320	285	338	339	356	120
Communities	39	38	35	33	40	39	41	20
R^2	0.253	0.292	0.182	0.239	0.250	0.235	0.229	0.491

Notes: Estimation method is OLS. All regressions include a full set of locality and time fixed effects, and all controls. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

A fifth concern could be that localities that switch their confession more than once could drive

the results. Column 5 shows that the results hold when those places are dropped that, after having introduced the Reformation, partially or completely reverse to Catholicism. Sixth, one might wonder about localities that belonged to a bishopric. Column 6 suggests that the results hold when dropping these places. Seventh, one might be wondering whether the results are driven by differences in the strength of serfdom, that is labour coercion in rural areas. The intensity of labour coercion is very difficult to measure precisely, but the areas east of the river Elbe are generally seen as the part of Germany with the strongest form of labour coercion, the so-called “Second Serfdom” (Bosshart and Dittmar 2021). But it does not seem to drive the results, as Column 7 suggests. Ultimately, one might be concerned about the unbalanced structure of the panel. This data structure is, at the current state of research, inevitable if one wants to analyse a phenomena for which data is so hard to find such as household-level wealth inequality since the late Middle Ages. The differences in the composition of the sample *might* influence the results. For that reason Column 8 reports results for a balanced sub-sample, from 1475 until 1600, just before the beginning of the Thirty Years’ War, which, as we will see in Figure 8, might have been an important shock. Unsurprisingly, the coefficient is even bigger than for the full sample, and highly statistically significant. This suggests that the changing composition of the sample does not drive the results. These checks reassure that the main results are robust to the composition of the sample and are not driven by outliers.

In the Appendix I report further heterogeneity tests. I look at urban versus rural, Northern versus Southern, and Western versus Eastern communities. None of these checks changes the treatment coefficient of interest substantially, but it seems that Protestant poor strata lost significantly less wealth shares in cities and Eastern communities. I also check for discontinuities in the number of taxpayers during the early phase of the Reformation, which would be indicative of systematic changes in Protestant tax systems, such as different reporting of poor people. But the results are practically zero and insignificant.

4.2.4 Flexible Difference-in-Differences Estimates

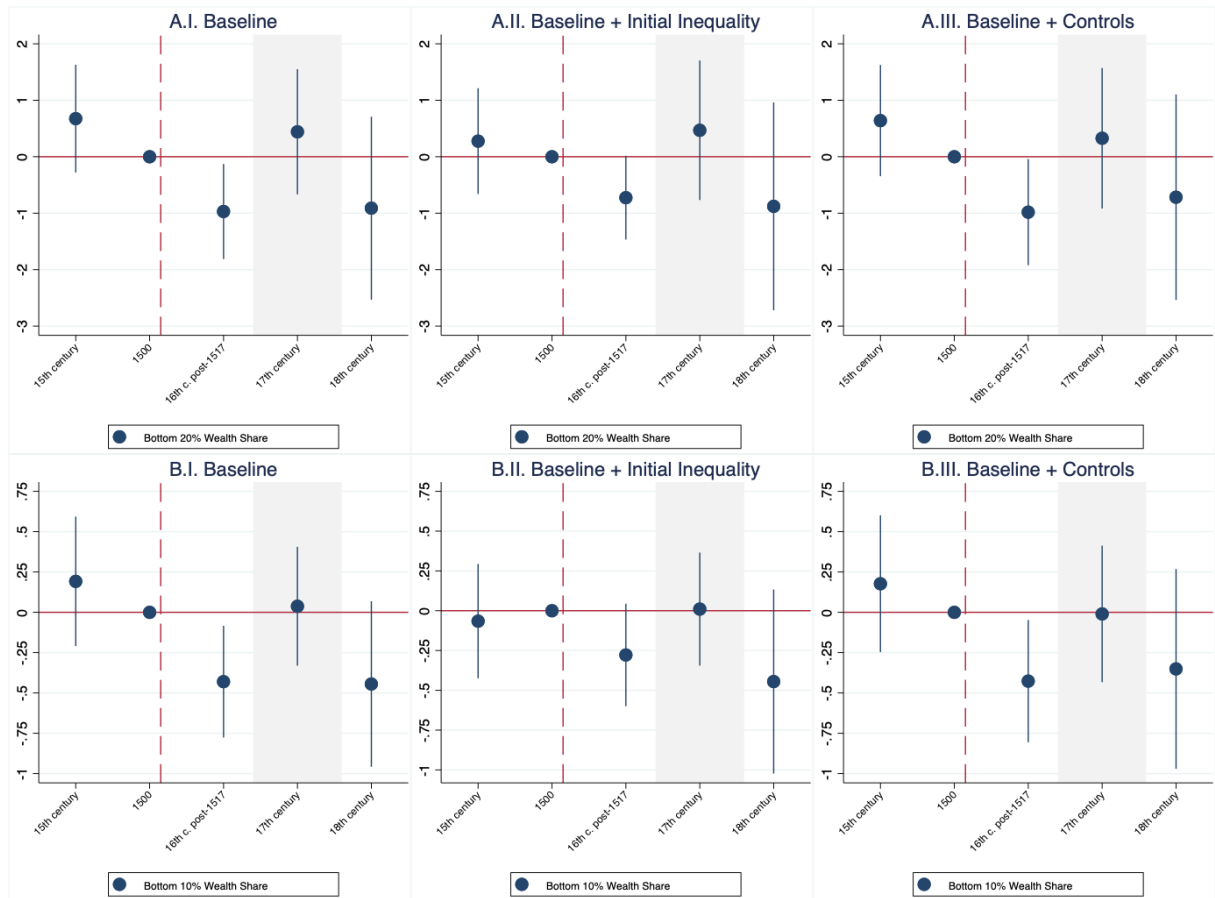
In Figure 8 I plot the estimation results of a flexible DD model, taking the wealth share of the bottom 20 and 10 percent as dependent variables. This setup allows me to evaluate, first, whether the relationship between Protestantism and the wealth share of the lower classes of society changed over time. Second, it allows for another test of the presence of differential pre-Reformation trends, and to control for such trends to some extent. The specification is almost identical to Equation 1:

$$I_{i,t} = \alpha_i + \pi_t + \sum_{t=1}^4 \beta_t (Prot_i \times Century_t) + \gamma' \mathbf{X}_{i,t} + \epsilon_{i,t} \quad (2)$$

The main difference is the inclusion of an interaction term between the treatment status indicator ($Prot_i$) and a set of four century dummies ($Century_t$) covering the pre- and post-treatment periods. The beta (β) coefficients are the main coefficients of interest. Unfortunately, the small dataset does not have enough power for estimating a fully flexible model in 25-year intervals. I therefore follow the example of Cantoni et al. (2018) and cluster observations into larger intervals, to increase the

number of observations for each interval. I take the year 1500 as reference category and divide the remaining period of observation into four intervals corresponding to centuries: from 1400 until about 1475; from 1525 just after the Reformation began until 1600, from 1625 until 1700; and from 1725 until the end of the period of observation in 1800. Note that this is still a demanding specification, considering the size of the dataset, especially towards the end of the period of study when the coverage of the dataset is less dense. The results are admittedly noisy, but still informative. Unfortunately, there is only one pre-Reformation period (15th century), because of a lack of sources for earlier periods. So the insignificance of the 15th-century dummy suggests the absence of a pre-trend, which will be further supported by the IV analysis.

Figure 8: Wealth Share of Bottom 20% and Bottom 10% in Protestant vs. Catholic Communities (Flexible Difference-in-Differences Estimates)



Notes: Regression estimates of the wealth share of the bottom 20 percent (Panel A) and the bottom 10 percent of the population (Panel B) in eventually Protestant communities before and after the Reformation (vertical red line), with respect to Catholic communities (horizontal red line). The omitted reference year is 1500. The estimation method is OLS. All specifications have a full set of community fixed effects and time fixed effects. Standard errors clustered at locality level. Confidence intervals indicate significance at the 95-percent level. The vertical red line represents the beginning of the Reformation period, the grey box the Thirty Years' War.

In Panel A I take the wealth share of the bottom 20 percent and in Panel B of the bottom 10 percent as dependent variable. Both patterns are almost identical. In Frame I the pre-Reformation period coefficients indicate that Protestant communities may not have been significantly different from Catholic ones. This supports the common-trend assumption. Yet the indicators of the first part of

the treatment period (1525-1600) point to a significant widening in outcomes between Protestant and Catholic communities once the Reformation began. Poor strata were relatively worse off in Protestant places, similar to the simple DD estimates in Table 3.

The coefficients for the period 1625 to 1700 indicate that the effect seems to disappear during the crisis-ridden seventeenth century, the century of the Thirty Years' War (1618-1648; approximately indicated by the grey box). Poor strata economically suffered the most during that war (Scheidel 2017: 339; see also Figure 1). This result might seem puzzling, but it is what one would expect if poor relief policies were one of the mechanisms causing lower wealth shares of Protestant poor strata, and considering the historical context. The war was exceptionally destructive, incomparably more than epidemics. About 40 percent of the German population died, and the prolonged macroeconomic recession that followed had long lasting negative budgetary effects, also for poor relief. This was particularly relevant in Catholic communities, as they had more generous poor relief provision (Alfani et al. 2022, Jütte 1994: 131). For example, in Catholic Trier the war led to a complete cessation of poor relief for decades. While hundreds of poor households had received support in the sixteenth and early seventeenth centuries, the number of recipients dropped to zero during the war (Ackels 1984). In other words, this extraordinary war eliminated the redistributive effect of universal and generous Catholic poor relief, which instead could have its effect during small crises.

If the war was an exceptional shock to poor relief, it is not surprising that the coefficients for the eighteenth century are clearly negative again, which suggests that Protestant poor-relief policy had a lasting effect. The new institutions remained substantially unchanged until at least the late eighteenth or early nineteenth centuries (Jütte 1994: 106-109). However, this result is not statistically significant. The large confidence intervals suggest that this is probably the result of the poorer coverage of the later early modern period in the dataset. From 1725 to 1800 the dataset only contains 50 observations.

In Frame II I interact every community's inequality level at the beginning of observation with a full set of year dummies. These interactions absorb a large amount of the pre-Reformation variation in the outcome, that is, the potential pre-trends that one might suspect looking at estimates in Frame I. The benefit of this demanding specification is to make communities even more comparable in the post-treatment period. Reassuringly, the patterns of redistribution observed in Frame I are qualitatively preserved in both panels and a point estimate for the pre-Reformation period around zero. In Frame III all controls are added to the baseline specification. In this demanding saturated specification, too, the picture of significantly lower wealth shares of Protestant poor strata is preserved in Panel A and B. That is an important result because one could argue that it was population decline leading to higher wages of lower strata, not the cessation of poor relief, that led to the vanishing difference between Catholic and Protestant communities during the seventeenth century. Since population change is controlled for in this specification, it cannot drive the results.

A general concern with the results presented in this section could be that Protestantism made poor people only relatively worse off, but richer in absolute terms due to positive effects on economic growth (see Becker and Woessmann 2009). This possibility cannot be entirely excluded, because

the only data systematically available refer to the distribution of wealth but, as mentioned, cannot be used to analyse absolute wealth levels. The main proxy used to control for economic growth — population size — is time-varying, but the saturated specifications only indicate that poor strata lost wealth shares, regardless of economic development. The regressions do not imply that growth was constant.

4.3 Endogenous Adoption of the Reformation: Instrumental Variable Analysis

Figure 1 and Table 2, and the flexible DD estimates did not provide evidence for significant pre-trends or pre-Reformation differences in inequality. But one might still be sceptical about whether this assumption really holds looking at some of the point estimates in Figure 8. More generally, one might be worried about selection bias, and ask whether the Reformation was adopted also because of unobserved economic factors, which could in turn be related to inequality after the beginning of the Reformation.²⁷ To rule out such potential remaining concerns I employ a second identification strategy. I follow an approach that has been employed in several previous studies (see Becker and Woessmann 2009, Cantoni 2012, 2015, Becker and Pascali 2019), and use the distance to the city of Wittenberg — the place where Luther started the Reformation — as an instrumental variable (IV). Since we cannot ultimately test the assumptions on which the method builds, the instrument *might* be imperfect (see Kersting et al. 2020). Hence, I view the results of the IV-analysis as a useful addition to the main DD results.

The IV isolates exogenous variation by exploiting that the Reformation was more likely to be adopted the closer a community was to the movements' starting point, possibly because of geopolitical considerations in the sixteenth century. There were numerous reasons why a prince or magistrate could have been in favour or against the Reformation (see Cantoni 2012). One of the reasons was that it was outright dangerous for princes and magistrates to officially introduce the new religious confession.²⁸ In that situation, having a powerful neighbouring polity made adoption less risky. Wittenberg was a centre of the first state within the Holy Roman Empire to adopt the Reformation, which was also a powerful polity: the Electorate of Saxony (Cantoni 2012). Additionally, being closer to Wittenberg and Saxony might have made it easier for imitators to observe how the Reformation was practically implemented (Becker and Woessmann 2009: 557-558).²⁹

In the Appendix I report first-stage regression results, which indicate that distance to Wittenberg is indeed a strong predictor of Reformation adoption, suggesting that the instrument-relevance condition is met. The exclusion restriction in this setting amounts to assuming that, conditional on locality and time fixed effects and several covariates, the distance to Wittenberg affected inequal-

²⁷For example, one could argue that more entrepreneurial places were more prone to the Protestant interpretation of Christianity but were also more likely to have higher inequality after 1517. Then the OLS results could overstate the actual effect. Moreover, one could be concerned about measurement error in the treatment variable.

²⁸Legally, Luther was banned, and his ideas and writings were prohibited in the Empire before the Peace of Augsburg. A military intervention by the Catholic Emperor and his allies or similar dire political consequences were an imminent threat. Therefore, allegiances among polities were important when introducing the Reformation (Scribner and Dixon 2003: 42-43).

²⁹Having contact with Luther himself was immensely important for the spreading of his new interpretation of Christianity. He had most frequent and intense contacts — either through writing or personal visits — with places close to Wittenberg (Becker et al. 2020: 868-869).

ity only through influencing the probability of a community of becoming Protestant. One possible threat could be that being closer to Wittenberg was related to better pre-existing economic development, and therefore to higher inequality. Then the instrument would not be independent. However, Wittenberg was not at all an economic hub. It was an economically backward and remote place, or to put it the words of Luther “on the edge of civilisation” (Becker et al. 2020: 861).

Figure 9: Instrument Exogeneity: Wealth Distribution in 1500



Figure 9 provides additional evidence supporting the assumption that the instrument is as good as randomly assigned. Wealth shares and inequality indicators in 1500 are plotted against the distance to Wittenberg. If the instrument really provides variation that is exogenous to pre-existing determinants of wealth distribution, then one would expect to find no significant correlation between the distance to Wittenberg and distributional outcomes in 1500. This is precisely what the results of this placebo test indicate. The correlations are practically zero, and not statistically significant (see additional results in the Appendix). Moreover, in most of the regressions presented below I control for several observable economic but also demographic and institutional time-variant characteristics of communities, in addition to time-invariant characteristics captured by the fixed effects. If these saturated specifications still produced significant results, one could be relatively confident that the exclusion restriction holds, and that the instrument is as good as randomly assigned, that is, independent of potential outcomes.

My instrumental variable setup is similar to Cantoni (2015). Compared to the previous DD setup, the interaction term of interest including the indicator of the adoption of the Reformation is now instrumented by an interaction term containing the IV instead. This is conceptually equivalent to a two-stage least squares setup. The first-stage is:

$$Prot_i \times P_t = \alpha_{1i} + \pi_{1t} + \phi DistanceWittenberg_i \times P_t + \epsilon_{1i,t} \quad (3)$$

Equation 4 represents the second-stage:

$$I_{i,t} = \alpha_{2i} + \pi_{2t} + \delta \widehat{Prot_i \times P_t} + \epsilon_{2i,t} \quad (4)$$

Panel A of Table 6 shows the results of the IV estimates (reduced-form estimates are reported in the Appendix, showing the same pattern as the second stage results). The estimates are local average treatment effects (LATE). They report the causal impact of the introduction of the Reformation on communities which for reasons of geographical positioning became Protestant.³⁰ Weakness of the instrument does not seem to be an issue: Kleibergen-Paap F-statistics are well above the conventional cutoff of 10. The picture of distributional differences revealed by the IV estimates is the same as in Tables 3 and 4. Columns 1 to 3 show a negative, substantial, and highly significant effect of the Reformation on the wealth share of lower classes of society. Importantly, this negative relationship also holds when controlling for several community-level covariates, including proxies for economic development. The highly significant estimates in the saturated specifications including covariates (Column 2 to 6) indicate that the estimates are robust to potential exclusion and independence restriction violations. Column 4 shows that the overall inequality changes in Protestant communities are now also captured by the Gini coefficient. Yet Columns 5 and 6 show that there is again no evidence of a significantly positive effect of the Reformation on top wealth shares, or on the wealth of the arithmetic middle of the population. Overall, the IV results confirm the main DD results. The IV-approach de facto randomises treatment, which suggests that the DD results are unlikely to be driven by selection bias.

4.4 Limitations of the Analysis

The above analysis is not without limitations, mostly related to the relatively small size of the dataset. First, the results show some sensitivity when dropping parts of the dataset in Table 5. Although the effect never statistically disappears, one might be wondering which characteristics of the dropped communities drive the sensitivity of the results. Unfortunately, beyond the additional results reported in the Appendix, the data limitations make it impossible to fully disentangle the variation of the effect. There was lots of variation in how Protestant poor relief systems were set up in cities and territories. For example, in some places begging was forbidden, in others only restricted or not regulated (see below). Unfortunately, at the current state of research, we do not have systematic information about all aspects of local poor relief systems to account for this variation.

More generally, one might, second, be concerned about the generalizability of the results. Without doubt, the 368 observations from 43 communities cover the German area considerably less systematically than other datasets of German town, such as Cantoni et al. (2018) or Dittmar and Meisenzahl (2020). However, these larger datasets do not contain any measure of inequality and

³⁰The coefficients are slightly larger than the OLS estimates (Panel B). The difference most likely is the result of proxy measurement error in the Reformation-indicator. The IV might recover a cleaner measure of the intensity of treatment, which is lost in the simple binary treatment variable employed in the OLS regressions. It might also be that the IV removes some of the endogeneity that plagues and thus attenuates the OLS estimates, for example due to unobserved community characteristics. The IV identification relies on compliers, and if these places are less likely to have characteristics that attenuate the effects of the Reformation, this would lead to larger IV estimates compared to OLS estimates.

Table 6: Instrumental Variable Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Bot. 20%	Bot. 20%	Bot. 10%	Gini	Top 10%	Mid. 40%
Panel A: IV-Estimates						
Protestant \times Post-Reform.	-1.995*** (0.483)	-1.896*** (0.502)	-0.744*** (0.211)	0.052* (0.028)	2.899 (3.097)	-0.391 (2.034)
Panel B: OLS-Estimates						
Protestant \times Post-Reform.	-1.004** (0.383)	-1.023*** (0.360)	-0.428*** (0.137)	0.031 (0.023)	3.493 (2.504)	-0.514 (1.580)
R^2	0.165	0.228	0.235	0.318	0.237	0.304
F-Stat. IV	75.68	65.35	65.35	65.35	65.35	65.35
Controls	NO	YES	YES	YES	YES	YES
Locality FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES
Observations	368	368	368	368	368	368
Communities	43	43	43	43	43	43

Notes: Estimation method is TSLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

can therefore not inform us about the household-level distribution of economic means in the distant past. The relatively small sample used here is due to the rarity of inequality data for the preindustrial period, when no censuses or similar large statistical sources exist. The rarity of the sources, and the high degree to which the available sources have been exhausted for constructing this dataset, make it unlikely that a substantially larger dataset of preindustrial household-level inequality in Germany will become available any time soon.

5 Mechanisms: Particularistic Protestant Poor Relief Policies

The analysis has shown so far that the Reformation made poor people relatively poorer in Protestant communities, thus increasing the gap between them and the rest of the population. What are the likely mechanisms behind this relationship? One of them being that inequality and relative poverty grew in Protestant places due to potential economic expansion, for example fuelled by upper-tail human capital growth. These could have been outcomes of the Reformation and could have led to social differentiation. But on closer examination this conclusion is unlikely to explain the results: first, because in most regressions economic development is controlled for; second, because growth and higher human capital would benefit the middling and upper classes of society disproportionately, leading to wealth concentration at the very top of the distribution (van Zanden 1995). But as we have seen, no other part of the population gained significant wealth shares, neither the middling parts nor the percentiles at the very top. Another explanation might be that it was not the Reformation, but rather the German Peasants' War (1524-25) that impoverished poor strata and happened around the time when the Reformation began. This is an unlikely explanation too, because I also control for exposure to warfare in most regressions. The theoretical framework in

Section 2 suggests a different hypothesis to explain the empirical patterns: new low-redistribution policies, such as exclusion of marginal poor people from poor relief, but also prohibitions of begging, the disincentivising of almsgiving and the stigmatisation of the poor in the labour market. Ideally, one would like to measure low-redistribution policies with data on effective redistribution to poor people. Unfortunately, such data are not available for the pre-statistical age, so I have to rely on indirect evidence. I study, first, the legal changes brought about by the Reformation in cities and territories, and, second, monastery closures during the Reformation.

5.1 New Poor Relief Institutions: Protestant Church Ordinances

Often, some years after the adoption of the Reformation, Protestant city magistrates and territorial rulers introduced laws in their communities, which regulated the areas of public life that used to be regulated by the Catholic Church, namely education, church governance and, importantly, the secular provision of poor relief (Dittmar and Meisenzahl 2020). These “church ordinances” (*Kirchenordnungen*) were the legal basis for such poor relief institutions as the common chest or the prohibition of begging. They often put the particularistic provision of poor relief into written law, by declaring non-working and non-resident poor as non-eligible for receiving support. Ordinances also conferred communities the task of managing poor relief locally. The texts of the laws together with their dates of introduction have been edited in a multi-volume series by Sehling (1902). I employ the introduction of a church ordinance in a community as a proxy variable for the establishment of the new systems of Protestant welfare and poor relief, similar to the study of Dittmar and Meisenzahl (2020).

In addition to the mere presence of a church ordinance, the editions of Sehling make it possible to code a variable indicating whether the content of the laws regulated begging in a community. The regulation of begging took different forms: begging could be limited to certain days of the week, the right to beg could be denied to specific groups (for example foreigners), begging could be restricted to certain areas of the locality or prohibited altogether. The important aspect that unites these forms of regulation is that they restricted the opportunities for begging, thus denying needy people an important source of income and limiting voluntary transfers from better-off to poor individuals.³¹ This reduced poor strata’s wealth share.

In order to test the presence of these different patterns more formally, I estimate again a DD setup that is similar to Equation 1. The main interaction term of interest includes now the ordinance indicator:³²

$$I_{i,t} = \alpha_i + \theta_1 \text{Ordinance}_i \times \text{PostIntro}_t + \gamma' \mathbf{X}_{i,t} + \epsilon_{i,t} \quad (5)$$

Columns 1 and 2 of Table 7 document that the introduction of an ordinance led to a sizeable

³¹It has to be kept in mind that the text of the lengthy ordinances has not always been published entirely in the Sehling-volumes (see Dittmar and Meisenzahl 2020: Appendix). Hence the variable has potential measurement error because it is possible that some communities that were coded as not regulating begging actually did so. The estimates therefore give only a lower bound estimate of the actual effect of begging regulation.

³²Since the first ordinance dates from 1522, the treatment variable is interacted with a post-treatment indicator that switches on in 1522.

and significant reduction of the wealth share of the bottom 20 percent of the population. The effect of Protestantism loses its statistical and economic significance. This is indicative of church ordinances and the poor relief institutions they introduced having large explanatory power and driving a major part of the observed differences in the wealth shares of poor people in Protestant and Catholic communities.

Table 7: Evidence on Laws Regulating Poor Relief (Difference-in-Differences Estimates)

	(1) Bot. 20%	(2) Bot. 20%	(3) Bot. 20%	(4) Bot. 20%	(5) Bot. 10%
Ordinance×Post-Intro.	-1.476*** (0.383)	-1.639*** (0.286)	-0.532 (0.441)	-0.843** (0.412)	-0.478** (0.212)
Ordinance×Post-Intro.×Begging			-1.443*** (0.430)	-1.181*** (0.438)	-0.594*** (0.185)
Protestant×Post-Reform.	0.375 (0.428)	0.518 (0.339)	0.212 (0.460)	0.349 (0.361)	0.313 (0.189)
Controls	NO	YES	NO	YES	YES
Locality FE	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
Observations	368	368	368	368	368
Communities	43	43	43	43	43
R^2	0.174	0.239	0.200	0.256	0.283

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In Columns 3 to 5 I augment the interaction term of interest with the begging restriction indicator.³³ The results show that, as expected, regulating begging had a sizeable and negative association with the wealth share of poor strata, resulting from the closure of an important income channel for the poor. As one would expect, the presence of a church ordinance still shows a negative effect, most likely because ordinances restricted poor relief in unmeasured ways other than limiting begging, for instance by excluding able-bodied but non-working people from welfare or by disincentivising almsgiving. Note that the effect on the bottom 10 percent (Column 5) being more than half as large as the effect on the bottom 20 percent (Column 4) suggests that poorer households were hit harder by a begging regulation. Intuitively, this is exactly what one would expect because households were more likely to receive money from begging the poorer they were. Other economically relevant effects of begging regulations among the poor were a higher propensity for out-migration (although this was in practice difficult for the poor), higher mortality and increased incentives to work. However, all these factors would have reduced the number of poor people in a place, reducing inequality, implying that the reported coefficients are lower-bound estimates. In general one has to interpret these results keeping in mind that regulating begging was likely correlated with other unmeasured characteristics of the poor relief system that limited redistribution to poor people in Protestant

³³Note that the lower-order interaction Ordinance×Begging is absorbed by the locality fixed effects, and Post-Intro.×Begging would be identical to the triple interaction term since there are no places that had a begging regulation but no ordinance.

communities.

Overall, the results are consistent with the hypothesis that particularistic Protestant poor relief policies were an important mechanism for why poor strata were worse off in Protestant places relative to the rest of the population. There is evidence for at least two particularistic, low-redistribution policies having an impact, the new exclusionary Protestant poor relief system, and the prohibition of begging. The insignificant coefficients on the Protestantism indicator across all specifications might be interpreted as evidence that institutions mattered more than ideology or culture, but it could also be that strong ideology is correlated with begging restrictions. These things are impossible to disentangle with the available data. In the Appendix I report again the distribution of poor strata's wealth shares, which suggest that outliers are not an issue. I also report flexible difference-in-differences estimates, which show an analogous pattern to Figure 8 and support the conclusions of this section.

5.2 The Disappearance of Catholic Social Welfare: Monastery Closures

I also proxy for one facet of the disappearance of the universal social welfare system of the Catholic Church during the Reformation: the expropriation and closure of monasteries by local rulers, thus confiscating valuable assets of the Old Church. Monasteries used to redistribute a substantial part of the Church's income — probably between one third and one fourth — to the poor through almsgiving and were an important element of the Catholic poor relief system. Often monasteries also functioned as hospitals, providing rudimentary healthcare to poor people. Reformers had envisaged that monasteries' expropriated assets would be reemployed in Protestant communities to provide relief for needy people from the poor chest (Kahl 2009: 270). In reality, these assets often went into the coffers of local rulers and were not employed for welfare but for waging wars, building roads, hiring bureaucrats, or building palaces. In Catholic territories these assets continued to be employed to benefit the poor (Cantoni et al. 2018). One would expect that this disappearance of Catholic social welfare *ceteris paribus* led to lower strata of society losing wealth shares. I interpret monastery closures as a mechanisms, but one could of course also interpret them as a proxy for the diffusion of the Reformation itself.

I gather information on the location and, if applicable, the closure date, of 3,094 monasteries from Cantoni et al. (2018) and Jürgensmeier and Schwerdtfeger (2005). For each locality in my dataset, I calculate the number of monasteries nearby that were closed during the Reformation period until 1600. Since the closure was a one-off reallocation of resources, usually happening when a ruler converted to Protestantism, it is likely that an effect could be observed most clearly in the early phase of the Reformation period. For that reason, I provide estimates covering a shorter period until about 1600 and estimates for the complete period until 1800.

I estimate regressions analogous to the DD setup in Equation 1, taking monastery closures as treatment variable.³⁴ Column 1 indicates that the closures of monasteries significantly reduced

³⁴Note that Pfaff and Corcoran (2012) find that a higher monastery density decreased the odds of a community adopting the Reformation. This implies that poor people might have had lower income and wealth shares in eventually Protestant places before 1517. However, such potential differences in *levels* are not a major issue since the estimator used here calculates the change in wealth share *trends*.

Table 8: Evidence on Monastery Closures (Difference-in-Differences Estimates)

	1400-1600			1400-1800		
	(1)	(2)	(3)	(4)	(5)	(6)
	Bot. 20%	Bot. 20%	Bot. 10%	Bot. 20%	Bot. 20%	Bot. 10%
Monasteries closed×Post-Reform.	-0.137** (0.062)	-0.167** (0.062)	-0.056** (0.027)	-0.076 (0.054)	-0.127** (0.056)	-0.050* (0.026)
Controls	NO	YES	YES	NO	YES	YES
Locality FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES
Observations	252	252	252	368	368	368
Communities	43	43	43	43	43	43
R^2	0.113	0.211	0.227	0.148	0.218	0.222

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

the wealth share of the bottom fifth of the population. For every monastery closed, people at the bottom of the population lost 0.137 percentage points of their wealth share until 1600. Columns 2 and 3 show that this result holds when controls are introduced. Columns 4 to 6 show that, as expected, the effect is somewhat smaller and less precisely estimated if we consider the entire period until 1800 but is qualitatively preserved. One might wonder whether rulers' propensity to close and expropriate monasteries was driven by an increase in the macro-level frequency of warfare in the Reformation period. However, such a macro effect should be captured fairly well by time fixed effects.

The results suggest that the disappearance of the universal Catholic social welfare system, triggered by the Reformation, and the subsequent closure of monasteries did not just redistribute economic resources between rulers and the Church (Cantoni et al. 2018). It also affected the distribution among individuals, by making poor strata relatively poorer and increasing inequality. This evidence is consistent with the universalism-particularism mechanism suggested by the theoretical framework.

6 Conclusion

This paper has investigated whether religious confession has an impact on the distribution of wealth and inequality, studying the emblematic case of the Protestant Reformation. I have presented a theoretical framework of the Reformation's inherent trade-off between, first, the expansion of public goods provision, especially in terms of social welfare, and, second, a more particularistic provision of poor relief to in-group members only. The framework suggests that the Reformation was quite ambiguous in its redistributive implications and for inequality. I have then argued that the Reformation overall reshuffled the lower part of the wealth distribution to increase the gap between poor strata and all others.

This hypothesis has then been tested empirically. Employing a DD and IV strategy, the analysis

finds strong evidence of a negative causal effect of the Reformation on the wealth share of lower classes of the population in Protestant communities between 1400 and 1800. It seems that the Reformation increased inequality, by making the poor relatively poorer compared to the rest of the population. This effect can be traced back through the early modern period, which suggests that the Reformation had a lasting impact that only began in the sixteenth century. Yet I do not find evidence for significantly higher wealth shares of the rich or middling parts of the population. This result is confirmed by a variety of specifications.

The empirical picture is consistent with a historical characterisation of the Reformation as the trigger of new low-redistribution policies in Protestant places, which embodied the particularistic character of the new religious confession towards parts of the poor. Evidence for the plausibility of this hypothesised mechanism comes from information about the introduction of church ordinances, begging prohibitions, and the closure of monasteries. The inequality-promoting character of Protestantism, typically observed in the modern day, seems to have deep historical roots. Protestantism might therefore be an important, hitherto underappreciated driver of rising preindustrial inequality, long before the onset of industrialisation and modern economic growth. Moreover, the case of the Reformation exemplifies that a key dimension of redistributive policies is how universal or particularistic societies provide social welfare (Enke et al. 2022), today as in the past.

References

- Abramitzky, R. (2008). The Limits of Equality: Insights from the Israeli Kibbutz. *Quarterly Journal of Economics* 123(3), 1111 – 1159.
- Acemoglu, D., S. Johnson, and J. Robinson (2005). Institutions as the fundamental cause of long-run growth. In P. Aghion and S. Durlauf (Eds.), *Handbook of Economic Growth, Volume 1A*, pp. 385–472. Amsterdam: Elsevier.
- Ackels, M. (1984). Das Trierer städtische Almosenamt im 16. und 17. Jahrhundert: Ein Beitrag zur Analyse sozialer Unterschichten. *Kurtrierisches Jahrbuch* 4, 75–103.
- Ager, P., L. Boustan, and K. Eriksson (2021). The Intergenerational Effects of a Large Wealth Shock: White Southerners After the Civil War. *American Economic Review* 11(11), 3767–3794.
- Alesina, A. and P. Giuliano (2010). Preferences for Redistribution. In J. Benhabib, A. Bisin, and M. Jackson (Eds.), *Handbook of Social Economics, Vol. 1A*, pp. 93–132. Amsterdam: North Holland.
- Alfani, G. (2021). Economic inequality in preindustrial times: Europe and beyond. *Journal of Economic Literature* 59(1), 3–44.
- Alfani, G., V. Gierok, and F. Schaff (2022). Economic Inequality in Preindustrial Germany, ca. 1300 – 1850. *Journal of Economic History* 82(1), 87–125.
- Bairoch, P., J. Batou, and P. Chèvre (1988). *The Population of European Cities from 1300 to 1850*. Geneva: Librairie Droz.
- Basten, C. and F. Betz (2013). Beyond Work Ethic: Religion, Individual, and Political Preferences. *American Economic Journal: Economic Policy* 5(3), 67–91.
- Battenberg, J. (1991). Obrigkeitsliche Sozialpolitik und Gesetzgebung: Einige Gedanken zu mittelrheinischen Bettel- und Almosenordnungen des 16. Jahrhunderts. *Zeitschrift für Historische Forschung* 18(1), 33–70.
- Becker, S., Y. Hsiao, S. Pfaff, and J. Rubin (2020). Multiplex Network Ties and the Spatial Diffusion of Radical Innovations: Martin Luther’s Leadership in the Early Reformation. *American Sociological Review* 85(5), 857–894.
- Becker, S. and L. Pascali (2019). Religion, Division of Labor and Conflict: Anti-Semitism in German Regions over 600 Years. *American Economic Review* 109(5), 1764– 1804.
- Becker, S., S. Pfaff, and J. Rubin (2016). Causes and consequences of the Protestant Reformation. *Explorations in Economic History* 62, 1–25.
- Becker, S. and L. Woessmann (2009). Was Weber Wrong? A Human Capital Theory of Protestant Economic Theory. *Quarterly Journal of Economics* 124(2), 531–596.
- Blickle, P. (2015). *Die Reformation im Reich*. Stuttgart: Kohlhammer.

- Bosshart, L. and J. Dittmar (2021). Pandemic Shock and Economic Divergence: Political Economy Before and After the Black Death. *Unpublished manuscript London School of Economics*.
- Cantoni, D. (2012). Adopting a New Religion: The Case of Protestantism in 16th Century Germany. *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.
- Clark, G. (2007). *A Farewell to Alms: A Brief Economic History of the World*. Princeton: Princeton University Press.
- Clasen, K. (1984). Armenfürsorge in Augsburg vor dem Dreißigjährigen Krieg. *Zeitschrift des Historischen Vereins für Schwaben* 78, 65–115.
- Deaton, A. (2015). *The Great Escape: Health, Wealth, and the Origins of Inequality*. Princeton: Princeton University Press.
- Dewan, T. and S. Wolton (2019). A Political Economy of Social Discrimination. *Unpublished manuscript*.
- Dittmar, J. and R. Meisenzahl (2020). Public Goods Institutions, Human Capital and Growth: Evidence from German History. *Review of Economic Studies* 87(2), 959–996.
- Dollar, D. and A. Kraay (2002). Growth Is Good for the Poor. *Journal of Economic Growth* 7(3), 195–225.
- Ekelund, R., R. Hébert, and R. Tollison (2002). An Economic Analysis of the Protestant Reformation. *Journal of Political Economy* 110(3), 646–671.
- Enke, B., R. Rodríguez-Padilla, and F. Zimmermann (2022). Moral Universalism and the Structure of Ideology. *Review of Economic Studies* forthcoming.
- Fischer, W. (1982). *Armut in der Geschichte. Erscheinungsformen und Lösungsversuche der Sozialen Frage in Europa seit dem Mittelalter*. Göttingen: Vandenhoeck & Ruprecht.
- Galor, O. and O. Moav (2004). From Physical to Human Capital Accumulation: Inequality and the Process of Development. *Review of Economic Studies* 71(4), 1001–1026.
- Galor, O., O. Moav, and D. Vollrath (2009). Inequality in Landownership, the Emergence of Human-Capital Promoting Institutions, and the Great Divergence. *Review of Economic Studies* 76(1), 143–179.
- Gerteis, K. (1998). Auswanderungsfreiheit und Freizügigkeit in ihrem Verhältnis zur Agrarverfassung: Deutschland, England, Frankreich im Vergleich. In O. Volckart (Ed.), *Frühneuzeitliche*

- Obrigkeiten im Wettbewerb: Institutioneller und Wirtschaftlicher Wandel zwischen dem 16. und 18. Jahrhundert*, pp. 82–98. Baden-Baden: Nomos.
- Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics* 225(2), 254–277.
- Guiso, L., P. Sapienza, and L. Zingales (2003). People’s opium? Religion and economic attitudes. *Journal of Monetary Economics* 50, 225–282.
- Hartung, W. (1989). Armut und Fürsorge: eine Herausforderung der Stadtgesellschaft im Übergang vom Spätmittelalter zur Frühen Neuzeit. In J. Jahn, W. Hartung, and I. Eberl (Eds.), *Oberdeutsche Städte im Vergleich. Mittelalter und Frühe Neuzeit*, pp. 158–181. Sigmaringen: Regio Verlag Glock und Lutz.
- Hudson, K. and A. Coukos (2005). The Dark Side of the Protestant Ethic: A Comparative Analysis of Welfare Reform. *Sociological Theory* 23(1), 1–24.
- Isenmann, E. (2014). *Die Deutsche Stadt im Mittelalter 1150-1550*. Köln: Böhlau.
- Jürgensmeier, F. and R. Schwerdtfeger (Eds.) (2005). *Orden und Klöster im Zeitalter von Reformation und katholischer Reform, 1500–1700 (3. Vol.)*. Münster: Aschendorff.
- Jütte, R. (1994). *Poverty and Deviance in Early Modern Europe*. Cambridge: Cambridge University Press.
- Jütte, R. (1996). Poverty and Poor Relief. In S. Ogilvie (Ed.), *Germany: A New Social and Economic History Volume II 1630-1800*, pp. 377–404. New York: Arnold.
- Kahl, S. (2009). Religious Doctrines and Poor Relief: A Different Causal Pathway. In K. van Kersbergen and P. Manow (Eds.), *Religion, Class Coalitions, and Welfare States*, pp. 267–296. Cambridge: Cambridge University Press.
- Kersting, F., I. Wohnsiedler, and N. Wolf (2020). Weber Revisited: The Protestant Ethic and the Spirit of Nationalism. *Journal of Economic History* 80(3), 710–745.
- Kim, H. and S. Pfaff (2012). Structure and Dynamics of Religious Insurgency: Students and the Spread of the Reformation. *American Sociological Review* 77(2), 188–215.
- Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review* 45(1), 1–28.
- Lindert, P. (2004). *Growing Public: Social Spending and Economic Growth Since the Eighteenth Century*. Cambridge: Cambridge University Press.
- Milanovic, B. (2016). *Global Inequality: A New Approach for the Age of Globalization*. Cambridge MA: Harvard University Press.
- Milanovic, B. (2018). Towards an explanation of inequality in premodern societies: the role of colonies, urbanization and high population density. *Economic History Review* 71(4), 1029–1047.

- Ogilvie, S. (1997). *State corporatism and proto-industry: The Württemberg Black Forest, 1580-1797*. Cambridge: Cambridge University Press.
- Persson, T. and G. Tabellini (1994). Is inequality Harmful for Growth? *American Economic Review* 84(3), 600–621.
- Pfaff, S. and K. Corcoran (2012). Piety, Power and the Purse: Religious Economies Theory and Urban Reform in the Holy Roman Empire. *Journal for the Scientific Study of Religion* 51(4), 757–776.
- Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge MA: Belknap Press.
- Piketty, T. (2020). *Capital and Ideology*. Cambridge MA: Belknap Press.
- Puga, D. and D. Treffer (2014). International Trade and Institutional Change: Medieval Venice’s Response to Globalization. *Quarterly Journal of Economics* 129(2), 753–821.
- Röck, B. (1989). *Eine Stadt in Krieg und Frieden – Studien zur Geschichte der Reichsstadt Augsburg zwischen Kalenderstreit und Parität*. Göttingen: Vandenhoeck & Ruprecht.
- 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.
- Saalfeld, D. (1980). Die ständische Gliederung der Gesellschaft Deutschlands im Zeitalter des Absolutismus. Ein Quantifizierungsversuch. *Vierteljahrschrift für Sozial- und Wirtschaftsgeschichte* 67(4), 457–483.
- Schaff, F. (2023). Warfare and Economic Inequality: Evidence from Preindustrial Germany (c. 1400-1800). *Explorations in Economic History* 89(2), 101495.
- Scheidel, W. (2017). *The Great Leveler*. Princeton: Princeton University Press.
- Schilling, H. (1994). *Aufbruch und Krise: Deutschland 1517-1648*. Berlin: Siedler.
- Scribner, B. and C. Dixon (2003). *The German Reformation, 2nd Edition*. Houndmills: Palgrave-Macmillan.
- Sehling, E. (1902). *Die Ordnungen Luthers, die Ernestinischen und Albertinischen Gebiete*. Leipzig: Mohr Siebeck.
- Spenkuch, J. (2017). Religion and Work: Micro Evidence from Contemporary Germany. *Journal of Economic Behavior and Organization* 135, 193–214.
- Tawney, R. (1926). *Religion and the Rise of Capitalism*. London: J. Murray.
- van Zanden, J. (1995). Tracing the beginning of the Kuznets Curve: Western Europe during the Early Modern Period. *Economic History Review* 48(4), 643–664.
- Weber, M. (1930). *The Protestant Ethic and the Spirit of Capitalism*. London: Routledge.

Appendices

A Historical Background: the Introduction of the Reformation in Early Modern Germany

The introduction of the Reformation took place over several decades. This long process can be divided into three phases: from 1517 until 1525, from 1525 until 1555, and from 1555 onward. An overview of the religious-political developments during this period is helpful for appreciating the identification strategy of the analysis in the main text.³⁵

The first phase begins with the formulation and circulation of Martin Luther's famous 95 theses in the Saxon town Wittenberg in 1517. They were a sharp critique of the Catholic Church and the Pope, especially of the practice of selling indulgences. Luther wanted to call attention to theological misconceptions and provoke a debate, but not split the Church. He was soon accused of heresy. Frederick the Wise, the Elector of Saxony, protected Luther, who was a theology professor at Frederick's newly established university in Wittenberg. The spread of Luther's critique in the Holy Roman Empire and beyond was rapid. It was made possible by the diffusion of Gutenberg's printing technology a few decades earlier, and by the fact that authorities in the Empire were distracted by political struggles and warfare. In these early years the Reformation was characterised by the absence of a legal provision in the Empire about whether or how to introduce the new faith. Several polities introduced it nevertheless, such as Electoral Saxony or Hesse. Some localities started to follow the reformatory vision of Christianity before their hesitant rulers officially introduced it, for example by substituting their Catholic priest with a Protestant preacher. However, this happened in a legal vacuum, and the Peasants' War in 1525 ended this early phase of the Reformation. A coalition of princes, Imperial cities and the Emperor crushed the revolt.

The war began the second phase of the Reformation, from 1525 until 1555, during which its introduction fell in the hands of political authorities. This was, at least implicitly, confirmed by the Imperial Diet in Speyer in 1526: as no Empire-wide agreement could be found yet, the decision to introduce the new faith was left to individual territorial rulers and magistrates. Still, this happened under great legal and political uncertainty.

The Peace of Augsburg was the beginning of the third phase of the Reformation, from 1555 onward. Only at this Imperial Diet a treaty between Lutherans, by far the largest branch of Protestantism in Germany, and Catholics was brokered. It officially confirmed the existence of a second protected religious confession in Germany. A principle later termed *cuius regio, eius religio* ("Whose rule, his religion") certified worldly rulers' monopolistic right to determine the confession in the communities of their territories and of their subjects, a right which they had practically exercised since 1526.³⁶ The treaty made the Reformation a project of the slowly emerging "states" within the Holy Roman

³⁵The following historical account is based on Schilling (1994: 85-116, 193-254, 445-464), Reinhard (2004) and Blickle (2015: 165-191).

³⁶In the case of the many ecclesiastical polities in the Empire, not even the rulers could choose the religious confession (*geistlicher Vorbehalt*). They were legally obliged to stick to the old confession or otherwise lose their dominion and possessions (Blickle 2015: 189-190).

Empire. Following the Peace of Augsburg in 1555, more polities became formally Protestant.

Even the Thirty Years' War (1618-1648) did not change the fact that two confessions existed in Germany. The Peace of Westphalia further consolidated the religious divide: it fixed polity's confessional denomination *ex post* to its status in 1624 and established that rulers could change their personal confession, but not force their territories' subjects to convert anymore.³⁷ As a consequence, for the overwhelming majority of subjects the religious denomination of the polities they were living in did not change until at least the nineteenth century.

B Additional Information on Inequality Data and Sources

B.1 Dataset

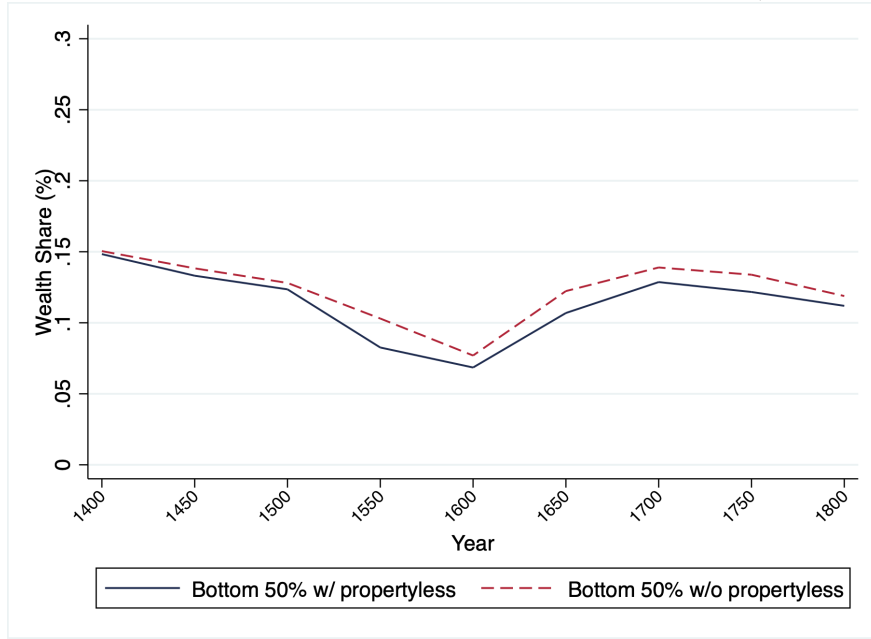
The inequality data and the sources used to construct the panel dataset have been extensively discussed by Alfani et al. (2022) and Schaff (2022). This Appendix provides some essential background information. The community-level inequality panel employed here differs from Alfani et al. (2022) in two substantial ways. First, it employs more household-level wealth data to obtain 25-year intervals between observations instead of 50-year intervals and is therefore more fine-grained. Second, it provides all local inequality estimates including the propertyless households — the only sub-category among the poor that was sometimes missing from the tax registers — where these were recorded in the tax registers. Of course, what constituted propertylessness was dependent on local understanding of the concept, and even if people without property were reported, we cannot exclude with absolute certainty that some poor people were not captured by the tax registers. Since Alfani et al. (2022) have been primarily interested in providing macro-level estimates of inequality for Germany comparable to estimates from other European areas, they have dropped the propertyless from their community-level distributions. However, for the purposes of this paper it seemed adequate to base the analysis on the complete local wealth distributions, including also the covered propertyless households.

The propertyless were a small part of the total population in preindustrial Germany. Their share ranged between 0.81 percent in 1400 and a high of 8.4 percent in 1550. Reassuringly, Figure 10 shows that the hypothetical exclusion of all the propertyless from the wealth distributions reduces the wealth share of the poor, as measured by the wealth share of the bottom 50 percent, only marginally in preindustrial Germany (data from Alfani et al. 2022). This implies that inequality estimates where some of the poor are missing should be interpreted as lower bound estimates. Most importantly for my analysis, that is, regressions with unit- and time-fixed effects, including or dropping the propertyless does not change the trend of the wealth share of the poor. Even more reassuringly, the propertyless were missing from the tax registers of one town in my sample only, Heilbronn. As shown below, dropping Heilbronn from the sample does not change the results.

This dataset also differs from Schaff (2023) in that the main outcome variable — the wealth share of the bottom 20 percent of the population — has been newly constructed from the household-level wealth information. Other outcomes, such as the Top 1 percent, the middling 40 percent and the

³⁷Arguably another important provision was the legalisation of Calvinism.

Figure 10: Poor Wealth Shares in Germany with Propertyless (Bottom 50%)



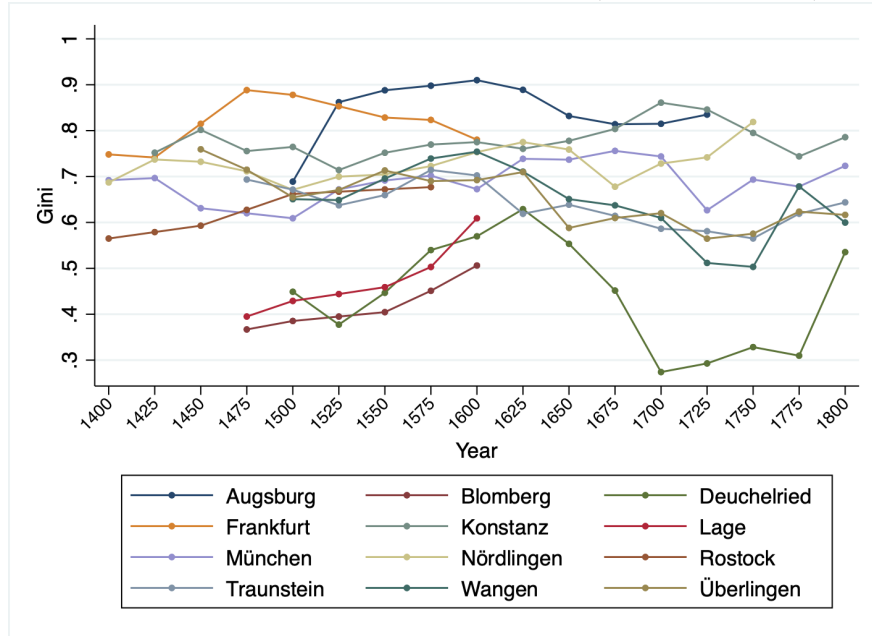
Source: See the text.

Top 90 percent wealth shares, have also been constructed newly. The use of data from communities that are observed before and after the beginning of the Reformation makes this dataset somewhat smaller than the one of Schaff (2023), which also includes communities that are only observed, for example, from 1600 onwards.

Since this study is interested in comparing localities' inequality before and after the beginning of the Protestant Reformation, I include only those 43 localities in the analysis that have before-and-after observations. All other localities are dropped from the analysis in the main text. (In a robustness check reported below I show that the main results hold when employing a different research design that makes it possible to include all localities from the original dataset.) The localities included in the analysis are: Augsburg, Blomberg, Boizenburg, Brake, Detmold, Deuchelried, Erfurt, Frankfurt a.M., Freiburg i.B., Gadebusch, Heiden, Heilbronn, Hildesheim, Horn, Kitzingen, Königshofen, Konstanz, Leonberg, Lage, Lübeck, Merhathshofen, Merkershausen, Mühlhausen, the suburbs of Mühlhausen, Munich, Niederwangen, Nördlingen, Oerlinghausen, Opfenbach, Ravensburg, Rostock, Rudlings, Schötmar, Schwäbisch Hall, Siggen, Straubing, Thann-Wohmbrechts, Traunstein, Überlingen, Umpferstedt, Wangen and Wanger Pfarr.

Figure 11 shows inequality estimates for some selected communities in the dataset. The inequality estimates are based on information from property tax registers that list all tax-paying households in a locality in a given year. Although these sources have limitations (see the main text, Schaff 2022 and Alfani et al. 2022), their use to study questions of inequality and distribution of wealth and income is a consolidated tradition in German historiography (see for example Friedrichs 1979, Warde 2006). Part of the dataset builds on secondary sources, where the data has usually been published in the form of summary tables divided into wealth brackets. A substantial part of the data come from primary sources, collected from numerous communal archives, for example in

Figure 11: Wealth Inequality in Selected Communities (Gini Coefficients), c. 1400-1800



Source: See the text.

Munich, Konstanz and Traunstein, among others (see Schaff 2022 for the places for which primary or secondary sources have been employed). The localities have been chosen based on the availability of sources for multiple periods. For several interesting places data sources have survived for single years only, for instance for Stuttgart or Cologne. These single observations were not included in the panel dataset.

The dataset covers the period from 1400 until 1800, in steps of 25 years, clustered around reference years. As a general rule, data points were clustered to their nearest cluster year, but in few cases also to the second nearest, when this was necessary to extend the time series by another period. For example, if an observation actually referred to the year 1555, then it was clustered to 1550. But an observation that referred to the year 1562 was clustered to 1575 if there was already another observation for 1550. This handling was necessary because archival tax registers are overall rare sources and not always available for the exact year one might be interested in. For the same reason, first, missing values had to be linearly interpolated (but never extrapolated),³⁸ and, second, the panel ends up being unbalanced.

B.2 Representativeness of Sample: Calculation of Average Population Size and Share of Protestant Localities

To support the argument that at least the urban part of my sample is broadly representative of the universe of localities in Germany, I have compared the population development in my sample of towns with that of a “universe” of towns. That universe comes from information about urban population in the “Bairoch-dataset” (Bairoch et al. 1988). I have extended this dataset, by adding to the Bairoch-towns population figures from De Vries (1984) and Wahl (2019). Based on this

³⁸Overall, 9.5 percent of all data points have been obtained through linear interpolation

Figure 13: Archival Sources: Tax Registers from Nördlingen (1575 and 1725)

Source: Stadtarchiv Nördlingen.

Überlingen was a Catholic city. For comparison, Figure 13 shows tax registers from Protestant Nördlingen. Simply eye-balling the documents shows a high degree of similarity between the sources, although both cities were different Imperial estates. The tax registers list household heads and their respective tax estimate. The third entry of the register on the left shows a propertyless household. Similar to Überlingen, it was marked as having “*nil*”. The high degree of homogeneity of the tax registers in (Catholic) Überlingen and (Protestant) Nördlingen does not provide any visual evidence for the notion of differential Protestant tax record-keeping.

Table 9 provides summary statistics by confessional groups of the main variables of interest, divided into the time before and after the beginning of the Reformation. Importantly, Catholic and eventually Protestant communities appear balanced in terms of pre-treatment inequality levels. For example, the Gini coefficient was 0.587 in Catholic, and 0.580 in eventually Protestant communities. The population size, soil quality and the share of urban communities are also similar in the

two samples. The differences in coordinates reflect the fact that Catholic communities tended to be more concentrated in the south of Germany.

B.5 Real Wealth Calculation for Poor Strata in Augsburg

This Appendix describes the calculation of the real wealth per poor household — understood as a household in the bottom 20 percent of the wealth distribution — in Augsburg around the beginning of the Reformation. In 1512 the total amount of wealth tax paid in Augsburg was 17,990 fl., and there were 5,479 households recorded including the propertyless (Hartung 1898: 191-192). The tax rate in that year was 0.38 percent (Gierok 2022: Appendix), which implies that the total (taxable) wealth was 47,342.11 fl. Calculations reported in the main text suggest that the average wealth share of the bottom 20 percent was 2.18 percent in my sample, which would correspond to 1,032.06 fl. held by all poor households in Augsburg, or 0.94 fl. per household. Note that this value is strikingly close to the real wealth poverty threshold of 1 fl. set by many studies of poverty in preindustrial Germany (see Dirlmeier 1978, Fischer 1982).

This number does not say much about the real value of poor households' property, so I decided to calculate the corresponding number of day wages of an unskilled labourer. An unskilled labourer earned 1.51 grams of silver per day in 1513 Augsburg (Pfister 2017: Appendix). We know that 1 guilders corresponded to 64 kreuzer in Augsburg in the period 1525–1530, and 1 kreuzer contained de jure 0.43 grams of silver (Geffcken and Häberlein 2014: 10-5-152, Weisenstein 2002: 106, Newald 1883: 6). That implies that the average taxable wealth of a household in the bottom 20 percent in Augsburg corresponded to 25.90 grams of silver, or 17.14 day wages of an unskilled worker.

C Coding of Independent Variables

This appendix describes how the main independent variables employed in the analysis in the main text and the appendix have been coded.

Protestant Reformation. A dummy that indicates whether the Protestant Reformation has been introduced in a locality by a town council or local ruler after 1517. I have also recorded the official year of introduction, and when no precise year is indicated I took as alternative date the appointment of a protestant priest by the town council. When no introduction is mentioned, or the source indicates that the Reformation had “no substantial impact”, I code the locality as Catholic. Information was taken from the *Städtebuch* (Keyser 1939, 1941, 1952, 1954, 1956, 1957, 1959, 1962, 1964, Keyser and Stooß 1971, 1974, Baltzarek et al. 1973). For communities without entry in the *Städtebuch* usually the Imperial Estate introduced the Protestant Reformation. This information has been taken from the *Städtebuch* and the *Historisches Lexikon der deutschen Länder* (Köbler 2007).

Log-taxpayers. The number of taxpayers comes from the same local tax registers that provide information about the distribution of wealth.

Log-population size. The population size of a locality has been obtained by multiplying the number of taxpayers in a given year with the presumed average household size. The household size typically

Table 9: Community Characteristics by Confessional Group

	1400-1500 (pre-treatment period)		1525-1800 (treatment period)	
	N	Mean	N	Mean
ALL COMMUNITIES:				
Log-Population Size	99	7.664	269	7.452
Soil Quality	99	0.704	269	0.726
Longitude	99	10.07	269	10.03
Latitude	99	50.28	269	49.21
Urban Communities	99	0.626	269	0.606
Wealth Share Top 1%	99	11.70	269	12.19
Wealth Share Top 10%	99	45.87	269	45.93
Wealth Share Middle 40%	99	21.87	269	21.72
Wealth Share Bottom 20%	99	2.433	269	2.089
Wealth Share 2. Decile	99	1.593	269	1.473
Wealth Share Bottom 10%	99	0.840	269	0.616
Gini	99	0.582	269	0.592
CATHOLIC COMMUNITIES:				
Log-Population Size	31	7.454	149	7.139
Soil Quality	31	0.701	149	0.717
Longitude	31	10.33	149	10.19
Latitude	31	48.08	149	48.06
Urban Communities	31	0.677	149	0.577
Wealth Share Top 1%	31	12.59	149	10.76
Wealth Share Top 10%	31	48.10	149	42.85
Wealth Share Middle 40%	31	21.94	149	23.72
Wealth Share Bottom 20%	31	2.177	149	2.406
Wealth Share 2. Decile	31	1.631	149	1.768
Wealth Share Bottom 10%	31	0.546	149	0.638
Gini	31	0.587	149	0.559
PROTESTANT COMMUNITIES:				
Log-Population Size	68	7.760	120	7.840
Soil Quality	68	0.705	120	0.738
Longitude	68	9.952	120	9.829
Latitude	68	51.29	120	50.65
Urban Communities	68	0.603	120	0.642
Wealth Share Top 1%	68	11.29	120	13.95
Wealth Share Top 10%	68	44.86	120	49.76
Wealth Share Middle 40%	68	21.83	120	19.24
Wealth Share Bottom 20%	68	2.549	120	1.694
Wealth Share 2. Decile	68	1.576	120	1.106
Wealth Share Bottom 10%	68	0.973	120	0.588
Gini	68	0.580	120	0.632

Notes: The variables “Soil Quality”, “Longitude”, “Latitude” and “Urban Communities” are time-invariant. Differences in the means between the two periods are due to changes in the composition of the sample.

assumed for preindustrial German towns is 4.5 (Minns et al. 2020: 611).

Epidemic. A dummy that indicates whether there was an outbreak of an epidemic in a locality in the previous period. Information on major outbreaks of epidemics has been taken from the *Städtebuch*. Epidemics indicated by the *Städtebuch* are for example smallpox, syphilis and plague. For those rural communities in the dataset that have no entry in the *Städtebuch* I had to make an assumption about plague occurrence. I assumed that the rural communities had the same plague occurrence as the nearest town for which an entry in the *Städtebuch* and information about the outbreak of epidemics is available. These assumptions are based the regular interaction between village and town inhabitants via urban markets in preindustrial times. Towns were daily markets where peasants from surrounding villages regularly sold agricultural products and bought goods that they could not produce themselves (Isenmann 2014: 673). For those villages that were under the administrative authority and were taxed by a nearby city that is part of the dataset, I have assumed the same occurrence of epidemics as in the city. For example, for the rural community of Niederwangen I assume the same plague occurrence as for the nearby city of Wangen. For those villages that were not under the administrative authority of a city in the dataset I have assumed the same occurrence of epidemics as in the closest town with an entry in the *Städtebuch*.

Log-university distance. Log-distance (km) of a locality to the closest University in every given year (own calculations). Locations and opening years of German universities are taken from Schilling (1994: 330).

Conflict exposure. A dummy that indicates whether a locality was exposed to battle action or a siege within a radius of 25 km. The data are from Schaff (2020).

Urban community. A dummy that indicates whether a locality was a city. I consider all those localities as cities that have an entry in the *Städtebuch*.

Soil quality. Index of agricultural potential of a locality. Data was taken from Ramankutty et al. (2002). The index is a composite indicator that takes into account soil quality itself, but also climatic conditions.

Distance to Wittenberg. Distance of a locality to the city of Wittenberg, measured in kilometres (own calculations).

Distance to Geneva. Distance of a locality to the city of Geneva, measured in kilometres (own calculations).

Distance to Zurich. Distance of a locality to the city of Zurich, measured in kilometres (own calculations).

Number of Monasteries closed. A continuous variable indicating the number of monasteries within 20km from a locality, that were closed during the Reformation period (beginning in 1517) until 1600. Data on the geographical position, opening, and, if applicable, closure date of 3094 monasteries in Germany (as reported in Jürgensmeier and Schwerdtfeger (2005)) was taken from Cantoni et al. (2018). For those single entries without geographical position in the Cantoni et al. (2018) dataset,

I cross-checked against the entries in Jürgensmeier and Schwerdtfeger (2005) (from where Cantoni et al. (2018) have obtained the information) to locate and geo-reference the single monasteries.

Church ordinance. A dummy that indicates whether a locality had a so-called church ordinance (*Kirchenordnung*) from 1522 onward. The information come from Sehling (1902), who provides word-by-word editions of the ordinances. I count a community as having a church ordinance if it introduced the ordinance itself, which was usually the case for urban communities, or if the Imperial estate to which the locality belonged introduced an ordinance.

Church ordinance with begging restriction. A dummy that indicates whether a locality had a church ordinance that includes restrictions of begging. These restrictions could be, for instance, denying specific groups the right to beg, restricting begging to certain areas of the locality, or prohibiting begging altogether. The information come from Sehling (1902).

Seaside locality. A dummy that indicates localities that lie within 10 kilometres of the seaside.

Imperial city. A dummy that indicates whether a locality was an Imperial city. This information has been obtained from the *Städtebuch*. Information about the imperial rank of each estate (that is, elector, bishop, imperial city, and so on) have been obtained from Zeumer (1913).

Hanse city. A dummy that indicates whether a city was member of the Hanseatic league, as indicated by Dollinger (1981: 68).

D Additional Results

D.1 Alternative Clustering of Standard Errors

In the main analysis standard errors have been clustered at the locality level. In this section I report results with standard errors clustered at the level of imperial estates, or in other words territories. The results in 10 suggest that the clustering of standard errors does not make a substantial difference. Standard errors are slightly larger, but significance levels are still high, for the bottom 20 percent wealth share as well as for the bottom 10 percent wealth share.

D.2 Communities with Propertyless in the Wealth Distributions

In this section I estimate the effect of Protestantism on the wealth share of the poor, including only those localities in the sample that record the propertyless population in the tax registers. Luckily, only one community in the sample does not record the propertyless, Heilbronn. Intuitively one would expect that dropping this community increases the negative effect of the Reformation on the wealth share of the poor. The results below show that this is the case. The effect of Protestantism on the wealth share of the lowest fifth of the population is slightly larger and more significant, compared to the baseline results in the main text. This result holds regardless of whether controls are added or not, and suggests that the baseline results are to be interpreted as lower-bound estimates.

Table 10: Alternative Clustering of Standard Errors

	(1)	(2)	(3)	(4)
	Bot. 20%	Bot. 10%	Bot. 20%	Bot. 10%
Protestant \times Post-Reform.	-1.004** (0.460)	-0.431** (0.184)	-1.023** (0.411)	-0.428** (0.163)
Controls	NO	NO	YES	YES
Locality FE	YES	YES	YES	YES
Time FE	YES	YES	YES	YES
Observations	368	368	368	368
Communities	43	43	43	43
R^2	0.165	0.147	0.228	0.235

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 11: Communities with Propertyless in the Wealth Distributions

	(1)	(2)
	Bot. 20%	Bot. 20%
Protestant \times Post-Reform.	-1.076*** (0.386)	-1.108*** (0.361)
Localities w/o propertyless dropped	YES	YES
Controls	NO	YES
Locality FE	YES	YES
Time FE	YES	YES
Observations	354	354
Communities	42	42
R^2	0.163	0.230

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

D.3 Alternative Treatment Variable Coding

In this section I employ an alternatively coded treatment variable. In the main text I employ a treatment variable that switches on in the year 1517, the year in which the Reformation period begins with the publication of Luther's 95 theses, if a community eventually became Protestant until 1600. This is the standard approach in the published literature (Rubin 2014, Cantoni 2015, Dittmar and Meisenzahl 2020). The reasoning behind this coding is that communities might have lived *de facto* as if it were Protestant before the official introduction, but often political authorities delayed the official adherence to Protestantism because of geo-political considerations. Here I estimate the baseline specification employing a treatment variable that switches on differentially for each community, in the year it became officially Protestant.

Table 12 shows that the alternative coding of the treatment variable does not change the DD

Table 12: Alternative Treatment Variable Coding

	(1) Top 10%	(2) Mid. 40%	(3) Gini	(4) Bot. 20%	(5) Bot. 10%
Protestant \times Post $_{i,t}$	1.366 (2.006)	0.599 (1.378)	0.012 (0.019)	-0.761** (0.364)	-0.356** (0.150)
Locality FE	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
Observations	368	368	368	368	368
Communities	43	43	43	43	43
R^2	0.189	0.287	0.257	0.159	0.144

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

results. There is still no significant effect of becoming Protestant on top and middling wealth shares. Instead, there is a negative and sizeable effect on the wealth shares of the poor strata of the population. These results reinforce the findings in the main text.

D.4 Spatial Dependence

Table 13 reports the baseline specifications when standard errors are adjusted for spatial autocorrelation (Conley 1999; see also Gibbons et al. 2015: 137). The Conley-correction is considered sufficient to assess the significance of regression results in spatial datasets (Voth 2021: 259). I calculate spatial correlation-adjusted standard errors following the routine of Fetzner (2014), who builds on Hsiao (2010). I assume that spatial autocorrelation linearly decreases with distance between localities. Standard errors have been adjusted for different cutoff-distances until which spatial correlation is assumed to decrease to zero, from 50 to 1000 kilometres.

Table 13: Baselines Estimates: Spatial Autocorrelation-Adjusted Std. Errors

	(1) Bot. 20%	(2) Bot. 20%	(3) Bot. 20%	(4) Bot. 20%
Protestant \times Post-Reform.	-1.004*** (0.293)	-1.004*** (0.306)	-1.004*** (0.277)	-1.004*** (0.285)
Cutoff	50km	100km	500km	1000km
Observations	368	368	368	368
R^2	0.025	0.025	0.025	0.025

Notes: Estimation method is OLS. All regressions include a full set of locality and time fixed effects. Conley-standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Columns 1 to 4 show the results of the baseline specification in the main text, taking the wealth share of the bottom 20 percent of the population as dependent variable. The difference-in-difference estimates remain unchanged and significance remains very high, regardless of the cutoff distance.

These results suggests that the baseline results are robust to adjusting standard errors to allow for spatial autocorrelation and that spatial dependence is not a problem for this study.

One might still wonder whether there existed actual spatial autocorrelation among localities, which could inflate t-statistics in the regressions. For addressing this potential concern, Moran's I has been calculated to evaluate whether there is spatial autocorrelation in the regression residuals. This has been done for the respective cross-sections of the panel-dataset, as done by Kelly (2019: 16-17, 27).

Table 14: Spatial Dependence in Regression Residuals (Global Morans'I-Statistic)

	(1)	(2)	(3)
Year	I-statistic	Z-score	P-value
1400	-0.218	-0.868	0.193
1425	-0.156	-0.308	0.379
1450	-0.014	0.423	0.336
1475	-0.373	-3.442	0.000
1500	-0.124	-1.277	0.101
1525	0.001	0.334	0.369
1550	0.092	1.734	0.041
1575	-0.061	-0.87	0.192
1600	0.011	0.612	0.270
1625	0.466	6.518	0.000
1650	0.726	9.243	0.000
1675	-0.059	-0.195	0.423
1700	-0.052	-0.045	0.482
1725	-0.075	-0.272	0.393
1750	0.071	1.252	0.105
1775	0.064	1.111	0.133
1800	0.009	0.684	0.247

The specification includes all controls of the main analysis, but without any of the fixed effects, due to the cross-sectional data structure. Instead I have added longitude, latitude and the interaction of longitude and latitude. Kelly (2019: 17, 22) considers positive z-scores of two and greater as indicative of relevant spatial autocorrelation. Table 14 shows that the z-score for almost all years of the analysis is considerably below the critical value of two. This indicates that positive spatial autocorrelation, which would be a reason of concern, is not an issue for this study. Only in two years, 1625 and 1650, are there signs of significant spatial autocorrelation, that is, a tendency that localities with similar values cluster. However, the degree of spatial-autocorrelation is quite low. It can only be speculated why spatial dependence occurs precisely in this period. One possibility could be that the Thirty Years' War (1618-50) played a role. Since this war was a global negative shock to inequality across Germany, affecting practically all communities (Alfani et al. 2022), a shock that due to the cross-sectional setting cannot be absorbed by time fixed effects, it could be that it induced spatial dependence in this particular period. Beyond this historically exceptional period

and considering the results using the Conley-correction of standard errors, spatial dependence seems to be a negligible concern for this analysis. Moreover, note that more sophisticated methods to deal with spatial dependence are not feasible because of the highly unbalanced structure of the panel. For example, because a spatial weighting matrix cannot be calculated with the available data, estimating a spatial autoregressive (SAR) model is impossible (Gibbons et al. 2015).

D.5 Wild Cluster Bootstrap Statistics

To address concerns about uncertainty of inference because of the small sample size, I calculate p-values and confidence intervals obtained with Wild Cluster Bootstrapping — a re-sampling technique aimed at obtaining statistics and distributional properties that reflect the true underlying distribution of the data when we have a small number of cluster — for the baseline DD estimates in the main text. I follow the routine of Roodman et al. (2019).

Table 15: Baselines Estimates: Wild Cluster Bootstrap Statistics

	(1) Top 10%	(2) Mid. 40%	(3) Gini	(4) Bot. 20%	(5) Bot. 10%
Protestant×Post-Reform.	3.072 (2.334)	-0.161 (1.608)	0.027 (0.023)	-1.004** (0.383)	-0.431*** (0.144)
Wild Cluster Bootstrap p-value	0.195	0.921	0.268	0.019	0.006
Wild Cluster Bootstrap CI (lower bound)	-1.835	-3.351	-0.021	-1.783	-0.739
Wild Cluster Bootstrap CI (upper bound)	8.034	3.092	0.078	-0.219	-0.147
Communities	43	43	43	43	43
Observations	368	368	368	368	368
R^2	0.196	0.286	0.262	0.165	0.147

Notes: Estimation method is OLS. All regressions include a full set of locality and time fixed effects. Normal standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

To assess whether the small number of clusters in the dataset causes significant uncertainty, one can compare the normal significance levels (as indicated by asterisks) and the Wild Cluster Bootstrap p-values and confidence intervals reported in Table 15. None of the significance levels change when considering the alternative statistics. The effect of the introduction of the Reformation remains insignificant when taking top or middling wealth shares or the Gini coefficient as dependent variables. However, there is a negative effect at the same significance level when considering the wealth shares of poor strata of the population. These results suggest that the small sample size and number of clusters in the data do not cause relevant uncertainty of inference.

D.6 Extended Sample

As a further robustness check, I employ an extended sample to test the effect of Protestantism, including all communities that are part of the dataset of Schaff (2022), also those for which only information before or after the beginning of the Reformation is available. For example, I include also information about the cities of Koblenz (Catholic) and Hersfeld (Protestant), although they

are only observed after the Reformation began. This implies that I am interested not only in estimating differences within localities, but also differences between localities. For that reason I follow the random effects approach (Wooldridge 2002, Greene 2012). This approach is often considered inferior to the fixed-effects approach, but it has the advantage that it allows me to increase the size of the sample considerably, by 22 localities, and capture between-unit variation. The results should therefore be interpreted with caution.

Table 16: Extended Sample: Random Effects Estimates

	(1) Top 10%	(2) Mid. 40%	(3) Gini	(4) Bot. 20%	(5) Bot. 10%
Protestant \times Post-Reform.	2.101 (2.097)	0.290 (1.410)	0.018 (0.020)	-0.673** (0.313)	-0.240** (0.121)
Random effects	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES
Observations	504	504	504	504	504
Communities	75	75	75	75	75
R^2	0.198	0.273	0.258	0.118	0.116

Notes: Estimation method is GLS. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

The coefficients in Table 16 suggest that the main results in the main text hold, even in this extended dataset using between and within-variation of units. There is no statistically significant effect of the Reformation on top or middling wealth share. Yet there is a significant reduction of the wealth shares of the poor parts of the population.

D.7 Time-Invariant Controls

The evidence reported in the main text suggests that the main results of the paper are robust to the inclusion of some economic, demographic and institutional observable time-variant characteristics. In this section I assess the robustness of the main results to the inclusion of several time-invariant locality-specific characteristics. Because of the panel-structure of the dataset I interact these time-invariant characteristics with time dummies. This is a demanding specification given the small size of the dataset, because the interaction terms consume many degrees of freedom.

I test for variables that might have influenced inequality and the adoption of the Reformation, and might have therefore induced selection bias. I first focus on variables related to economic fundamentals and city status: agricultural potential, seaside location, whether a locality was an urban community, an Imperial city, a member of the Hanseatic League or was part of a bishopric (Columns 1 to 6). All regressions include a full set of locality and time fixed effects. The results show that the main finding is robust to the inclusion of all these controls: the introduction of the Reformation led to a reduction of the wealth share of poor strata. The coefficient of interest remains negative, sizeable and significant.

Second, I focus on geographic location: whether community was up in the North of Germany, the

Table 17: Time-Invariant Controls

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%
Protestant \times Post-Reform.	-1.126*** (0.327)	-1.034** (0.401)	-0.724* (0.379)	-1.355*** (0.345)	-1.011** (0.413)	-0.942** (0.390)	-1.038** (0.420)	0.154 (0.427)	-0.030 (0.410)
Soil quality \times time	YES	NO	NO	NO	NO	NO	NO	NO	NO
Seaside location \times time	NO	YES	NO	NO	NO	NO	NO	NO	NO
Urban community \times time	NO	NO	YES	NO	NO	NO	NO	NO	NO
Imperial city \times time	NO	NO	NO	YES	NO	NO	NO	NO	NO
Hanse city \times time	NO	NO	NO	NO	YES	NO	NO	NO	NO
Bishopric \times time	NO	NO	NO	NO	NO	YES	NO	NO	NO
Northern community \times time	NO	NO	NO	NO	NO	NO	YES	NO	NO
Dist. to Wittenberg \times time	NO	NO	NO	NO	NO	NO	NO	YES	NO
Longitude & latitude \times time	NO	NO	NO	NO	NO	NO	NO	NO	YES
Observations	368	368	368	368	368	368	368	368	368
Communities	43	43	43	43	43	43	43	43	43
R^2	0.211	0.215	0.389	0.303	0.220	0.207	0.215	0.295	0.345

Notes: Estimation method is OLS. All regressions include a full set of locality and time fixed effects. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

distance to the “capital” of the Reformation Wittenberg and geographic coordinates (Columns 7 to 9). Whether a place was up in the North does not seem to matter, but once the distance to Wittenberg or coordinates are introduced the effect disappears. This is, of course, little surprising because Reformation adoption depended heavily on the distance to Wittenberg and, therefore, geographic location, as argued in the main text. This fact is exploited in the instrumental variable analysis.

D.8 Heterogeneity Analysis

In the main text I have reported robustness checks where I drop different part of the dataset, to evaluate whether these drive the results. I have also shown the distribution of wealth shares in Protestant and Catholic communities, which provided reassurance that the results are not driven by outliers. In this Appendix I explore further potential heterogeneity among the communities in the dataset.

I investigate urban versus rural, Northern versus Southern and Eastern versus Southern Protestant communities, in Table 18. I interact these characteristics with being Protestant in the post-period, holding the respective lower-order interactions constant. A first observation across Columns 1 to 6 is that none of the included interactions leads to the main coefficient of interest losing significance or size. This implies that geographic location or city status do not drive the main results. However, the coefficients of the interaction terms provide further insights. In Columns 1 and 2 I check whether poor strata experienced a differential development of their wealth shares in cities. The coefficients suggest that Protestant poor strata were better off in cities than in rural communities. This is unsurprising given that cities were economically vibrant places where, for example, also corporate organisations helped poor people (van Bavel and Rijpma 2016: 181).

In Columns 3 and 4 I evaluate whether Protestant poor strata had a differential wealth share development in Northern localities. I define as “Northern” all localities that on the map in Figure 4 have the same or lie above the latitude of Hildesheim. Places in the north of Germany were in

several ways different. They were often associated with the German Hanse, that is oriented towards long-distance trade, but they were also particularly detached from the influence of the Holy Roman Emperor and Imperial institutions (Whaley 2012: 39). The results indicate, however, no significant difference. In Columns 5 and 6 I evaluate whether Protestant localities in the Eastern half of Germany experienced a differential wealth share change of poor strata. I defined as “Eastern” all localities to the East of Mühlhausen on the map in Figure 4. This geographic dividing line reveals some significant differences. Eastern places have higher wealth shares of poor strata.

Table 18: Difference-in-Differences Estimates: Heterogeneity of Wealth Share Changes of Poor Strata

	(1)	(2)	(3)	(4)	(5)	(6)
	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%	Bot. 20%
Protestant×Post-Reform.×City	2.041** (0.821)	1.914** (0.860)				
Protestant×Post-Reform.×North			-0.218 (0.505)	-0.429 (0.467)		
Protestant×Post-Reform.×East					1.658** (0.618)	1.116* (0.622)
Protestant×Post-Reform.	-2.276*** (0.671)	-2.260*** (0.755)	-0.958** (0.417)	-0.931** (0.389)	-1.541*** (0.497)	-1.399*** (0.512)
Controls	NO	YES	NO	YES	NO	YES
Locality FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES
Observations	368	368	368	368	368	368
Communities	43	43	43	43	43	43
R^2	0.198	0.248	0.165	0.230	0.179	0.234

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. In Columns 1 and 2 I also control for Post-Reform.×City. In Columns 3 and 4 I also control for Post-Reform.×North. In Columns 5 and 6 I also control for Post-Reform.×East. Note that the interaction between being Protestant and City, North and East respectively is absorbed by the locality fixed effects. *** p<0.01, ** p<0.05, * p<0.1

D.9 Continuity in the Number of Taxpayers

One might wonder whether the lower wealth shares of poor strata in Protestant places were the result of different taxation principles in localities that introduced the Reformation, which could have had an impact on the recorded distribution of wealth. For example, it could be that local bureaucracies in Protestant places recorded poor people more rigorously in the tax registers. Alternatively, one could conjecture that Protestant places recorded serfs, which on average were most likely poorer than the rest of the population, more thoroughly than Catholic administrations. If that were the case, one would expect a higher number of taxpayers in Protestant compared to Catholic places.

In Table 19 I check whether there are any discontinuities in the number of taxpayers during the early phase of the Reformation. The set-up is again similar to Equation 1, but I take the local number of taxpayers (in logarithm) as dependent variable, which can be easily calculated from the tax registers

Table 19: Difference-in-Differences Estimates: Number of Taxpayers

	(1)	(2)	(3)
	ln-Taxpayers	ln-Taxpayers	ln-Taxpayers
Protestant \times Post-Reform.	-0.029 (0.082)	-0.032 (0.084)	-0.002 (0.097)
Period	1500-1550	1500-1575	1500-1600
Observations	127	163	196
Communities	43	43	43
R^2	0.221	0.211	0.208
Mean of dependent variable	5.575	5.601	5.615

Notes: Estimation method is OLS. All regressions include a full set of locality and time fixed effects. Standard errors clustered at locality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

that were used to calculate inequality measures. Moreover, I restrict the dataset to three different periods, from the time just before the Reformation began until three different cutoff years in the early phase of the Reformation: 1500-1550, 1500-1575 and 1500-1600. All coefficients are close to zero and slightly negative, but not statistically significant. This suggests that if anything there were slightly less taxpayers in Protestant places. These results provide no evidence for the idea that there existed differential recording of taxpayers in Protestant compared to Catholic places. Of course, this does not rule out entirely that Protestant bureaucracies recorded poor people more thoroughly, but it makes it unlikely.

One might think that another way to test for changes in the recording of poor people between the two confessions is to compare the share of recorded taxpayers over the total population, and use this as dependent variable. In practice, however, this comparison would not be informative, because population figures for preindustrial communities, such as the widely used ones reported in [Bairoch, 1988], are based on the number of taxpayers (multiplied with an assumed average household size) c

D.10 Additional Results for the IV Analysis

This Appendix reports several additional results for the instrumental variable analysis, which exploits the distance to Wittenberg as a source of exogenous variation of the adoption of Protestantism. Table 20 shows the correlations between the potentially endogenous variable, the binary indicator of the adoption of Protestantism in a community until 1600, and several possible predictors, including the distance to Wittenberg. The set-up is conceptually equivalent to a first-stage regression.³⁹

The results indicate that distance to Wittenberg is indeed a strong predictor of Reformation adoption, suggesting that the instrument-relevance condition is met. The further a community was away from Wittenberg, the less likely it was to be Protestant in 1600. This relationship is robust

³⁹Note that the first-stage varies when the complete panel of communities is employed to estimate the causal effect of the Reformation on the wealth distribution over time. The set-up of the main analysis is slightly different, containing various fixed effects and the instrumented variable interacted with the post-treatment indicator. The relationship between the instrument and the instrumented variable is nevertheless strong.

Table 20: Correlates of the Adoption of Protestantism until 1600

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Protestant	Protestant	Protestant	Protestant	Protestant	Protestant	Protestant	Protestant	Protestant
Dist. to Wittenberg	-0.003*** (0.000)	-0.002** (0.001)	-0.002*** (0.001)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.003*** (0.000)	-0.002** (0.001)
Dist. to Zurich		0.001 (0.000)							
Dist. to Geneva			0.000 (0.000)						
Soil Quality				0.318 (0.423)					
ln-Pop. in 1500					0.105 (0.078)				
Top 1%						0.016 (0.014)			
Urban Community					-0.198 (0.269)	-0.059 (0.207)	0.150 (0.148)	0.113 (0.111)	
Hanse City							-0.105 (0.200)		
Imperial City								0.116 (0.124)	
Longitude									-0.109*** (0.031)
Latitude									0.080* (0.040)
Constant	1.534*** (0.135)	1.086** (0.453)	1.165** (0.552)	1.338*** (0.226)	0.835 (0.570)	1.428*** (0.168)	1.527*** (0.193)	1.511*** (0.148)	-1.721 (2.336)
Observations	43	43	43	43	43	43	43	43	43
R^2	0.496	0.509	0.501	0.503	0.547	0.536	0.514	0.521	0.600

Notes: Estimation method is OLS (linear probability model). Standard errors clustered at locality level in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

to the inclusion of other theoretically plausible predictors, such as the distance to other centres of the Reformation in Europe, Zurich and Geneva (Columns 2 and 3). Similarly, several economic indicators — soil quality, the population size of a community in 1500, the top 1% wealth share in a community in 1500 — do not eliminate the significance of the distance to Wittenberg (Columns 4 to 6).⁴⁰ The inclusion of whether a city had a special constitutional status, that is, was member of the Hanse or an independent Imperial city (Columns 7 and 8), does not change the picture either, as do geographic coordinates (Column 9). Cantoni (2012, 2015) performs similar analyses and reports analogous results.

In the instrumental variable analysis the exclusion restriction amounts to assuming that, conditional on locality and time fixed effects and several covariates, the distance to Wittenberg affected inequality only through influencing the probability of a community of becoming Protestant. The assumption would be violated if the distance to Wittenberg caused higher inequality independently of the Protestant Reformation. While this assumption is itself untestable, Table 21 provides further evidence that the exclusion restriction might hold. I have regressed several wealth share and inequality indicators in 1500 on the distance to Wittenberg.

⁴⁰Note that when regressing Protestantism on the population size and top wealth shares one has to control for whether a community was a city. The reason being that cities had almost mechanically a larger population, and also higher top wealth shares, but cities were also more likely to adopt the Reformation (Scribner 1994: 7-14).

Table 21: Distance to Wittenberg and Wealth Distribution in 1500

	(1) Top 10%	(2) Mid. 40%	(3) Bot. 20%	(4) Gini
Dist. to Wittenberg	0.018 (0.020)	0.002 (0.013)	-0.003 (0.003)	0.000 (0.000)
Observations	43	43	43	43
R^2	0.020	0.001	0.037	0.003

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

An instrument that provides exogenous variation to preexisting determinants of unequal wealth distribution should not produce coefficients that point towards significantly lower inequality the further a locality is away from Wittenberg in 1500. This is what the coefficients indicate. They are statistically not different from zero, and almost zero. This result supports the exclusion restriction of the distance-to-Wittenberg instrument.

Table 22 reports reduced form or “intention to treat” estimates of the effect of the instrument for Protestantism on wealth distribution and inequality. Intuitively, one can think of these estimates as the effect of “encouraging” a locality to adopt the Protestant Reformation on inequality. Significant reduced-form estimates make the causal interpretation of instrumental variable estimates more credible (Angrist and Pischke 2009: 213).

Table 22: Reduced Form Estimates

	(1) Top 10%	(2) Top 1%	(3) Mid. 40%	(4) Gini	(5) Bot. 20%	(6) Bot. 10%
Dist. to Wittenberg×P _t	-0.008 (0.008)	-0.003 (0.005)	0.001 (0.006)	-0.000* (0.000)	0.006*** (0.001)	0.002*** (0.001)
Locality FE	YES	YES	YES	YES	YES	YES
Time FE	YES	YES	YES	YES	YES	YES
Observations	368	368	368	368	368	368
Communities	43	43	43	43	43	43
R^2	0.191	0.073	0.286	0.268	0.185	0.166

Notes: Estimation method is OLS. Standard errors clustered at locality level in parentheses. *** p<0.01, ** p<0.05, * p<0.1

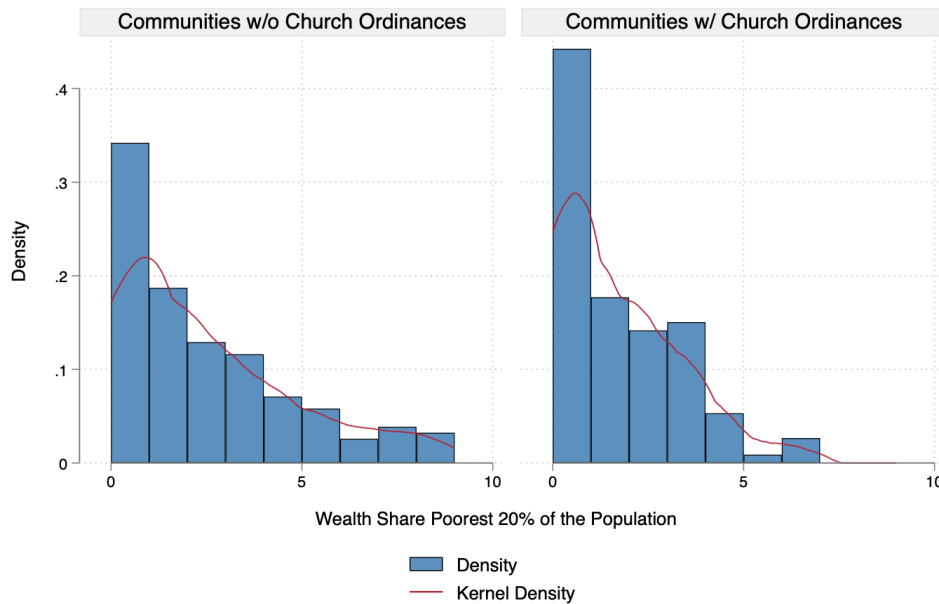
The results are in line with the pattern of the second stage and the DD results in the main text. While there was no significant effect on top or middling wealth shares, inequality as measured with the wealth shares of the poor strata got larger the closer a locality was to Wittenberg after the Reformation began. This result reinforces the plausibility of the findings in the main text. More evidence for Wittenberg’s economic insignificance comes from Becker and Woessmann (2009: 559-563). They test whether the distance to Wittenberg was related to a battery of pre-Reformation

economic and human-capital indicators, such as urban population or the presence of a school in 1517, but do not find any significant relationship. This further suggests that the distance to Wittenberg did not matter for economic outcomes and thus for economic inequality via other channels than the Reformation in preindustrial Germany.

D.11 Additional Results for the Analysis of Church Ordinances

This Appendix provides more insights about the effect of Protestant Church Ordinances on poor strata's wealth share. One might again be concerned that the effect is driven by outliers. Figure 14 therefore shows the distribution of wealth shares of the bottom 20 percent in communities with and without ordinances. Outliers do not seem to be an issue. But already from the raw data it is visible that places that had an ordinance also had a distribution of bottom wealth shares that was more skewed towards higher inequality. The right tail of the distribution above five percent is almost entirely missing in places with ordinance.

Figure 14: Distribution of Wealth Shares of the Bottom 20% in Communities with and without Church Ordinances

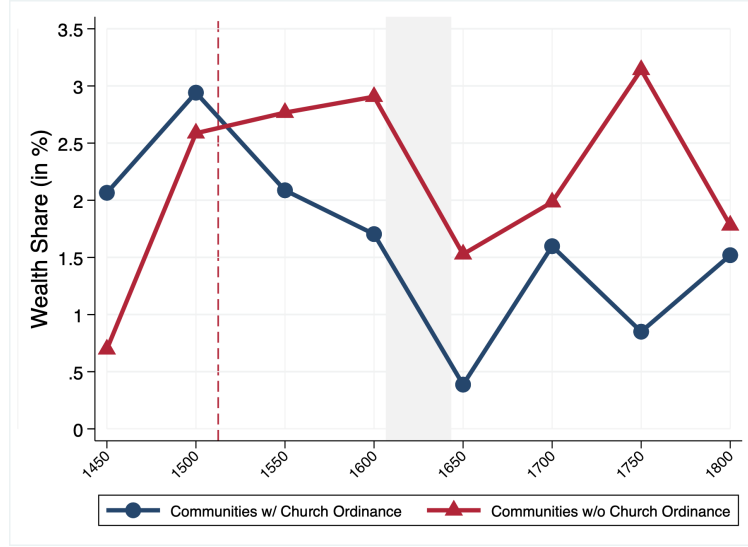


Notes: Distribution of wealth shares of the bottom 20 percent of the population in Communities with and without Church Ordinances (1525-1800). The red line is the kernel density estimate. The sample size is 269 observations.

Figure 15 report the wealth shares of poor strata (calculated analogously to Figure 1) for communities with and without ordinances. They show a clear pattern: a decline of the wealth shares of poor people from the sixteenth to the eighteenth century relative to the initial value in places with church ordinances. At the same time wealth shares of the poor stagnate in places without ordinances over the centuries. Instead, before the beginning of the Reformation both groups seem to have followed similar developments.

In Figure 16 I plot the estimation results of a flexible DD model, taking the wealth share of the bottom 20 and 10 percent as dependent variables, and the introduction of a church ordinance in

Figure 15: Wealth Share of the Bottom 20% in Communities with and without Protestant Church Ordinances



Notes: Values were collapsed into 50-year intervals and represent half-century-averages. To avoid that communities with more observations dominate the trend, every community has the same weight in the average. Because of the uneven number of years and the low number of observations in 1400, I collapsed the values for the years 1400 until 1450 into one data point. The vertical red line represents the beginning of the Reformation period, the grey box the Thirty Years' War.

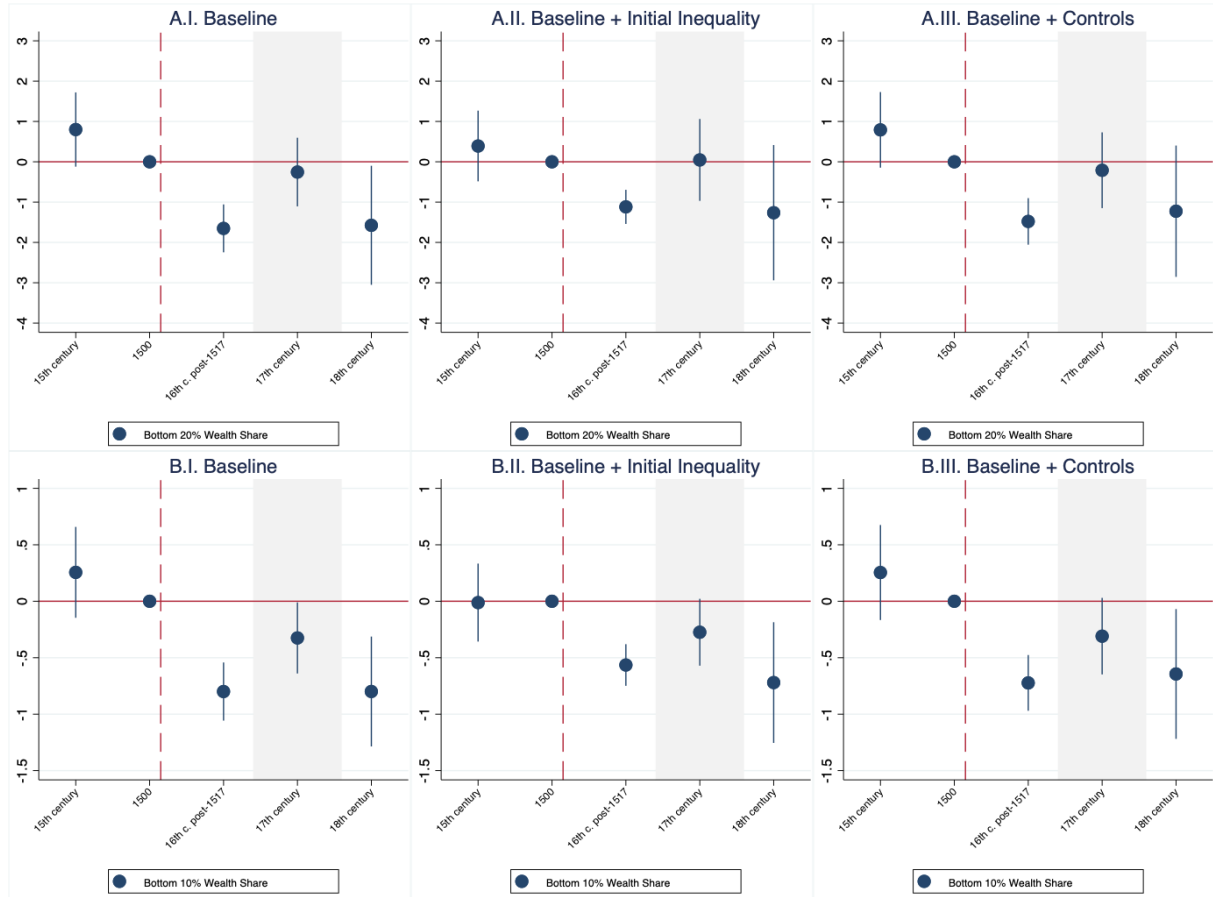
Protestant places as treatment. I estimate variants of the following specification, which is similar to the triple difference-in-difference specification in the main text:

$$I_{i,t} = \alpha_i + \pi_t + \sum_{t=1}^4 \beta_t \text{Prot}_i \times \text{Century}_t \times \text{Ordinance}_i + \theta \text{Prot}_i \times \text{PostReform}_t + \gamma' \mathbf{X}_{i,t} + \epsilon_{i,t} \quad (6)$$

The main difference is the inclusion of an interaction term between the treatment status indicator and a set of four century dummies (Century_t) covering the pre- and post-treatment periods. The beta (β) coefficients are the main coefficients of interest. This setup allows me to evaluate, first, whether the relationship between the introduction of a church ordinance and the wealth share of the lower classes of society changed over time and whether short- and long-term effects were similar. As for the flexible estimates of the effect of the Reformation in the main text, I take the year 1500 as reference category and divide the remaining period of observation into four intervals corresponding to centuries: from 1400 until about 1475; from 1525 just after the Reformation began until 1600, from 1625 until 1700; and from 1725 until the end of the period of observation in 1800.

The results show that the introduction of church ordinances in Protestant places led to a reduction of the wealth share of the bottom fifth and the bottom 10 percent of the population. The crisis-ridden seventeenth century interrupted that pattern, possibly because of the impact of the exceptionally destructive Thirty Years' War, as argued in the main text. The coefficient for the fifteenth century indicates insignificant larger values in eventually Protestant communities. However, in Frame 2 such initial wealth share differences are at least partially controlled for, leaving still a sizeable and

Figure 16: Wealth Share of Bottom 20% and Bottom 10% in Communities with and without Church Ordinance (Flexible Difference-in-Differences Estimates)



Notes: Regression estimates of the wealth share of the bottom 20 percent (Panel A) and the bottom 10 percent of the population (Panel B) in communities with Church ordinance before and after the Reformation (vertical red line), with respect to communities without church ordinance (horizontal red line). The omitted reference year is 1500. The estimation method is OLS. All specifications have a full set of community fixed effects and time fixed effects. Standard errors clustered at locality level. Confidence intervals indicate significance at the 95-percent level. The vertical red line represents the beginning of the Reformation period, the grey box the Thirty Years' War.

negative effect of church ordinances on the wealth shares of poor strata. Overall, the results are similar to the flexible DD estimates of the effect of Protestantism, and confirm the DD estimates of the effect of church Ordinances in the main text.

References Appendices

- Alfani, G., V. Gierok, and F. Schaff (2022). Economic Inequality in Preindustrial Germany, ca. 1300 – 1850. *Journal of Economic History* 82(1), 87–125.
- Angrist, J. and J. Pischke (2009). *Mostly Harmless Econometrics*. Princeton: Princeton University Press.
- Bairoch, P., J. Batou, and P. Chèvre (1988). *The Population of European Cittyies from 1300 to 1850*. Geneva: Librairie Droz.
- Baltzarek, F., J. Pradel, and R. Sandgruber (1973). *Österreichisches Städtebuch, Bd. 3: Die Städte Vorarlbergs*. Wien: Verlag der Österreichischen Akademie der Wissenschaften.
- Blickle, P. (2015). *Die Reformation im Reich*. Stuttgart: Kohlhammer.
- Cantoni, D. (2012). Adopting a New Religion: The Case of Protestantism in 16th Century Germany. *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.
- Conley, T. (1999). GMM estimation with cross-sectional dependence. *Journal of Econometrics* (92), 1–45.
- De Vries, J. (1984). *European Urbanization, 1500-1800*. London: Methuen.
- Dirlmeier, U. (1978). *Untersuchungen zu Einkommensverhältnissen und Lebenshaltungskosten in oberdeutschen Städten des Spätmittelalters*. Heidelberg: Universitätsverlag.
- Dittmar, J. and R. Meisenzahl (2020). Public Goods Institutions, Human Capital and Growth: Evidence from German History. *Review of Economic Studies* 87(2), 959–996.
- Dollinger, P. (1981). *Die Hanse*. Stuttgart: Kröner.
- Fetzer, T. (2014). Can workfare programs moderate violence? Evidence from India. *Unpublished manuscript*, 1–63.
- Fischer, W. (1982). *Armut in der Geschichte. Erscheinungsformen und Lösungsversuche der Sozialen Frage in Europa seit dem Mittelalter*. Göttingen: Vandenhoeck & Ruprecht.
- Friedrichs, C. (1979). *Urban Society in an Age of War: Nördlingen, 1580-1720*. Princeton: Princeton University Press.
- Geffcken, P. and M. Häberlein (2014). *Rechnungsfragmente Der Augsburger Welser-Gesellschaft*

- (1496-1551): *Oberdeutscher Fernhandel Am Beginn Der Neuzeitlichen Weltwirtschaft*. Stuttgart: Franz Steiner Verlag.
- Gibbons, S., H. Overman, and E. Patacchini (2015). Spatial Methods. In G. Duranton, J. Henderson, and W. Strange (Eds.), *Handbook of Regional and Urban Economics. Volume 5A*, pp. 115–168. Amsterdam.
- Gierok, V. (2022). Taxation, Credit and Public Expenditure in Urban Germany, 1400-1800. *Unpublished manuscript*.
- Greene, W. (2012). *Econometric Analysis, 7th edition*. Harlow: Pearson.
- Hartung, J. (1898). Die direkten Steuern und die Vermögensentwicklung in Augsburg von der Mitte des 16. bis zum 18. Jahrhundert. *Jahrbuch für Gesetzgebung, Verwaltung und Volkswirtschaft im Deutschen Reich* 22(1), 167–209.
- Hsiang, S. (2010). Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences* 107(35), 15367–15372.
- Isenmann, E. (2014). *Die Deutsche Stadt im Mittelalter 1150-1550*. Köln: Böhlau.
- Kelly, M. (2019). The Standard Errors of Persistence. *Unpublished manuscript*, 1–45.
- Keyser, E. (1939). *Deutsches Städtebuch, Bd. 1 Nordostdeutschland*. Stuttgart: Kohlhammer.
- Keyser, E. (1941). *Deutsches Städtebuch, Bd. 2 Mitteldeutschland*. Stuttgart: Kohlhammer.
- Keyser, E. (1952). *Deutsches Städtebuch, Bd. 3.1 Niedersachsen und Bremen*. Stuttgart: Kohlhammer.
- Keyser, E. (1954). *Deutsches Städtebuch, Bd. 3.2 Westfalen*. Stuttgart: Kohlhammer.
- Keyser, E. (1956). *Deutsches Städtebuch, Bd. 3.3 Landschaftsverband Rheinland*. Stuttgart: Kohlhammer.
- Keyser, E. (1957). *Deutsches Städtebuch, Bd. 4.1 Hessen*. Sigmaringen: Kohlhammer.
- Keyser, E. (1959). *Deutsches Städtebuch, Bd. 4.2 Teilband Baden*. Stuttgart: Kohlhammer.
- Keyser, E. (1962). *Deutsches Städtebuch, Bd. 4.2 Teilband Württemberg*. Stuttgart: Kohlhammer.
- Keyser, E. (1964). *Deutsches Städtebuch, Bd. 4.3 Rheinland-Pfalz und Saarland*. Stuttgart: Kohlhammer.
- Keyser, E. and H. Stoob (1971). *Deutsches Städtebuch, Bd. 5.1 Bayern*. Stuttgart: Kohlhammer.
- Keyser, E. and H. Stoob (1974). *Deutsches Städtebuch, Bd. 5.2 Bayern*. Stuttgart: Kohlhammer.
- Köbler, G. (2007). *Historisches Lexikon der deutschen Länder: Die deutschen Territorien vom Mittelalter bis zur Gegenwart, 7. Auflage*. München: C. H. Beck.

- Minns, C., C. Crowston, R. de Kerf, B. de Munck, M. Hoogenboom, C. Kissane, M. Prak, and P. Wallis (2020). The extent of citizenship in pre-industrial England, Germany, and the Low Countries. *European Review of Economic History* 24(3), 601–625.
- Newald, J. (1883). *Das Österreichische Münzwesen Unter Ferdinand I.: Eine Münzgeschichtliche Studie*. Wien: Kubasta & Voigt.
- Pfister, U. (2017). The timing and pattern of real wage divergence in pre-industrial Europe: evidence from Germany, c. 1500-1850. *Economic History Review* 70(3), 701–729.
- Ramankutty, N., J. Foley, J. Norman, and K. McSweeney (2002). The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Global Ecology & Biogeography* 11, 377–392.
- Reinhard, W. (2004). Reichsreform und Reformation 1495-1555. In W. Reinhard (Ed.), *Gebhardt: Handbuch der deutschen Geschichte Band 9*. Stuttgart: Klett-Cotta.
- Roodman, D., J. MacKinnon, J. Nielsen, and M. Webb (2019). Fast and wild: bootstrap inference in Stata using boottest. *Stata Journal* 19(1), 4–60.
- 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.
- Schaff, F. (2020). When ‘The State Made War’, what Happened to Economic Inequality? Evidence from Preindustrial Germany (c.1400-1800). *LSE Economic History Working Paper* 311.
- Schaff, F. (2022). Exploring the Political Economy Causes of Inequality in Pre-Industrial Germany (c. 1350-1850). *Unpublished Ph.D. Dissertation, London School of Economics*.
- Schaff, F. (2023). Warfare and Economic Inequality: Evidence from Preindustrial Germany (c. 1400-1800). *Explorations in Economic History* 89(2), 101495.
- Schilling, H. (1994). *Aufbruch und Krise: Deutschland 1517-1648*. Berlin: Siedler.
- Scribner, B. (1994). Germany. In B. Scribner, R. Porter, and M. Teich (Eds.), *The Reformation in National Context*, pp. 4–29. Cambridge: Cambridge University Press.
- Sehling, E. (1902). *Die Ordnungen Luthers, die Ernestinischen und Albertinischen Gebiete*. Leipzig: Mohr Siebeck.
- van Bavel, B. and A. Rijpma (2016). How important were formalized charity and social spending before the rise of the welfare state? A long-run analysis of selected western European cases, 1400-1850. *Economic History Review* 69(1), 159–187.
- Voth, H. (2021). Persistence – Myth and Mystery. In A. Bisin and G. Federico (Eds.), *Handbook of Historical Economics*, pp. 243–267. New York: Academic Press.
- Wahl, F. (2019). Political participation and economic development. Evidence from the rise of participative political institutions in the late medieval German Lands. *European Review of Economic History* 23(2), 193–213.

- Warde, P. (2006). *Ecology, Economy and State Formation in Early Modern Germany*. Cambridge: Cambridge University Press.
- Weisenstein, K. (2002). Die Münzpolitik Der Rheinischen Kurfürsten (Rheinischer Münzverein) Unter Besonderer Berücksichtigung Der Einflüsse Von Reich Und Städten. In R. Cunz (Ed.), *Währungsunionen: Beiträge Zur Geschichte Überregionaler Münz- Und Geldpolitik*, pp. 105–143. Hamburg: Museum für Hamburgische Geschichte.
- Whaley, J. (2012). *Germany and the Holy Roman Empire Volume 1: From Maximilian I to the Peace of Westphalia 1493-1648*. Oxford: Oxford University Press.
- Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge MA: MIT Press.
- Zeumer, K. (1913). *Quellensammlung zur Geschichte der Deutschen Reichsverfassung in Mittelalter und Neuzeit*. Tübingen: Mohr Siebeck.

EHES Working Paper Series

Recent EHES Working Papers

2023

- EHES 238 The persistence of social inequality in adult mortality in rural Spain, death cohorts 1546-2010
Francisco J. Marco-Garcia, Víctor A. Luque de Haro
- EHES 237 Capital, Productivity, and Human Welfare since 1870
Daniel Gallardo-Albarrán
- EHES 236 Economic Freedom in Retrospect
Leandro Prados de la Escosura
- EHES 235 Subjective Well-being in Spain's Decline
Carlos Álvarez-Nogal, Leandro Prados de la Escosura
- EHES 234 Exploring the Drivers of Spain's Nutritional Transition: From Meat Shortages to Excess(1958-1990)
Pablo Delgado
- EHES 233 Risk Management in Traditional Agriculture: Intercropping in Italian Wine Production
Giovanni Federico, Pablo Martinelli Lasheras
- EHES 232 Spatial inequality in prices and wages: Town-level evidence from the First Globalisation
Stefan Nikolić
- EHES 231 Leviathan's Shadow: The Imperial Legacy of State Capacity and Economic Development in the Kingdom of Yugoslavia
Magnus Neubert

2022

- EHES 230 Demography and age heaping: solving Ireland's post-famine digit preference puzzle
Eoin McLaughlin

All papers may be downloaded free of charge from: <http://www.ehes.org/>

The European Historical Economics Society is concerned with advancing education in European economic history through study of European economies and economic history. The society is registered with the Charity Commissioners of England and Wales number: **1052680**