KU LEUVEN

DEPARTMENT OF ECONOMICS

Doomsday to Today. 1000 Years of Spatial Inequality in England

Luke Heath Milsom

FACULTY OF ECONOMICS AND BUSINESS

DISCUSSION PAPER SERIES DPS 24.02

January 2024

Doomsday to Today.
1,000 Years of Spatial Inequality in England.

Luke Heath Milsom*

8th January, 2024

Abstract

Using data from the Doomsday Book, I find that areas of England that were 10% richer in 1086 are on average almost 2% richer today. Using a natural experiment and a dynamic quantitative spatial economics model I show that this persistence is not due to path dependency. Instead, the 1086 economy was moving towards a different, but correlated, long-run spatial equilibrium than that observed in 2020. This correlation in spatial equilibria can in part be explained by local market access. Modern place-based policies aiming to shift the spatial distribution of economic activity should focus on changing location fundamentals if they are to have long-run impacts.

*The Institute for Fiscal Studies, luke_m@ifs.org.uk

Today England is arguably one of the most regionally unequal industrialized economies [McCann, 2020], with the highest coefficient of variation in regional (NUTS1) gross value added in Western Europe [Stansbury et al., 2023] and a Northeast-Southeast healthy life gap of over 6 years [Bambra, 2016]. These differences persist in the face of considerable investment in relatively deprived areas culminating in the recent "levelling up" initiative. In this paper, I study the potential long run origins of this inequality, and by analysing its roots aim to shed some light on what types of policy could be most effective in equalising regional differences.

First, I show that areas that are richer today were also richer 1,000 years ago by leveraging unique data from the Doomsday Book. In 1086, after his successful invasion of England, William the Conqueror commissioned a survey of the value of his new English possessions so that he might better tax them. The resulting document, the Doomsday Book, (mainly) survives and gives a unique insight into the population and income of England in 1086 at a granular geography. Using this data I find that areas that were 10% richer in 1086 remain on average almost 2% richer today. This relationship is robust to a barrage of corrections for the possibility of spatial correlation in both series causing a spurious relationship. In particular, I fit a spatial autoregressive model, adjust for Conley [Conley, 1999] standard errors, use the approach suggested by Müller and Watson [2022], and finally perform the semi-parametric thin plate spline correction due to Kelly et al. [2023].

There are two competing explanations as to why areas that are richer today could also be richer in 1086 Lin and Rauch [2022]. It may be *path-dependence* "we are richer today because we were richer yesterday" or on the other hand, it could be due to location *fundamentals* "we are richer today because of local characteristic X, and this was also why we were rich yesterday". Distinguishing between these two forces helps us understand how the current spatial distribution of economic activity came to be, and the potential welfare gains to an alternative distribution. In addition, each potential mechanism implies that a different set of policies would be effective in tackling spatial inequality. Path dependency implies that big-push one-shot interventions that shock a location into increasing its population would be effective, such as local (individual or business) tax breaks as is common in "opportunity zones" or similar place-based policies. On the other hand, a prominent role for local fundamentals implies policies should aim to increase the innate attractiveness of a location for consumers or businesses, perhaps through building a large park or a new rail connection.

First I trace the impacts of a large-scale spatial shock over time. If the impact of such a shock can not be felt today, it provides evidence against path dependence and for local fundamentals, as the shocked area eventually returns to its position in the distribution. This is a common strategy employed in the literature, for example, see Davis and Weinstein [2002]. I leverage the large-scale shock of the "Harrying of the North". In the early years of the conquest, many areas of England rebelled against William's rule. However, only one of these rebellions was brutally put down and incensed William to exact reprisals on the general population. This rebellion was one in the north of England in 1069-70. The reaction from Willam was catastrophic, some commentators estimate that as much as three-quarters of the population of Yorkshire was killed or fled as a result. Comparing regions affected by the Harrying to other rebellious locations that didn't experience such reprisals, I find that the impact can still very much be felt some 16 years later using data from the Doomsday Book. However, the impact can not be detected using more modern data, implying that the Harrying did not result in a long-run change. This gives empirical evidence against path dependence and for fundamentals in this context. However, this relates only to a specific example and relies on the quasi-randomness of William's decision "Harry" only one rebellious geography. To provide more general evidence I next turn to conditioning directly on local fundamentals leveraging a dynamic quantitative spatial economics model.

One direct way to rule in, or out, the role of fundamentals in explaining the observed correlation, would be to directly control for them. However, the spatial distribution of local fundamentals is not observed. Additionally, controlling for some historical characteristics and observing whether the relationship attenuates does not in general allow one to separately identify path dependence from fundamentals. Controlling for a characteristic may attenuate the relationship because in both 1086 and today that characteristic was/is predictive of local incomes. However, it could also be that in 1086 the characteristic was important, and this led to the formation of agglomerations and infrastructure which perpetuates today. This fundamental identification issue will plague the use of any observed historical variable.

To overcome this, I add structure and estimate a dynamic quantitative spatial economics model on data from 1086 and 2020 following Allen and Donaldson [2020]. This framework endogenizes the spatial distribution of population and economic activity and allows the past to affect today through productivity and amenity spillovers. It also allows for path-dependence as multiple long-run spatial equilibria exist — a large shock to the current distribution of economic activity could cause the economy to switch equilibrium. Data from the Doomsday Book is sufficiently rich to allow me to invert the model and back out location-specific fundamental productivities and amenities in 1086, and modern data allows a similar procedure for 2020. These fundamentals are structural residuals where the effect of contemporaneous and historical trade and migration links as well as agglomeration spillovers have been conditioned upon. A location in 2020 with high fundamental amenities would have a population higher than that which can be explained by contemporaneous trade/migration links, local wages, contemporaneous agglomeration benefits (or negatives), or, crucially, historical agglomeration benefits. The model therefore overcomes the identification problem and provides a set of local fundamentals purged of the possible effects of path dependency. Of course, there may be other angles not captured by the model, through which path dependency operates such as cultural or political channels.

First, I find that location-specific fundamental productivities and amenities in 1086 and 2020 are correlated. This implies that, given the same initial distribution, the two economies could converge to correlated spatial equilibria. Second, I regress income in 2020 against that in 1086 conditioning on local fundamentals and find that this conditioning completely kills the previous persistence. That is, the observed long-run persistence between 1086 and 2020 operates through the correlation in fundamentals over this time. This gives strong evidence for the fundamentals explanation for the underlying correlation: Areas that are richer today are richer due to some characteristics X, and these characteristics (or some correlated set) are also why areas were richer in 1086.

I next build a dataset of location-specific time-invariant characteristics to identify which features of a location drive the observed correlation in fundamentals and thus persistence in outcomes over 1,000 years. I find that ruggedness, availability of running water for mills, pre-existing fisheries, Roman roads, Roman towns, proximity to the coast, soil fertility, and proximity to London all do not explain the correlation between fundamentals. Instead, I find that variables capturing local (time-invariant) market access go some way to explaining the correlation in fundamentals and therefore the long-run persistence uncovered. Areas with better market access were richer in 1086 and are also richer today.

England is today a highly unequal economy, this paper shows that part of that inequality was determined some one thousand years ago. I also show that correlation in local fundamentals, rather than economic path dependency explains this long-run persistence. Recently (and historically) politicians have championed spatially redistributive policies such as the Northern Powerhouse and the Leveling up initiative. This work shows that such initiatives will not rebalance the long-run distribution of economic activity unless they alter local fundamentals — and it is here that governments aiming for spatial equality should focus.

This paper contributes to the developed and growing literature on the determinants of the current spatial distribution of economic activity, and in particular on the path dependence vs fundamentals debate. For example, using shocks to city population Davis and Weinstein [2002], Miguel and Roland [2011], Jedwab et al. [2019] find evidence of quick rebounding suggesting a strong role for fundamentals. Turning to path-dependence, Bleakley and Lin [2012] and Jedwab et al. [2017] find evidence of persistence in the face of infrastructure obsoletism, which is evidence for path-dependence. However, Michaels and Rauch [2018], Redding and Sturm [2008], Gibbons et al. [2018] and Dell [2010] find evidence that large shocks can change long-run outcomes, providing evidence against path dependence. In this paper, I disentangle these two possibilities in my setting by combining evidence from a historical natural experiment with that from a dynamic quantitative spatial economics model. Lin and Rauch [2022] provide a modern overview of the literature on how looking to history can inform this discussion. I build on this literature by combining evidence from a natural experiment with that from a dynamic quantitative spatial economics model. I also speak to the literature on spatial inequality in general [Gaubert et al., 2021, Stansbury et al., 2023, Milsom, 2023], and in England specifically [Bell et al., 2023, Britton et al., 2021, Commission, 2019], highlighting that to achieve long-run spatial redistribution policymakers need to target underlying location fundamentals.

The rest of this paper proceeds as follows. Section 1 describes the Doomsday Book and gives historical context. Section 2 details the main long run spatial persistence result. Section 3 then analysis the mechanisms for this persistence. Finally section 4 concludes.

1 The Doomsday book

In 1066 William the Conqueror successfully invaded England from Normandy defeating the Anglo-Saxon forces of Harold Godwinson at the Battle of Hastings and therefore settling the succession dispute that followed the death of Edward the Confessor. Some twenty years after his victory, and unsatisfied with the relatively low tax revenue from his new land, William commissioned a census of the value of everything in England.

The resulting book came to be known in subsequent centuries as the Domesday[1] Book, reflecting its comprehensive nature. It represents a fantastically detailed and complete account of the economy of Medieval England. Only relatively recently with the painstaking work of the Hull digitization project [Palmer, 2016] has the Doomsday Book been digitized in a way suitable for quantitative analysis. For a more detailed discussion of the Doomsday

[1]In this paper I will from now on use the modern spelling Doomsday.

Book for economic research see Walker [2015] and Delabastita and Maes [2023]. Data is arranged at the level of the manor and covers ownership, population by type, plows, mills, land area, churches, and value. Manors were the level at which the agricultural feudal economy was organized and typically corresponded to a small settlement or part of a larger village.

The act of collecting data for the Doomsday Book was itself a remarkable act of central administration, unheard of in Medieval Europe. William sent commissioners to each corner of his new kingdom and interrogated local rulers on their holdings. Boards of English and Norman jurors were then tasked with verifying the validity of answers given by local lords. Although it was potentially collected, no data for major cities survives and therefore some large agglomerations in 1086 are not included in this analysis. In 1086 however, it is estimated only eight agglomerations exceeded 2,000 in population [Russell, 1948][2]. The Doomsday Book suggests a population of England of around 1.5 million implying that roughly 3% of the population is missing due to the omission of some major cities.

Crucial to my analysis is the manor-level recorded "valets" i.e. values. Some debate among historians remains as to whether this variable can be interpreted as a measure of manor-level income or only as a measure that captures money rents. As in Delabastita and Maes [2023], Walker [2015], and McDonald and Snooks [1985] as well as most modern historians, in this paper I take the former view. Evidence in favor of this interpretation comes first from Galbraith [1929] who shows how a manor was rented out at the rate exactly that of its value. This is also the conclusion reached by Britnell and Campbell [1995] and Roffe [2015] the latter of whom states, as noted by Delabastita and Maes [2023] on page 241 that: "Doomsday values are a more or less accurate index of the productive capacity of estates". For this reason, I shall refer to value per capita as a measure of local income and discuss inequality in terms of income throughout this paper. I will also take as holistic a view of "per capita" as possible, and include all souls recorded as living in a locality in my denominator. This includes those denoted as slaves but is unlikely to include many women [Stafford, 1989]. It should also be noted that in this feudal economy, within-location inequality is likely to be extremely high with local rulers enjoying most of any surplus generated.

Manors can be geolocated and thanks to the work of the Hull digitization project we know the location of manors across England and therefore could analyze spatial inequality in 1086 at this highly granular level. However, modern data doesn't live up to this degree of granularity and spatial precision. Instead, I aggregate Norman variables to the lowest

[2]In descending order: London (17,850-10,000), Winchester (6,000-6,750), Norwich (4,444-4,750), York (4,134-5,000), Lincoln (3,560-4,500), Thetford (2,681-4,000), Bristol (2,310), Gloucester (2,146-2,750). Estimates from Russell [1948] and Darby and Darby [1986]

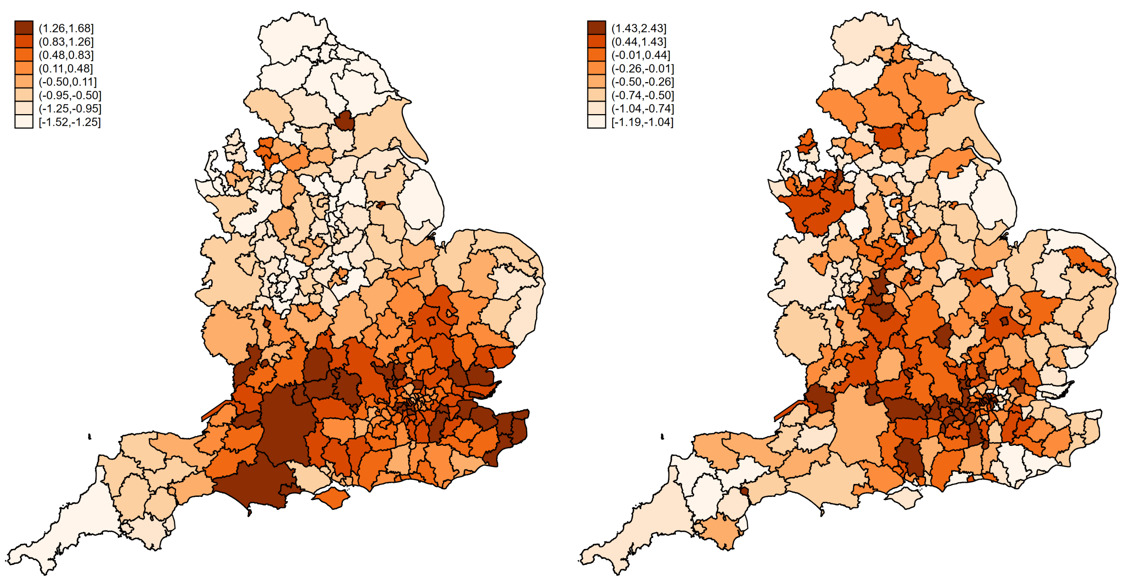
level at which modern data is representative, that of modern local authorities. Throughout this paper, my main level of analysis will therefore be these modern local authorities. As mentioned above the Doomsday Book for some major agglomerations doesn't survive, in addition, some areas in the North West of England are unaccounted for. In total, I have data for 283 modern local authorities across England. I winsorize incomes in 1086 and 2020 at the 5 and 95 percentile levels to remove outliers, but the results are unaffected by this.

2 Persistence in income inequality from Doomsday to Today

Figure 1 shows the spatial distribution of incomes in 1086 and 2020 measured in period-specific standard deviations, over modern local authorities. It is immediately clear from these figures that the two series show some correlation over space. In 1086 the dominance of the south was perhaps even more stark than it is today, and in both periods the relative poverty of the extremes of the country is evident.

Figure 1 Spatial distribution of incomes

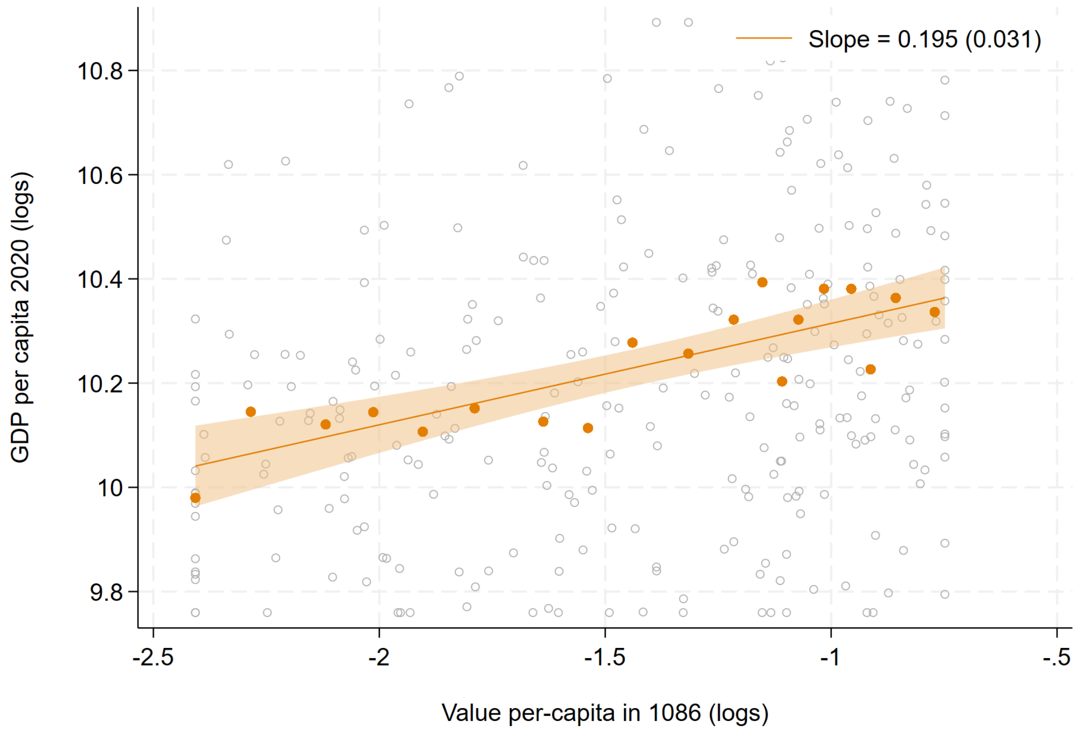
(a) 1086 (b) 2020



Notes: This figure shows the spatial distribution of (standardised) value per capita in panel (a) and (standardised) GDP per capita in panel (b). In both cases, the distribution is displayed over modern local authorities. Both series have been winsorised at the 5 and 95 percentile level. Both figures show the local authorities that appear in my main sample, i.e. those in England for which data from the Doomsday Book is available.

Figure 2 presents the correlation between log value per capita in 1086 and log GDP per capita in 2020. Figure 2 shows a clear positive correlation between 1086 and 2020. Areas that were 10% richer in 1086 remain on average almost 2% richer in 2020. The relationship appears fairly linear and tight, with the 95% confidence interval indicating an elasticity between 0.135 and 0.257.

Figure 2 Spatial persistence over the very long run



Notes: This figure plots log value per capita in 1086 against log GDP per capita in 2020 in a binscatter figure with the raw scatter plot displayed underneath. A linear line of best fit is shown in orange the slope and robust standard errors of which are given. The geographical unit is modern local authorities. Standard errors are robust.

Table 1 shows that this long-run relationship is robust to the exact specification used. In column one, I replicate the baseline results as shown in figure 2. In column two I weigh by 1086 population and recover a slightly higher estimate, not shown is that weighting by the 2020 population also recovers a similar estimate. In column three I include a second-order polynomial in local authority centroid latitude and longitude coordinates. In a crude manner, this specification controls somewhat for spatial dependencies anticipating the discussion on spatial autocorrelation picked up in section 2.1. Anticipating the results from section 2.1, in column three the correlation is somewhat attenuated, but remains significant both statistically and economically. In column four I include fixed effects for the nine English regions. Controlling for such cross-regional variation only attenuates the estimated coefficient by 27%, indicating that the broad correlation is not driven by large-scale differences such as north vs south, but is also present in within-regional variation. Finally, column five presents a rank-rank regression and recovers an even stronger correlation in ranks.

Table 1 Spatial persistence over the very long run

	Log GDP per-capita 2020	Log GDP per-capita 2020	Log GDP per-capita 2020	Log GDP per-capita 2020	LA rank 2020
Log values per capita 1086	0.195***	0.231***	0.143***	0.136***	
	(0.0311)	(0.0356)	(0.0450)	(0.0468)	
Rank values per capita 1086					0.300***
					(0.0523)
Weighting	None	1086 Pop	None	None	None
Lat-lng polynomial			Yes		
Region FE				Yes	
Observations	283	283	283	283	283
R2	0.0971	0.135	0.194	0.131	0.0900

Notes: This table shows the results from regressing the spatial distribution of income in 2020 against that in 1086 across various specifications. In the first column, I show the raw results in a log-log specification that corresponds to figure 2. Column two weights the regression by the 1086 population. Column three includes a second-order interacted polynomial in local-authority centroids. Column four includes fixed effects for the 9 high-level regions of England. Finally, column five performs a rank-rank regression.

Table 7 in the appendix considers the robustness of the relationship to further specifications. It shows that the result of significant long-run persistence is robust to: not winsorising, excluding Harried local authorities, excluding modern London, using average or median wages from the 2020 ASHE data, or performing the analysis in levels (rather than logs).

2.1 Spatial correlation

Correlating two spatial variables can lead to a spurious relationship due to spatial autocorrelation, in much the same way that correlating two time series can also lead to nonsensical relationships [Kelly, 2019].

Spatial data is naturally ordered (A is close to B) in a similar way to that which time series data is naturally ordered (A is soon after B). However spatial data operates in a two (or three, but I will not explicitly consider elevation) dimensional space, whereas time-series data operates over only one dimension. Spatial units also tend to not be uniformly spread over the topology they operate within and are measured on a continuous scale, whereas time series is often uniform and measured on a discrete scale. These complications mean that to the authors' reading, the literature has not settled on one single approach to accounting for the possibility of spatial correlation. Instead here I present four leading approaches each of which either accounts for spatial dependency directly or adjusts standard errors ex-post to account for it.

Table 2 shows the results from each approach in the upper panel and the baseline results in the lower panel. Coefficient estimates are indicated above standard errors which are given in brackets. Stars appending to standard errors indicate the usual significant levels. In column one, I present results using a spatial autoregressive model. This approach is analogous to using an autoregressive model to account for autocorrelation in time series data. When using an AR model in time series data, one must specify the number of lags to be used. Similarly, when using a spatial autoregressive model one must specify the spatial weighting or decay matrix. One can make various choices here, but for the results presented in table 2 I've used the simplest approach which is a linear inverse-distance weighted spatial decay matrix. By controlling for spatial correlation, much like when using an AR model on time series data, the coefficient of interest changes as well as its standard errors. The estimated elasticity is attenuated which suggests that spatial correlation is causing some of the initially estimated relationship. However, the coefficient remains statistically and economically significant.

The second approach I adopt is that of Conley [1999], the results of which are presented in columns two, three, and four of table 2. Conley standard errors use a kernel to give more weight to observations close together when calculating standard errors, in a manner analogous to how White heteroskedasticity robust standard errors are calculated. Following the standard specification, I use a uniform kernel and vary the cutoff. In column two I use a 10km cutoff, in column 3 a 50km cutoff and in column 4 a 100km cutoff, in appendix figure 9 I show the estimated t-statistic on the coefficient of interest at cutoff intervals between 10 and 500km. As the cutoff is increased the standard errors also increase, although not sufficiently to threaten statistical significance at conventional levels. Note that Conley standard errors are an ex-post correction (much like robust standard errors) so this procedure does not alter the estimated coefficient.

Thirdly, I control for spatial correlation by performing the procedure suggested in Müller and Watson [2022]. This paper proposes a procedure for constructing confidence intervals that account for many forms of spatial correlation by adjusting both the standard errors and critical values. As this procedure does not affect the point estimate, I only report the estimated confidence interval in table 2. The standard error in the Müller and Watson [2022] approach is calculated as the principle component from a given worst-case spatial correlation model, the critical value is then chosen to ensure correct coverage in some benchmark parametric model. This approach has the added advantage over Conley standard errors or Spatial autoregressive models, as it allows for the unequal placement of geographic units. In table 2 I display the estimated 95% confidence intervals for three such worst-case spatial correlation matrices. Although the confidence intervals sometimes include 0, in all cases the vast majority of the mass is in the positive range.

In the final column, I correct for spatial autocorrelation using the semi-parametric thin plate spline methodology proposed in Kelly et al. [2023]. This process fits a two-dimensional non-parametric spline in latitude and longitude to directly and flexibly control for spatial dependencies. As discussed in Kelly et al. [2023] this approach allows me to separate out the spatial structure of the regression as a nuisance variable and carry out standard inference on the remaining parameters. This procedure attempts to directly control for the spatial dependencies and so will alter the coefficient estimates and standard errors. This can be seen in column 8 of table 2 where the coefficient estimate is smaller, and the standard errors larger. Despite these corrections, the resulting point estimate remains economically significant and statistically significant at the 10% level (pvalue 0.079).

Table 2 Accounting for spatial correlation

	Spatial Autoregressive Model	Conley Standard Errors			Müller and Watson (2022)			Kelly, Mokyr, and, Ó Gráda (2023)
		d=10km	d=50km	d=100km	sc=0.03	sc=0.02	sc=0.01	
GDP per-capita 1086 (SD)	0.0902 (0.0448)**	0.1945 (0.0331)***	0.1945 (0.0520)***	0.1945 (0.0637)**	[-0.0579, 0.4470]	[-0.0081,0.3972]	[0.0467,0.3424]	0.0934 (0.0530)*
Observations	283	283	283	283	283	283	283	283
R2	0.1817	0.0971	0.0971	0.0971	0.0971	0.0971	0.0971	0.307

Results from the baseline model	
Coefficient	0.1945
Standard Error	0.0311***
Observations	283
R2	0.0971

Notes: This table the results of performing various corrections for the presence of spatial correlation. In the lower panel, I re-print the results from performing the baseline regression as shown in column one of table 1. In the upper panel, I perform 8 additional regressions adjusting for spatial correlation via four different approaches. In the first column, I fit a spatial autoregressive model with inverse-distance weighted spatial lags. In columns two, three, and four I calculate Conley standard errors using a 10km, 50km, and 100km cutoff. In columns five, six, seven, and eight I perform the spatial correlation robust inference procedure due to Müller and Watson [2022]. I specify a worst-case spatial correlation matrix with correlation values 0.03, 0.02, and 0.01 respectively. Finally, in column eight I perform the semi-parametric thin plate spline correction due to Kelly et al. [2023].

3 What can explain the long-run persistence in income inequality

Section 2 shows that areas that were richer 1,000 years ago are on average still richer today, here I turn to ask why, and in particular distinguish between two mechanisms highlighted in the literature.

1. **Path dependency**. "We are rich today because we were rich yesterday".

2. **Fundamentals**. "We are rich today because of local characteristics X, this was also why we were rich yesterday".

Disentangling these two potential mechanisms is crucial both for our theoretical understanding of the spatial distribution of economic activity, but also from a policymakers perspective. From the theoretical side, if path dependency is shown to be important vis-a-vis fundamentals, it implies that the most productive locations may have been passed by, due to some fluke of history causing self-reinforcing agglomerations elsewhere. Therefore, there could be large productivity gains to redistributing economic activity. From the perspective of a policymaker aiming for spatial equality (or less inequality), path dependency implies the effectiveness of big-push one-shot policies. However, a large role for fundamentals implies that the underlying characteristics of a location that make it attractive to individuals or firms need to be addressed, and big-push policies that fail to do this will not be effective in the long run.

Distinguishing between these channels empirically can be thought of as an issue of identification. The presence of a long-run correlation obviously does not by itself allow one to distinguish between these two possibilities. However, the challenge is greater. At first glance, a promising strategy may be to pick some historical local characteristic X, control for it in the persistence regression, and note whether the coefficient on historical values is attenuated i.e. if X mediates the relationship. To see this suppose that we find some such X and conclude that fundamentals drive the long-run persistence through X. However, we cannot rule out that X was historically important, and that this led to the development of cities and infrastructure which then led to prosperity today, although X itself is no longer directly important. The identification challenge persists. To circumvent this issue, the literature has relied on transient exogenous local shocks (e.g. Davis and Weinstein [2002], Miguel and Roland [2011], Jedwab et al. [2019]). As these do not cause any permanent changes in local fundamentals, permanent changes in local economic outcomes are taken as evidence for path dependency whereas reversion is evidence for the role of such local fundamentals.

In this section, I combine two approaches to provide evidence on which of the two channels is driving the observed long-run persistence. First, I follow the literature and leverage a natural experiment that decimated the population in part of the country and show that this transient quasi-exogenous local shock did not lead to a permanent change in local economic outcomes. Second, I provide more general and direct evidence, by turning to theory and imposing structure on the data using a dynamic quantitative spatial economics model [Allen and Donaldson, 2020] which allows me to directly identify local fundamentals and circumvent the identification problem posited above. By combining evidence from these two sources, I argue that in the case of 1,000 years of spatial income persistence in England, the role of fundamentals trumps that of path dependence.

3.1 The Harrying of the North as a "natural experiment"

A typical approach in the literature to distinguishing between path dependence and fundamentals is to track the relative performance of local outcomes after a local shock. Intuitively, if after a shock cities recover back to their pre-shock trend this is evidence of fundamentals playing a key role. On the other hand, if after a shock, the shocked area continues above or below trend, this is taken as evidence of path dependence. For example, Davis and Weinstein [2002] finds that Hiroshima and Nagasaki returned to their pre-war position in the urban hierarchy soon after the horrifying and devastating effects of nuclear warfare. Here, I will take a similar approach, using a historical "natural experiment" and tracing the impact on affected local authorities over time.

The natural experiment I employ is the "Harrying of the North" — the brutal reprisals visited on the North in response to a rebellion against William's rule in 1069/70 (see for example Strickland [1998], or Dalton [2002] for more details on the Harrying or Vitalis [1854] for an almost first-hand account). In 1069 the last Wessex claimant Edgar Ætheling incited a rebellion in the North of England, centered in York. William raised an army and marched on York but the rebels refused to meet him in open battle. William then decided to punish the northern shires and prevent future rebellions by using scorched earth tactics. Contemporary accounts, archeological evidence, and evidence from the Doomsday Book attest to the scale of the resulting destruction. By some accounts around 3/4 of the population of Yorkshire was killed or fled. For example, the Anglo-Norman chronicler Orderic Vitalis wrote:

> The King stopped at nothing to hunt his enemies. He cut down many people and destroyed homes and land. Nowhere else had he shown such cruelty. This

made a real change. To his shame, William made no effort to control his fury, punishing the innocent with the guilty. He ordered that crops and herds, tools, and food be burned to ashes. More than 100,000 people perished of starvation.

It should be noted that not all historians are in agreement as to the scale of destruction. Dalton [2002] suggests that given limited time and troops destruction on the scale suggested by Vitalis would not have been possible. Similarly, Hagger [2021] and Horspool [2009] suggest that the scale of destruction was not so abnormal relative to other contemporary conflicts. However, even if exact numbers are disputed, all authors maintain that the scale was such as to shock the North and create scars for generations and that the Harrying was certainly qualitatively and quantitively different from how William reacted to any other rebellions.

Of course, one may be concerned that the North's rebellion was not a random event. Perhaps this area chose to rebel because it was more negatively affected by the conquest or was more remote relative to the centers of Norman power and so considered a rebellion to be more likely to be successful. To combat these concerns I leverage a control group of local authorities where other rebellions against William's rule occurred. Within a similar time scale, Kent, Northumbria, the Welsh borders, and East Anglia all rebelled against Williams's rule. See appendix section A.2.1 for details on these other rebellions.[3] Figure 10 in the appendix shows on a map of England the Harried and other rebellious control modern local authorities.

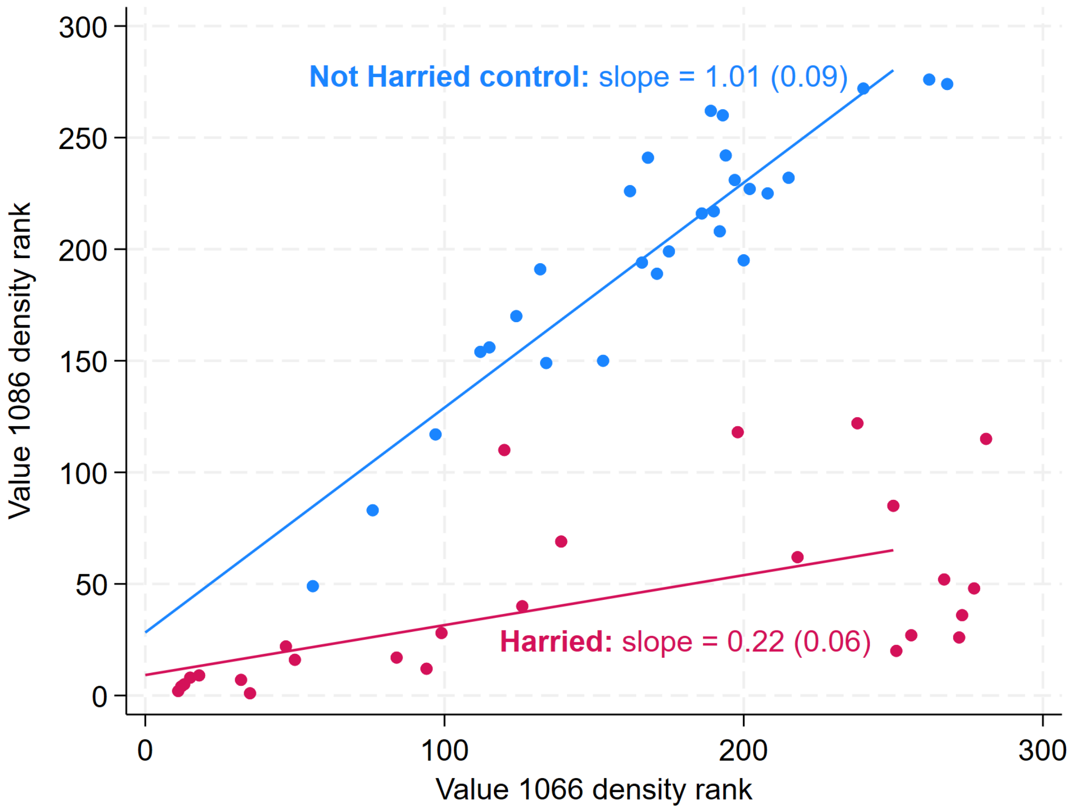
To investigate the initial impact of the Harrying I use data available in the Doomsday Book on local values in 1066. These values were recorded via recall in the census of 1086. However, population estimates are not available in 1066 so instead I consider a measure of value per area. This measures the value created in 1066 or 1086 in each modern local authority by square kilometer. In figure 3 I then plot each local authority's 1066 value density rank (inverse rank — higher is better) against its 1086 value density rank, where ranks run over all 264 local authorities. In red I plot Harried locations, and in blue I plot not Harried but otherwise rebellious control locations. Focusing first on the non-Harried locations we see a fairly precisely estimated coefficient of unity, implying that otherwise rebellious, but not harried, locations maintained their position in the overall value density distribution.[4] For Harried locations the picture could not look more different, these locations are positioned much lower in the 1086 value-density distribution than in the 1066 distribution. Harried

[3]I don't include the rebellion in Exeter as it was confined to the city, however, results are robust to its inclusion.

[4]Note that as Harried locations fell down the rank distribution one would mechanically expect other locations to move up the distribution. The fact that the control locations didn't move up the distribution implies that there may have been some, relatively small, penalty associated with rebelling.

locations that had among the highest value density in 1066, ranked around 250, were among the lowest in 1086, ranked around 50.

Figure 3 Short(er) run impact of the Harrying of the North



Notes: This figure shows the impact of the Harrying of the north on modern local authorities that were "harried" in 1086 some 16 or so years after the event. The figure plots, in red, Harried local authorities and in blue non-harried control local authorities that also rebelled against William's rule. On the x-axis is the 1066 value rank (scaled by the size of the local authority), and on the y-axis the same variable for 1086. The slope and associated robust standard errors are given in the figure for each line.

Figure 3 shows a large impact of the Harrying of the north some 16 years or so after it occurred. In table 3 I formalise the graphical results shown in figure 3 by performing the following regression on a sample of Harried treatment locations and rebellious but not Harried control locations

$$\text{Value by area}_{it} = \alpha_t + \beta_t \cdot \text{Harried}_i + \varepsilon_{it} \tag{1}$$

and reporting the β_t coefficients. Table 3 shows the results from estimating this one regression. In column one, I consider the impact in 1066, before the Harrying occurred, and find no effect. This can be considered as a placebo exercise, or as a rather crude test for parallel pre-trends in a DiD analysis. In column two I then confirm the large negative impact of the Harrying on the harried locations shown in figure 3. Finally, in column three I consider whether the impact of the Harrying can still be felt in 2020. I find no measurable impact of the Harrying today. This provides some suggestive evidence that Harried locations (eventually) recovered and returned to their position in the spatial distribution. To get a better handle on this I am currently collecting data covering the intervening periods.

Table 3 The Impact of the Harrying of the North over time

	Value 1066 by area (SD)	Value 1086 by area (SD)	GDP 2020 by area (SD)
Harried	0.107	-1.537***	0.00215
	(0.300)	(0.172)	(0.0238)
Constant	0.123	0.576***	-0.241***
	(0.115)	(0.164)	(0.0207)
Observations	54	54	54
R2	0.00255	0.590	0.000151

Notes: This table shows the estimated impact of the Harrying of the North on the Harried areas relative to other rebellious areas. The geographic unit of analysis is modern local authorities. Standard errors are robust. In column one, I show results for 1066 value by area. In Column two for 1086 value by area, and in column three for 2020 GDP by area. Observations and R-squared are given below. There are 25 treated (Harried) local authorities and 29 control (rebellious but not Harried) local authorities.

We can also consider results from the more demanding two-way fixed effects specification, controlling for local authority-level time-invariant characteristics. Table 8 in the appendix reports the results of running a 2WFE specification. Relative to 1066 this specification implies that Harried areas in 1086 are 1.78 (0.289) standard deviations poorer (in terms of value per area) and Harried areas in 2020 are not statistically significantly different (coefficient of -0.25 (0.33)). This difference in differences analysis relies on the usual assumption that, in the absence of treatment (the Harrying) treated locations (the North) would have on average evolved in a manner similar to that of the control locations. Over 1,000 years this is obviously a strong assumption, although similar such assumptions (or RD equivalents) are made in various leading papers in this literature see for example: Dell [2010], and Nunn and Qian [2011]. These papers, and others in this literature provide as evidence for the parallel trends assumption the randomness of treatment. This is the same argument I take, arguing that William could have "Harried" any of the areas that revolted, but only did so in the North for reasons orthogonal to the North's counterfactual outcome. The DiD analysis also requires the assumption of no contemporaneous shocks, although this was certainly a tumultuous period, no other event in the North stands out in the historical record as being on anywhere near the same scale as Williams Harrying.

3.2 Controlling directly for local fundamentals using a dynamic quantitative spatial economics model

The natural experiment afforded by the Harrying of the North gives some evidence against long-run path dependency in local value per capita across England. However, the evidence it provides is specific to the variation it uses i.e. from one specific case study. Although that might be indicative of more general forces, I now turn to providing direct evidence for this.

A good strategy for distinguishing between two potential causes of some phenomenon would be to control for one potential cause and check whether the phenomenon is still observed. Therefore, one approach could be to control directly for local fundamentals. If, once local fundamentals are conditioned upon, I find no correlation between values today and in 1086 this would imply that fundamentals rather than path-dependence are driving the persistence. Intuitively fundamentals can then be said to explain or mediate the relationship. On the other hand, if the estimated persistence remains even conditional on local fundamentals this gives evidence that fundamentals are not driving the observed relationship and instead, it may be due to path dependence. As described above, we need to condition on local fundamentals and not just on some predetermined observed characteristic as this does not allow one to separately identify the two potential mechanisms.

This strategy, although on the face of it promising, suffers from a potentially fatal downside: local fundamentals are not observed. However, such fundamentals can be estimated by adding structure and employing a dynamic quantitative spatial economics model. This framework allows me to condition on the observed spatial distribution of population and economic activity, as well as the interconnectedness and interdependencies between locations. The model allows for costly migration and trade as well as endogenous contemporaneous and historical agglomeration forces in local amenities and productivities. The rich data available in the Doomsday Book allows me to invert the model and back out local fundamental productivities and amenities that rationalize the observed distribution of economic activity and population both in 1086 and in 2020. By using a dynamic model I allow for the possibility that path dependence explains the observed distribution. Intuitively, a location in 2020 with high fundamental amenities would have a population higher than that which can be explained by contemporaneous trade/migration links, local wages, contemporaneous agglomeration benefits (or negatives), or crucially historical agglomeration benefits. These fundamentals therefore do not reflect the potential impacts of path dependency.

This model, although quite general (see Allen et al. [2020b], and Milsom [2023] for a discussion of generality in this type of model), does introduce considerable structure. The fundamentals uncovered are effectively structural residuals that allow us to rationalize the observed distribution of economic activity conditional on the model structure. The model necessarily cannot capture all potential sources of path dependency, and in particular cultural or political mechanisms are not accounted for, neither is the possibility of endogenous growth or structural transformation.

3.2.1 A dynamic quantitative spatial economics model

Here I follow Allen and Donaldson [2020] closely, and present a slightly simplified version of their model. The treatment of Allen and Donaldson [2020] presented here also closely follows that presented in Heath Milsom [2024]. Within this framework, cities are connected and goods and individuals are allowed to move (with some cost) between them. History impacts the future through dynamic agglomeration effects in productivity and amenities. Intuitively infrastructure built some time ago might enhance (or decrease) productivity today. The model also admits the potential for multiple long-run spatial equilibria whereby shocks can cause permanent changes to the distribution of economic activity.

There are arbitrarily many locations $i \in N$ and $t \in \mathcal{T}$ time periods. Each location i emits a unique good in an Armington fashion. A continuum of firms ω in i produce this homogeneous good ($q_{it}(\omega)$) under perfect competition and CRTS using labor ($l_i(\omega)$) as the only factor of production.

$$q_{it}(\omega) = A_{it} l_{it}(\omega), \qquad A_{it} = \bar{A}_{it} L_{it}^{\alpha_1} L_{it-1}^{\alpha_2} \tag{2}$$

Where $\bar{A}_{it}$ is exogenous productivity and L_{it} is the total number of workers. α_1 captures aggregate contemporaneous spillovers, α_2 captures aggregate historical productivity spillovers. Intuitively α_1 captures what is more traditionally thought of as agglomeration forces, whereas α_2 captures factors like historical infrastructure which remain productive in the next period.

Individuals have CES preferences over differentiated location-specific goods with the elasticity of substitution σ, therefore consumption welfare is captured by local real wages (w_{it}/P_{it}). A location also generates utility for individuals in the form of local amenities (u_{it}), and therefore location-time specific welfare is given by W_{it} in equation 3.

$$W_{it} = u_{it} \frac{w_{it}}{P_{it}}, \qquad u_{it} = \bar{u}_{it} L_{it}^{\beta_1} L_{it-1}^{\beta_2} \tag{3}$$

Where $\bar{u}_{it}$ is exogenous productivity and β_1, β_2 are analogous to α_1, α_2. β_1 captures contemporaneous congestion forces i.e. from non-tradeables or land and β_2 captures the impact of durable infrastructure on amenities, such as parks.

Bilateral trade from locations i to j incurs exogenous, symmetric, iceberg trade costs denoted by τ_{ijt}. Iceberg trade costs and CES demand generate the familiar gravity equation in trade [Allen et al., 2020b].

$$X_{ijt} = \tau_{ijt}^{1-\sigma} \left(\frac{w_{it}}{P_{it}} \right)^{1-\sigma} P_{jt}^{\sigma-1} w_{jt} L_{jt}, \qquad P_{it} = \left(\sum_{k=1}^{N} \tau_{kit} \left(\frac{w_{kt}}{A_{kt}} \right)^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \tag{4}$$

Individuals decide where to move to maximize utility given in 3 subject to iceberg moving costs μ_{ijt} and some idiosyncratic preference draw ε_{jt} which is drawn from a Frechet distribution with dispersion parameter θ. Given this distributional assumption, migration will also follow a gravity structure, and the number of people moving from i to j in period t will be given by equation 5.

$$L_{ijt} = \mu_{ijt}^{-\theta} \Pi_{it}^{-\theta} W_{it}^{\theta} L_{it-1} \tag{5}$$

Where $\Pi_{it} = \left(\sum_k \mu_{ikt}^{-\theta} W_{kt}^{\theta} \right)^{1/\theta}$, is a measure of labor market access.

The dynamic equilibrium of this model is described in the appendix section B. The model can be solved via a simple iterative algorithm [Donaldson and Hornbeck, 2016]. For details and an in-depth discussion of the equilibrium properties of this model, and models in this class, see Allen and Donaldson [2020] and Allen et al. [2020a].

Persistence.

The dynamic quantitative spatial economics model described above has the attractive property that, depending on the parameter values, it can exhibit within-period equilibrium uniqueness but long-run equilibrium multiplicity [Allen and Donaldson, 2020]. Intuitively α_1 and β_1 govern the within-period equilibrium properties of the model, if they are sufficiently small the dynamic equilibrium will exist and be unique. More specifically, if $\alpha_1 + \beta_1 < 1/\theta$, that is contemporaneous agglomeration forces are greater than the contemporaneous dispersion forces. If this condition holds the model will have a unique transition path.

Turning to the long-run equilibrium, that is the equilibrium to which the economy's transition path is converging. As discussed by Allen and Donaldson [2020], intuitively if agglomeration forces are strong enough location decisions may become self-reinforcing, and thus multiple equilibria could arise. The intuition here translates into a very similar condition as that discussed above, only now total agglomeration forces are what is important. That is multiple long-run equilibria will arise if $\alpha_1 + \alpha_2$ and $\beta_1 + \beta_2$ are sufficiently large.

If this is the case the economy has the potential to exhibit path dependence where different initial population distributions may cause the economy to converge to different long-run equilibria. Indeed, each potential long-run equilibrium will be associated with a "basin of attraction", that is a set of possible initial population distributions that converge to the specific equilibrium. Therefore, although multiple equilibria may exist, it remains an empirical question to ask whether any one specific shock would be sufficient to cause a shift from one equilibrium to the next.

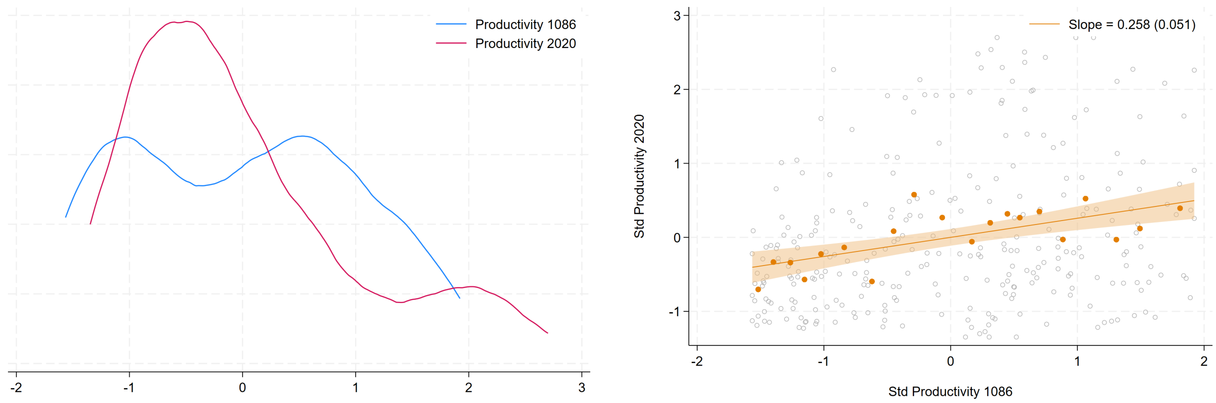
Estimation and calibration.

Although the Doomsday Book presents astonishingly detailed data from around a thousand years ago it does not present any information on flows (of individuals or goods) over space, nor any changes over time in local populations. Without such information, it is difficult to credibly identify parameters in the above-described dynamic quantitative spatial economics model. Rather than attempt to do so in-credibly, in this paper I take the approach of using previously-estimated parameters from the literature. I take the values used in Allen and Donaldson [2020] who consider the setting of the United States between 1800 and 2000.

3.2.2 Results

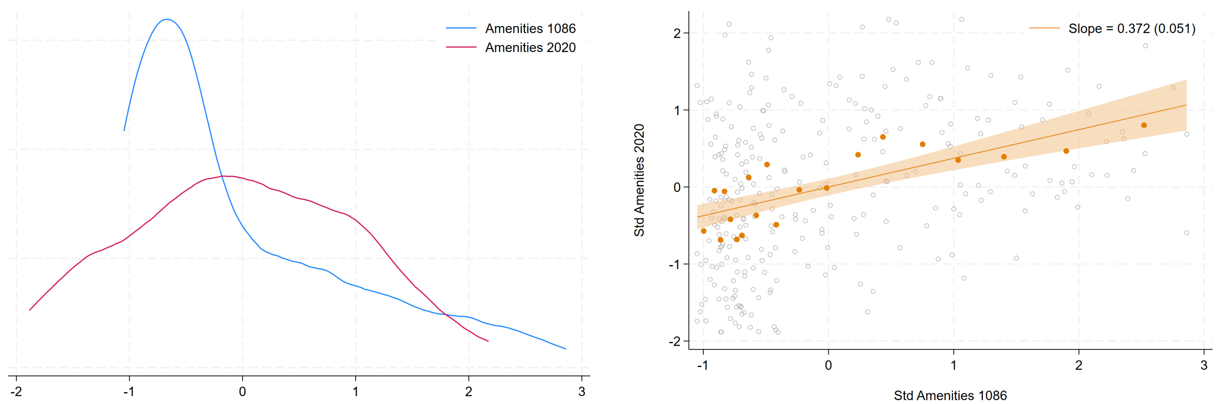
Figures 4 and 5 show the distribution of the estimated productivity and amenity fundamentals and along with table 4 their relationship to each other. Across time productivities and amenities show a clear correlation of around 0.25 — areas that are productive in 1086 are likely to also be productive in 2020 (and similarly for amenities). The right-hand side panel of each of these figures shows that this long-run temporal relationship can also be described to a first-order approximation as linear. Areas that were one standard deviation more productive in 1086 are on average 0.28 standard deviations more productive in 2020, and similarly, those with one standard deviation higher amenities in 1086 have on average 0.37 standard deviation higher amenities in 2020. Table 4 also shows that amenities and productivities are negatively correlated. This is intuitive, often the nicest areas to live in are not the most productive, but rather those with the most beautiful nature, best climate, or best schools/ public infrastructure.

Figure 4 Estimated local productivities in 1086 and 2020



Notes: This figure displays the estimated productivity fundamentals in both 1086 and 2020. In the left-hand panel, I plot the distribution of the standardised data series in blue in 1086 and in red in 2020. In the right-hand panel, I plot the relationship between these two series in a scatter plot with an overlayed binscatter. The x-axis plots 1086 values, and the y-axis 2020 values. In orange, I show ventile averages (binscatter plot) and overlay a corresponding linear line of best fit and confidence interval. The top right-hand corner gives the slope of this line and the associated standard errors.

Figure 5 Estimated local amenities in 1086 and 2020



Notes: This figure displays the estimated amenity fundamentals in both 1086 and 2020. In the left hand pane I plot the distribution of the standardised data series in blue in 1086 and in red in 2020. In the right hand panel I plot the relationship between these two series in a scatter plot with an overlayed binscatter. The x-axis plots 1086 values, and the y-axis 2020 values. In orange I show ventile averages (binscatter plot) and overlay a corresponding linear line of best fit and confidence interval. The top right hand corner gives the slope of this line and associated standard errors of this line.

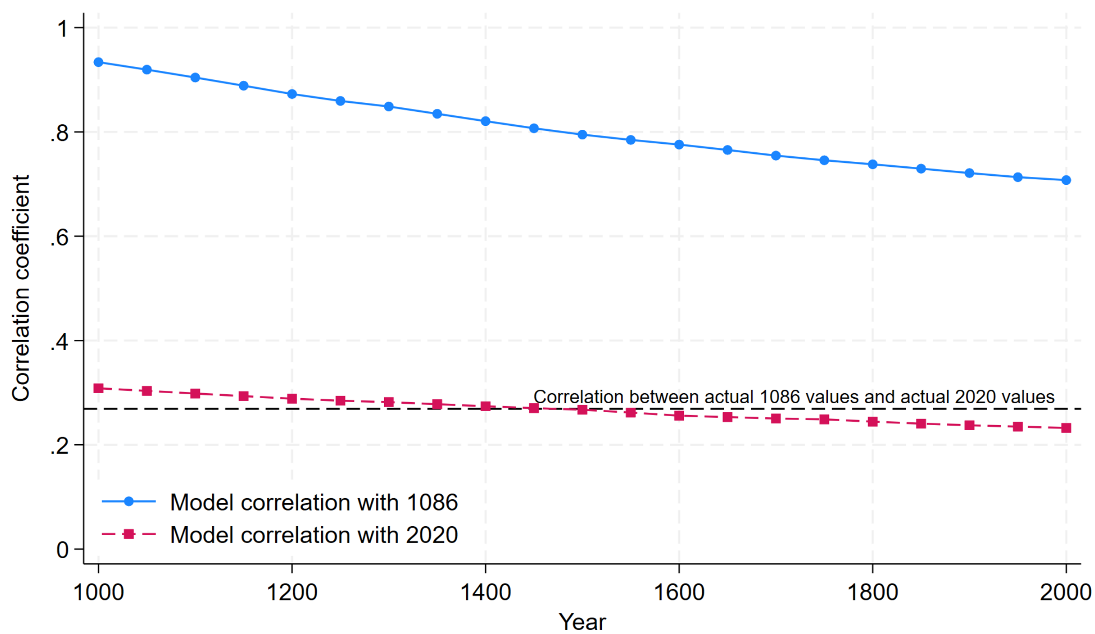
Table 4 Correlations between estimated fundamentals

	C Productivity 1086	Productivity 2020	Amenities 1086	Amenities 2020
Productivity 1086	1.00	0.26	-0.88	-0.29
Productivity 2020	0.26	1.00	-0.31	-0.87
Amenities 1086	-0.88	-0.31	1.00	0.37
Amenities 2020	-0.29	-0.87	0.37	1.00

Notes: The table reports the raw pairwise correlations between the four estimated location fundamentals.

A key question is whether this model-implied evolution of the spatial distribution of economic activity is converging towards that of 2020. Is it the case that the 2020 distribution looks more like the 1086 rolled forward to 2020 than it does the raw 1086 distribution? If so this would suggest that in 1086 the economy was converging towards a spatial equilibrium that more closely resembles the distribution in 2020. Figure 6 investigates this by correlating the model-implied values (keeping amenities and productivities fixed at 1086 values), with those from 2020. The estimated correlation in 2020 is lower than that between the raw 1086 values and 2020 values. This suggests that the 1086 distribution is tending towards a spatial equilibrium that is correlated with the 2020 distribution, but less so than the raw 1086 values. Figure 6 also shows that the correlation between model-implied and 1086 values decreases over time, indicating that the spatial economy in 1086 was not in spatial equilibrium. Note that this implies that even if local fundamentals had not changed if the economy had developed as suggested in figure 6 we would only expect a maximum correlation of around 0.7 between 1086 values and 2020 values. Reflecting the fact that the economy dynamically evolves over that period slowly toward spatial equilibrium.

Figure 6 Correlation between the actual and the model-implied spatial distributions of economic activity



Notes: This figure shows the cross local authority correlation between value per capita in 1086 in blue with circular marks, and in 2020 in red with square marks, with model-implied income per capita over time. The model implied distributions of economic activity over time are calculated by solving forward the dynamic quantitative spatial economics framework discussed in the main text. When performing this counterfactual I keep fundamental productivities and amenities fixed at their recovered 1086 distribution. The horizontal dashed line indicates the raw correlation between 1086 values and 2020 values.

Table 5, shows how controlling for estimated local fundamentals affects long-run persistence in incomes across space. Column one replicates the main results from section 2 showing long-run persistence in incomes over space. Column two then controls for (log) local amenities and productivities. The coefficient on log 1086 incomes decreases from 0.195 to 0.00006. If one were to look at two locations with comparable local fundamentals, additionally controlling for 1086 incomes adds no further explanatory power in predicting incomes today. The coefficient as well as significantly attenuating is also precisely estimated. The coefficients on local amenities and productivities are signed as one would expect. Amenities decrease local incomes as they incentivize "too many" people to move to a location causing negative agglomeration effects and putting downward pressure on wages through excessive labor supply. Productivities increase local incomes by increasing demand for local products and so local labor demand and wages which are not competed away due to costly migration and local amenities. Subsequent columns in table 5 show the robustness of this result to various alternative specifications.

Table 5 The impact of controlling for local fundamentals on long run persistence

	Log GDP per-capita 2020	Log GDP per-capita 2020	Log GDP per-capita 2020	Log GDP per-capita 2020	Log GDP per-capita 2020	LA rank 2020
Log values per capita 1086	0.195***	0.0000559	0.0140**	0.0116	0.00451	
	(0.0311)	(0.00518)	(0.00653)	(0.00743)	(0.00732)	
Rank values per capita 1086						0.0115
						(0.0139)
Log Amenities 2020		-0.267***	-0.268***	-0.281***	-0.201***	51.28***
		(0.0124)	(0.0131)	(0.0142)	(0.0165)	(7.231)
Log Productivity 2020		0.658***	0.659***	0.645***	0.729***	-182.7***
		(0.0141)	(0.0162)	(0.0154)	(0.0185)	(7.759)
Weighting		None	1086 Pop	None	None	None
Lat-lng polynomial				Yes		
Region FE					Yes	
Observations	283	283	283	283	283	283
R2	0.0971	0.987	0.987	0.987	0.989	0.947

Notes: This table shows the results from regressing the spatial distribution of income in 2020 against that in 1086 controlling for local fundamentals, across various specifications. For reference in the first column, I reprint the unconditional correlation. In the second column, I show the raw results in a log-log specification controlling for local fundamentals. Column three weights the regression by the 1086 population. Column four includes a second-order interacted polynomial in local-authority centroids. Column five includes fixed effects for the 9 high-level regions of England. Finally, column six performs a rank-rank regression.

In figure 15 in the appendix I show further that the result that conditing on fundamentals kills the long-run income persistence is robust to various alternative model parameters. The model relies on six parameters that I have taken from the literature. To check robustness to these I calculate local fundamentals over 1,000 alternative parameters. For each set of parameters I randomly vary each individual parameter to be between 0.25 and 2.25 of its original value — recalculate local amenities and productivities and then condition on them in the long-run persistence regression. I then only include in my analysis parameter combinations that result in a unique per-period equilibrium but possibly multiple long-run spatial equilbria. Figure 15 then shows the recovered coefficient on log values in 1086 in said regression. In all cases, this coefficient is no more than around 25% of the unconditional coefficient, and in only 10% of simulations is it significantly different from 0 at the 5% level. Figure 15 also shows which parameter constellations correspond to which coefficients. Particularly problematic are large values (in absolute terms) of α_1 and β_1, and small values of θ and σ. That is, larger contemporaneous agglomeration forces and more dispersed fundamentals.

3.3 What location characteristics mediate the correlated spatial equilibria

Table 5 shows that once estimated fundamentals have been controlled for, the previously observed long-run correlation becomes a precisely estimated zero. This, in conjunction with the reduced form evidence from the Harrying, suggests that fundamentals rather than path dependency can explain the observed relationship. Intuitively the fundamentals explanation suggested that some characteristics "X" were associated with richer areas in both 1086 and 2020. Therefore, the natural next question is: What are the X's? To answer this question I build a local-authority level database of time-invariant (since 1086) characteristics. Using this data I first attempt to explain the correlation between the estimated 1086 and 2020 amenities and productivities. This identifies a few variables that explain the correlation in fundamentals and so should explain the correlation in values if it is indeed fundamentals that are driving that correlation. Finally, I check that this is the case in table 6.

I consider four categories of local characteristics. First, ruggedness. Ruggedness captures how flat the local terrain is, flatter areas have lower ruggedness values. These areas are much more amenable to agriculture, much easier to build transport infrastructure to/from, and indeed to build on in general [Nunn and Puga, 2012, Henderson et al., 2018]. Therefore, one could expect ruggedness to be an important determinant of local fundamentals both in 1086 and 2020. In figures 7 and 8 I show in blue with circular markers the impact of controlling for ruggedness on the correlation between 1086 and 2020 amenities and productivities respectively. Both figures show that ruggedness has no impact on the long-run correlation of local fundamentals.

Second, I consider the role of infrastructure that already existed in 1086. From the

Doomsday Book itself, there is data on local mills and fisheries which I aggregate into the raw number in each modern local authority. I then also include a dummy variable that takes the value one if a major Roman town lies inside the local authority and another dummy variable that takes the value one if a major Roman road runs through the local authority. One can clearly see how previously existing infrastructure may have increased local productivity in 1086, but these areas may also be more productive in 2020. This could be the case if access to the following amenities is predictive of growth today: running water (mills), coastal access (fisheries), natural thoroughfares (Roman roads), or preexisting urban centers (Roman cities) [Dalgaard et al., 2022, Michaels and Rauch, 2018]. In figures 7 and 8 I show in red with diamond markers the impact of controlling for preexisting infrastructure on the correlation between 1086 and 2020 amenities and productivities respectively. Both figures show that preexisting infrastructure (measured in this way) has no impact on the long-run correlation of local fundamentals.

Third, I consider variables that capture the geographic centrality, or market access, of a location within England. The variables I use are: distance to the coast, distance to London[5], a measure of market potential based on Roman settlements, and a measure of market potential based on the 1086 population. Measures capturing some notion of local market access have been shown to have been associated with various local outcomes including local wages [Donaldson and Hornbeck, 2016, Redding and Sturm, 2008, Asher and Novosad, 2020]. In figures 7 and 8 I show in green with triangular markers the impact of controlling for market access on the correlation between 1086 and 2020 amenities and productivities respectively. Figure 7 shows that (other than the distance to the coast) market access variables do mediate the long-run correlation between estimated amenities. Although not quite statistically significant figure 8 also shows some mediating power for market potential terms in explaining the long-run productivity correlation.

Finally, I consider variables that capture measures of local agricultural productivity. Using data on attainable yields from the FAO's global agro-ecological zones database I construct measures of attainable yields for four key staple crops: wheat, oats, rye, and barley. FAO's measures use data on local climatic and geological features from the present day and so will not perfectly reflect conditions in 1086. However, I use their measure with no modern inputs, and consider it unlikely that local changes would be endogenous to local outcomes i.e. that such a measure would be a "bad control". In figures 7 and 8 I show in purple with square markers the impact of controlling for local agricultural productivity on

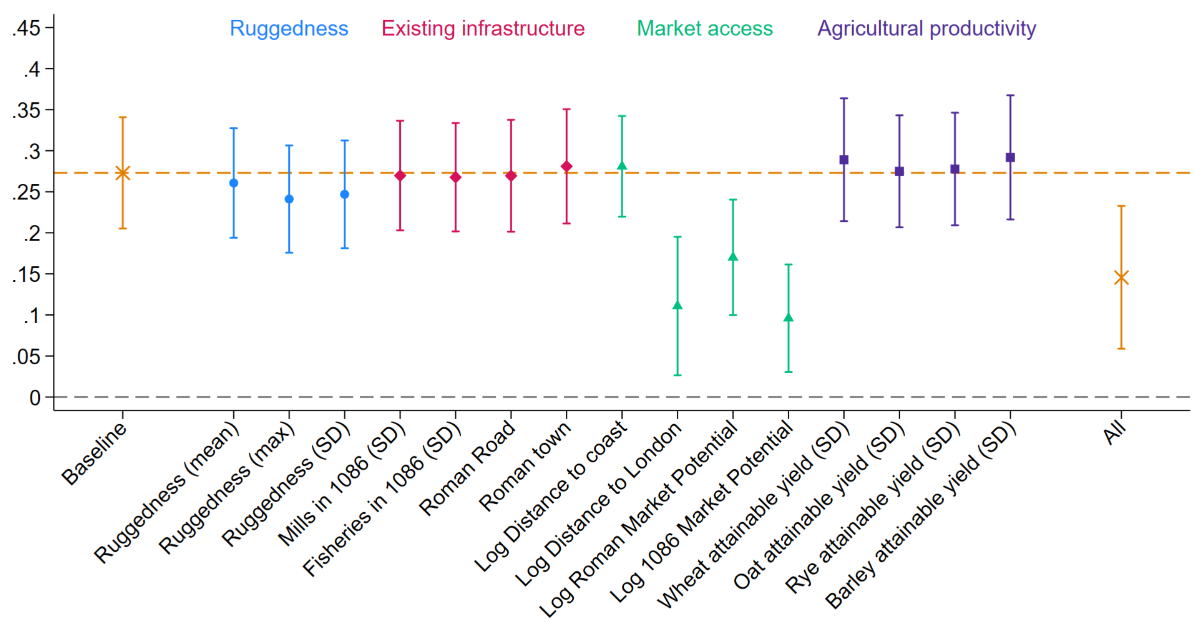
[5] London was already the most populous city in 1086.

the correlation between 1086 and 2020 amenities and productivities respectively. In both figures, one can see that local agricultural productivity is not a significant mediator of the respective long-run relationships.

In the final column in each of figures 7 and 8 I show the long-run relationship between estimated amenities and productivities conditing on all the previously discussed local characteristics.

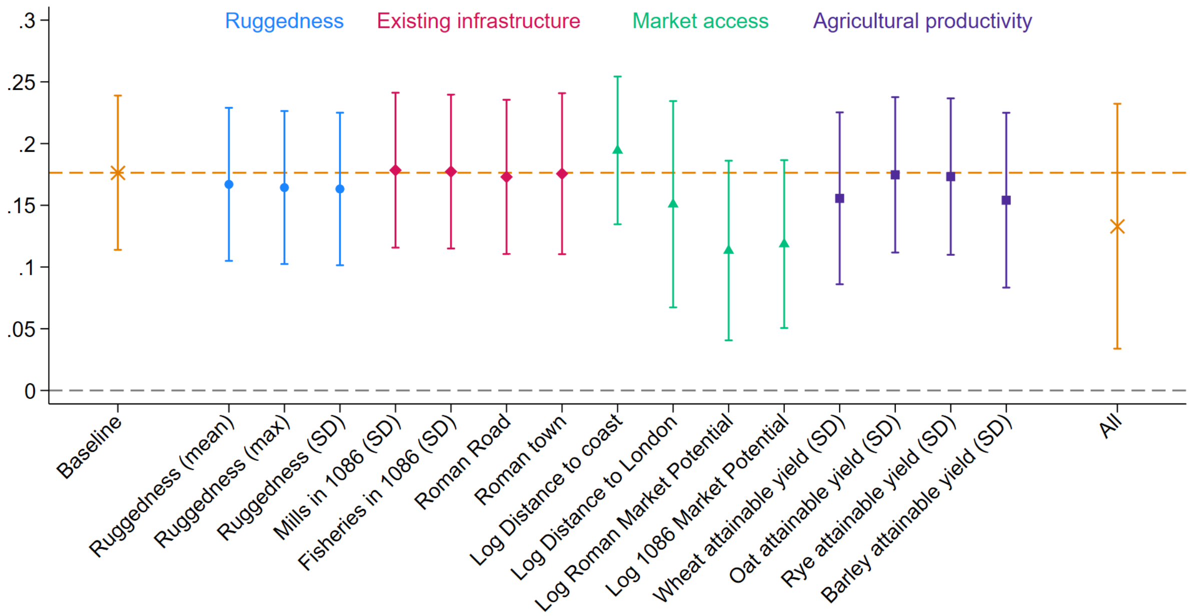
These characteristics are directly correlated with estimated amenities and productivities as one would intuitively expect. Figures 11, 12, 13, and 14 in the appendix show these raw relationships. Most variables show a significant relationship. Ruggedness and distance to London are positively correlated with local amenities in both periods, and mills, fisheries, and some measures of soil quality are also positively correlated in 2020. Market potential is however consistently negatively correlated with local amenities. Productivities show somewhat the mirror picture: ruggedness and distance to London is negatively correlated whereas market potential and soil quality is positively related.

Figure 7 Explaining the spatial correlation between 1086 and 2020 amenities



Notes: This figure displays β_x in the following regression $\text{Amenity}_{2020,i} = \beta_x \text{Amenity}_{1086,i} + \alpha_x X_i + \varepsilon_i$, where X_i denotes individual characteristics indicated on the x-axis. 95 percent confidence intervals are indicated by horizontal lines and are estimated using robust standard errors. "Baseline" indicates a lack of covariate and thus reports the raw correlation between 1086 and 2020 values.

Figure 8 Explaining the spatial correlation between 1086 and 2020 productivities



Notes: This figure displays β_x in the following regression $\text{Productivity}_{2020,i} = \beta_x \text{Productivity}_{1086,i} + \alpha_x X_i + \varepsilon_i$, where X_i denotes individual characteristics indicated on the x-axis. 95 percent confidence intervals are indicated by horizontal lines and are estimated using robust standard errors. "Baseline" indicates a lack of covariate and thus reports the raw correlation between 1086 and 2020 values.

The long-run relationship between local fundamentals can be in part mediated through the local characteristic of market access. I now turn to ask whether market access can also partly explain the observed long-run correlation in incomes. Table 6 provides the results from regressing log GDP per capita in 2020 against log value per capita in 1086 controlling for various combinations of the characteristics described above. In column one, I replicate the raw correlation. In column two I control for ruggedness, in column three for historic infrastructure, in column four for market potential, in column five for soil quality, in column six for soil quality and market potential, and finally in column seven for all of the above. Only when market access is included do we see any meaningful changes to the estimated persistence relationship. It appears that market access, measured in this way, explains about half of the long-run correlation in incomes over space.

Table 6 Accounting for long-run spatial persistence

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log value pc 1086	0.195***	0.182***	0.195***	0.0977***	0.194***	0.0851**	0.0894**
	(0.0311)	(0.0309)	(0.0317)	(0.0355)	(0.0357)	(0.0391)	(0.0417)
Log Ruggedness (mean)		-0.0517**					-0.0282
		(0.0225)					(0.0222)
Mills (SD)			-0.0298**				-0.0231*
			(0.0132)				(0.0126)
Fisheries (SD)			-0.0334***				-0.0275**
			(0.0128)				(0.0124)
Roman Road			0.0923**				0.0635
			(0.0406)				(0.0400)
Roman town			-0.00930				0.0468
			(0.0528)				(0.0579)
Log market potential				0.232***		0.237***	0.195***
				(0.0518)		(0.0513)	(0.0590)
Barley Attainable Yield (SD)					0.000950	0.0144	0.0174
					(0.0233)	(0.0221)	(0.0231)
Constant	10.51***	10.63***	10.45***	8.800***	10.51***	8.746***	9.071***
	(0.0518)	(0.0800)	(0.0606)	(0.375)	(0.0584)	(0.373)	(0.458)
Observations	283	282	283	283	283	283	282
R2	0.0971	0.118	0.132	0.168	0.0971	0.169	0.194

Notes: This table reports results from regressing log GDP per capita in 2020 against Log value per capita in 1086 controlling for various combinations of covariates. In column one, I report the raw relationship. In column two I control for the ruggedness of a location. In column three I control for measures of historic infrastructure. In column four I control for log market potential. In column five I control for the soil quality of a location proxied by the attainable yield of barley. Finally, in column six I control for both log market potential and soil quality. Standard errors are robust. Stars indicate usual significant levels.

4 Conclusion

In this paper, I have shown that the spatial distribution of income we see today in England was, at least partly, already in place 1,000 years ago. To do this I have leveraged unique data on the economic circumstances of individual manors in 1086 England, available in the Doomsday Book commissioned by William the Conqueror. I show that areas that were 10% richer in 1086 are on average almost 2% richer today and that this relationship is robust to a barrage of corrections for potential spatial correlation.

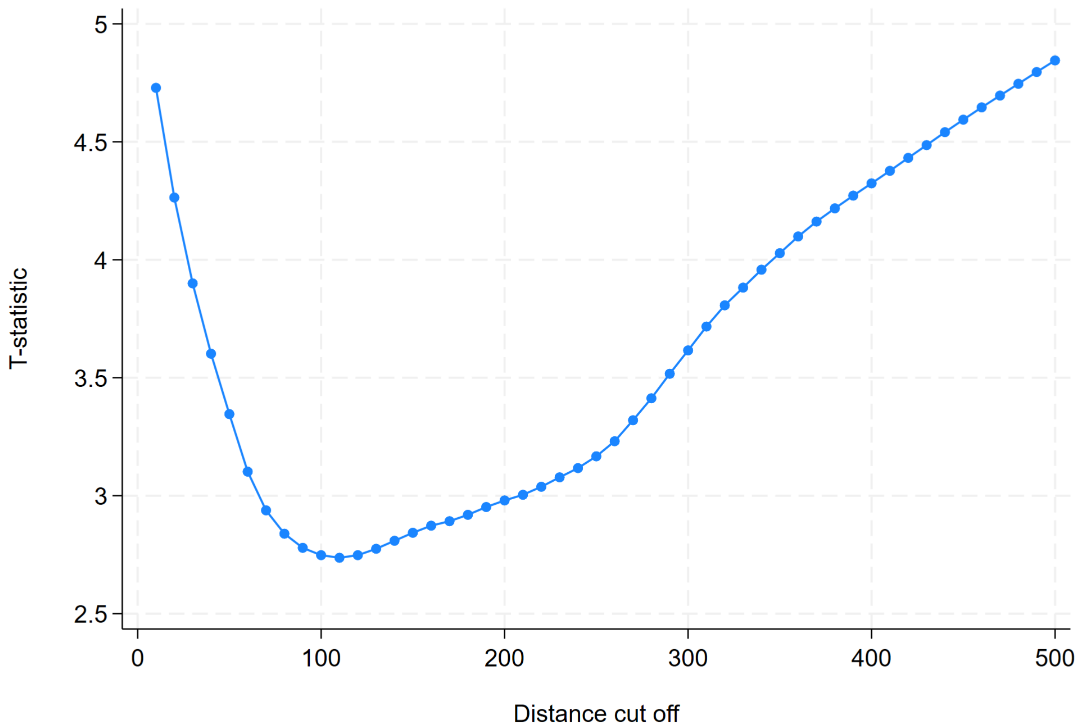
I then show evidence, using the natural experiment of the Harrying of the North and a dynamic quantitative spatial economics model, that this long-run correlation is not driven by path dependence, but rather by local fundamentals, and in particular local market access.

Recently (and historically) politicians have championed spatially redistributive policies such as the Northern Powerhouse and the Leveling up initiative. This work shows that such initiatives will not rebalance the long-run distribution of economic activity unless they alter local fundamentals — and it is here that governments aiming for spatial inequality should focus.

A Appendix

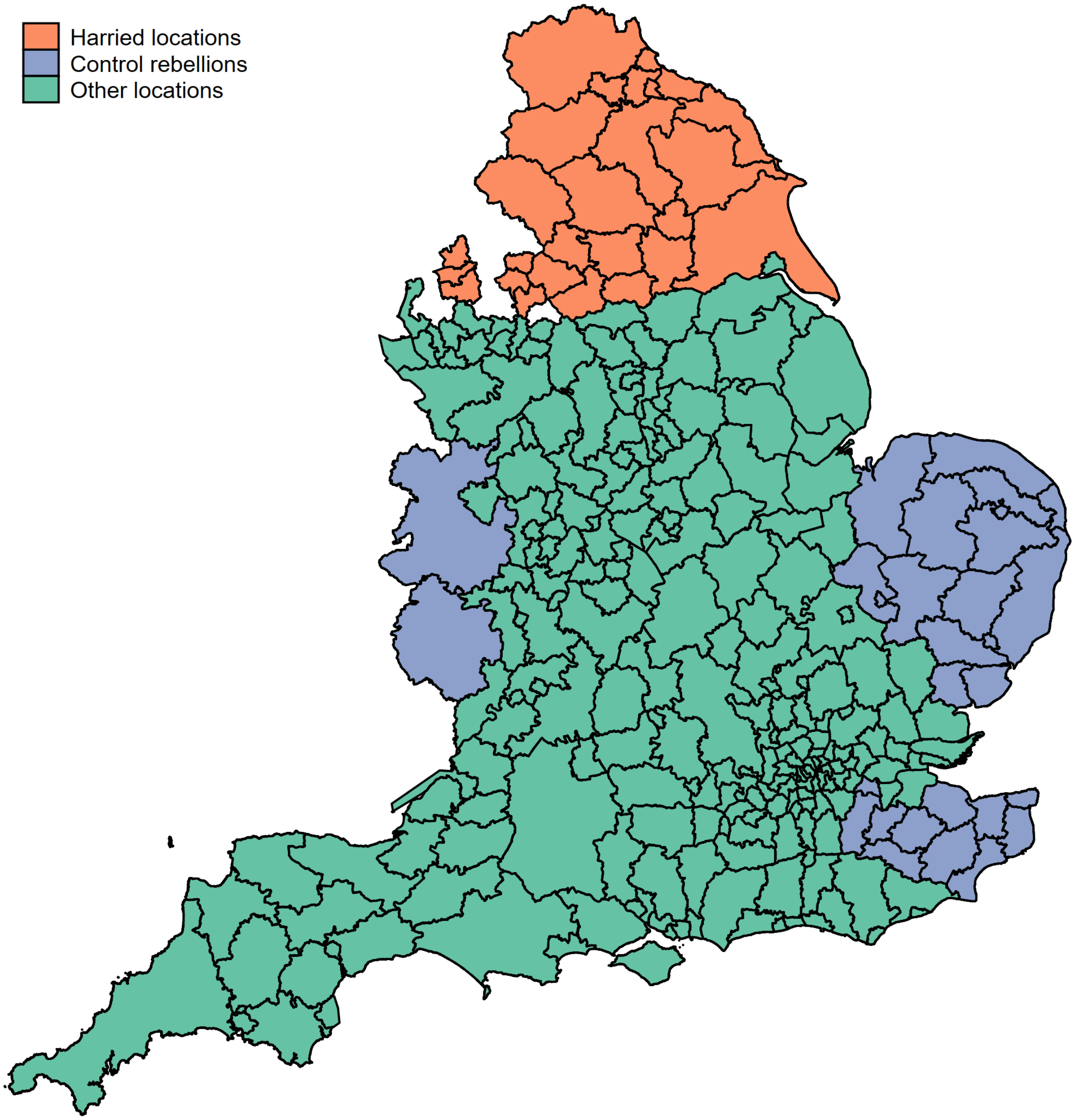
A.1 Additional figures and tables

Figure 9 Conley T-statistic over specified distance cutoff



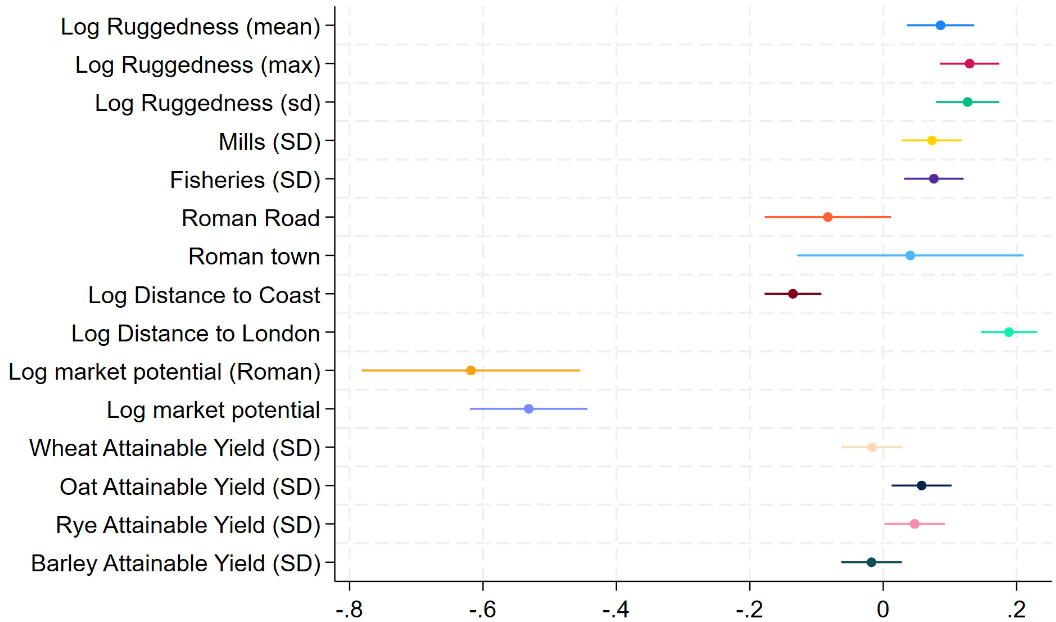
Notes: This figure shows the estimated t-statistics on the main long-run persistence coefficient of interest, β in the following regression: $\ln(2020 GDP per capita)_i = \beta \ln(1086 Value per capita)_i + \varepsilon_i$. For each regression I adjust the standard errors following the procedure due to Conley [1999] and report the recovered test statistic over various cutoffs as indicated on the x-axis.

Figure 10 Harried and other rebellious control local authorities



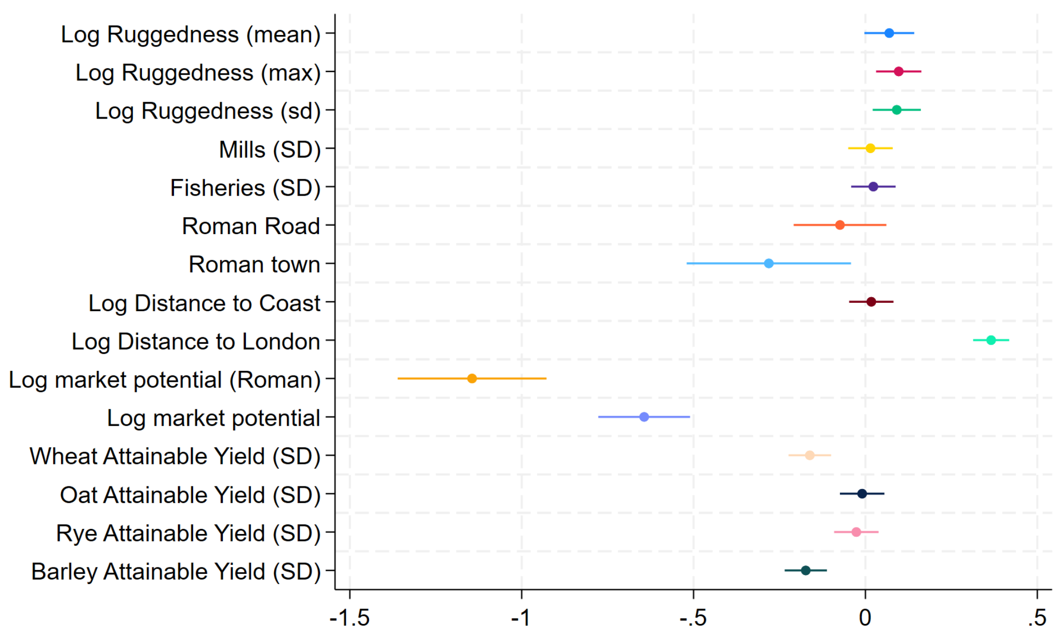
Notes: This figure shows the Harried and other rebellious control local authorities in orange and blue respectively.

Figure 11 Correlation between 2020 amenities and local characteristics



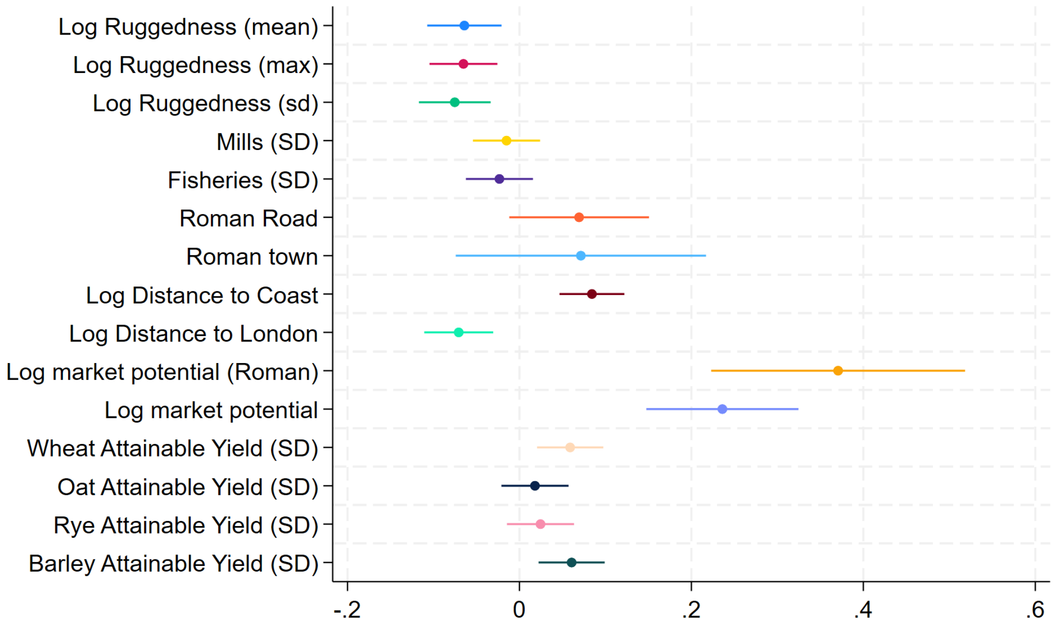
Notes: This figure shows the correlation between the estimated local amenities in 2020 and time-invariant local characteristics. Standard errors are robust and 95% confidence intervals are shown.

Figure 12 Correlation between 1086 amenities and local characteristics



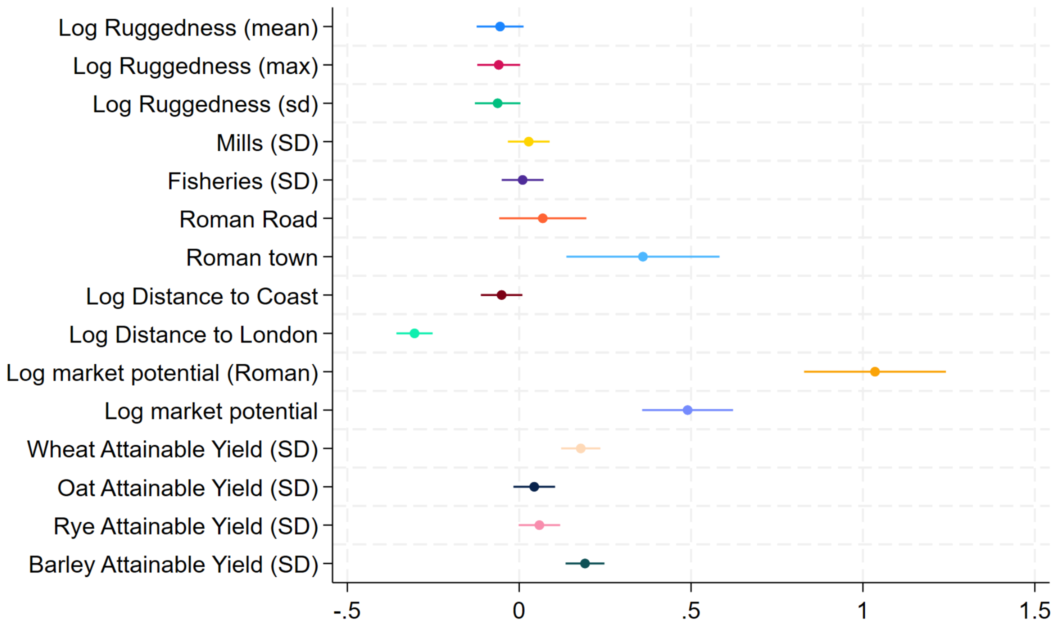
Notes: This figure shows the correlation between the estimated local amenities in 1086 and time-invariant local characteristics. Standard errors are robust and 95% confidence intervals are shown.

Figure 13 Correlation between 2020 productivities and local characteristics



Notes: This figure shows the correlation between the estimated local productivities in 2020 and time-invariant local characteristics. Standard errors are robust and 95% confidence intervals are shown.

Figure 14 Correlation between 1086 productivities and local characteristics



Notes: This figure shows the correlation between the estimated local productivities in 1086 and time-invariant local characteristics. Standard errors are robust and 95% confidence intervals are shown.

Table 7 Spatial persistence over the very long run robustness

	(1) Without winsorizing	(2) Excluding Harried LA's	(3) Excluding Modern London	(4) Median 2020 Wages	(5) Average 2020 Wages	(6) In Levels
Value per-capita 1086 (logs)	0.159***	0.204***	0.189***	0.117***	0.153***	
	(0.0261)	(0.0353)	(0.0325)	(0.0133)	(0.0161)	
Value per-capita 1086						25154.1***
						(4271.0)
Constant	10.46***	10.53***	10.49***	6.352***	6.570***	22471.0***
	(0.0454)	(0.0556)	(0.0552)	(0.0225)	(0.0274)	(1102.9)
Observations	283	257	251	283	282	283
R^2	0.094	0.093	0.103	0.185	0.214	0.090

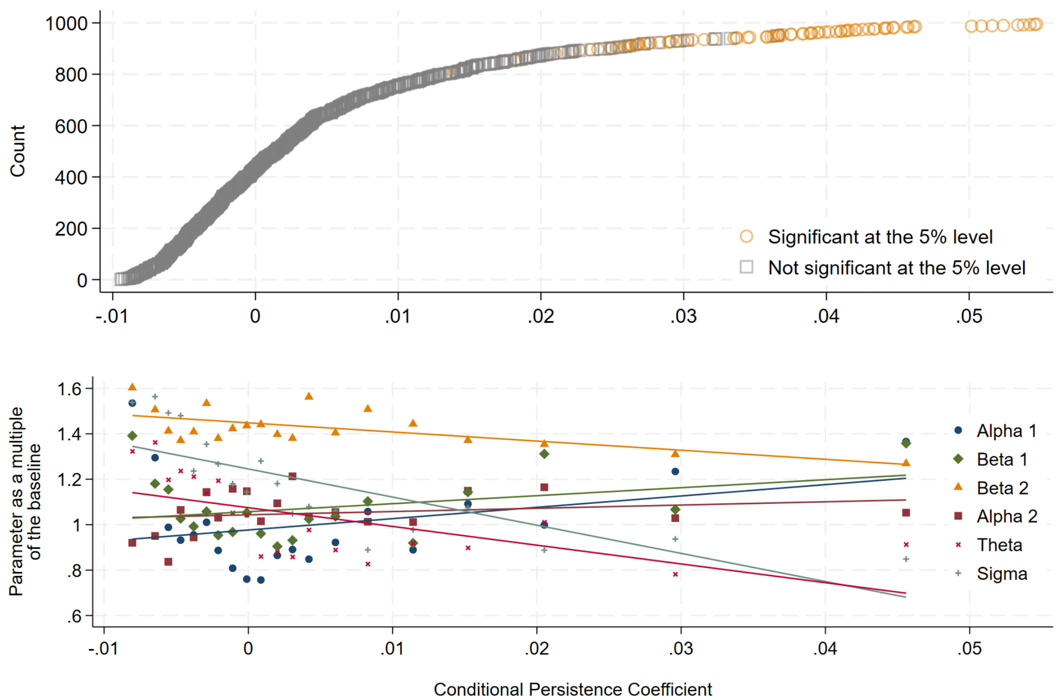
Notes: This table shows the robustness of the main persistence result to various alternative specifications. In column one I show the results without first winsorising either variable. In column two I exclude local authorities in the North that are Harried. In column three I exclude local authorities that consist of modern-day London. In column four I use modern median wages in a local authority from the ASHE data as the dependent variable. In column five I use modern average wages in a local authority from the ASHE data as the dependent variable. In column six I show the results using each variable in levels as opposed to logged values.

Table 8 Harried two way fixed effects table

	Baseline	Weight by 1086 pop	Weight by 2020 pop
Harried × 1086	-1.644***	-2.323***	-1.574***
	(0.269)	(0.370)	(0.271)
Harried × 2020	-0.104	-0.745*	-0.197
	(0.301)	(0.400)	(0.323)
Observations	162	162	162
R2	0.644	0.727	0.633

Notes: This table shows the results from estimating a two way fixed effects model with local authority and period fixed effects. It considers value per density over three time periods, 1066 (omitted category), 1086, and 2020. Standard errors are clustered at the local authority level. In column one the regression is unweighted. In column two I weight by 1086 population. In column three I weight by 2020 population.

Figure 15 Robustness to model parameters



Notes: This figure shows the robustness of the result that conditioning on local fundamentals explains the long run persistence to various model parameters. In the upper panel, I show the CDF of the coefficient on 1086 incomes in a regression of said incomes on 2020 incomes controlling for local fundamentals estimated using a random set of parameters. The parameters used are given in the lower panel which is a binscatter of said parameters over the estimated coefficient (each figure shares the same x-axis). In the upper panel orange circular markers indicate that the coefficient on 1086 incomes is significant at the 5% level, whereas a gray square indicates that it is not. Overall 10% of the parameter specifications considered result in a significant effect for 1086 incomes. In the lower panel lines of best fit are given. In blue with circular markers, I plot the values of α_1. In green with diamond markers, I plot the values of β_1. In orange with triangular markers, I plot the values of β_2. In red with square markers, I plot the values of α_2. In orange with x markers, I plot the values of θ. In green with $+$ markers, I pot the values of σ.

A.2 Historical details

A.2.1 Rebellions against Williams rule

Barlow [2014], Rex [2014]

- North: In 1068 Edwin and Morcar, the earls of Mercia and Northumbria started a rebellion supported by Edgar who was crowned king by the English after the battle of Hastings and subsequently gave up his crown to William. The rebellion started because William had made the earls lands smaller and imposed a heavy tax. William crushed the rebellion but forgave Edwin and Morcar who were kept as guests in his court.

- North: In 1069 Edgar joined forces with Sweyn of Denmark and attacked York, afterwards their forces scattered causing small rebellions all over the country. William paid the Danes to leave and then pursued a campaign of destruction on the North — the Harrying of the north.

- East Anglia: In 1070-71 the Danes arrived in the fens in the east of England and were helped by a local ruler Hereward the Wake. The rebellion ended when the Normans captured their center of power in Ely.

- East Anglia and Northumbria: The revolt of the earls 1075. Ralph de Guader Earl of East Anglia, Roger de Breteuil Earl of Hereford, and Waltheof Earl of Northumberland revolted. Waltheof almost immediately gave up and confessed the rebellion. Rodger was stopped by the English bishop Wulfstan's fyrd. Ralph was similarly stopped near Cambridge and eventually fled to Denmark.

- Exeter: After the battle of Hastings Gytha, Harold's mother, fled to Exeter and from there fermented rebellion against William. After Exeter refused to swear fealty, or pay taxes, to William in 1068 he lay siege to Exeter. After 18 days William and his forces entered the city. By all accounts William gave very generous terms, allowing the city to not pay taxes in return for pledging fealty, and the Williams soldiers were denied their traditional right of looting the surrendered city.

- Welsh borders: William never conquered Wales, and indeed never appeared to intend to. However, Welsh forays and attacks in the border regions caused William to establish a series of earldoms in the borderlands at Chester, Shrewsbury, and Hereford.

- Kent: In Kent, a rebellion quickly surfaced against the local ruler Odo in 1067 who was extremely unpopular with his new subjects. This was put down by Odo without the need for William to personally intervene.

B Dynamic quantitative spatial economics model equilibrium

As described in Allen and Donaldson [2020] the equilibrium of the DQSE can be set out as follows. For any initial population vector $\{L_{i0}\}$ and vectors of geographic fundementals $\{\bar{A}_{it}, \bar{u}_{it}, \tau_{ijt}, \mu_{ijt}\}$ such that for all i, t the following holds.

1. A locations income equals the value of purchases from it: $w_{it}L_{it} = \sum_j X_{ijt}$. Which implies that $w_{it}^{\sigma}L_{it}^{1-\alpha_1(\sigma-1)} = \sum_j K_{ijt}L_{jt}^{\beta_1(\sigma-1)}W_{jt}^{1-\sigma}w_{jt}^{\sigma}L_{jt}$. Where all exogenous and predetermined variables have been bundled together into the Kernel:
   $K_{ijt} = (\tau_{ijt}(\bar{A}_{it}L_{it-1}^{\alpha_2}\bar{u}_{jt}L_{jt-1}^{\beta_2})^{-1})^{1-\sigma}$.

2. Trade is balanced. Income is fully spent $w_{it}L_{it} = \sum_j X_{jit}$. Which implies that: $w_{it}^{1-\sigma}L_{it}^{\beta_1(1-\sigma)}w_{it}^{\sigma-1} = \sum_j K_{ijt}L_{jt}^{\alpha_1(\sigma-1)}w_{jt}^{1-\sigma}$.

3. Total population equals the sum of those arriving. $L_{it} = \sum_j L_{ijt}$. This implies that: $L_{it}V_{it}^{-\theta} = \sum_j \mu_{ijt}^{-\theta}\Pi_{jt}^{-\theta}L_{jt-1}$.

4. Total population in the previous period equals the sum of those leaving. $L_{it-1} = \sum_j Lijt$. Which implies that: $\Pi_{it}^{\theta} = \sum_j \mu_{ijt}^{-\theta}W_{jt}^{\theta}$.

We can simplify this system by imposing symmetry in trade costs which implies that 1. and 2. can be combined. Thus we are left with a 3-equation model in each i, t with three unknowns $\{L_{it}, W_{it}, \Pi_{it}\}$, and unknown parameters $\{\alpha_1, \alpha_2, \beta_1, \beta_2, \sigma, \theta\}$.

References

Treb Allen and Dave Donaldson. Persistence and path dependence in the spatial economy. Technical report, National Bureau of Economic Research, 2020.

Treb Allen, Costas Arkolakis, and Xiangliang Li. On the equilibrium properties of network models with heterogeneous agents. Technical report, National Bureau of Economic Research, 2020a.

Treb Allen, Costas Arkolakis, and Yuta Takahashi. Universal gravity. *Journal of Political Economy*, 128(2):393–433, 2020b.

Sam Asher and Paul Novosad. Rural roads and local economic development. *American economic review*, 110(3):797–823, 2020.

Clare Bambra. *Health divides: where you live can kill you.* Policy Press, 2016.

Frank Barlow. *The feudal kingdom of England: 1042-1216.* Routledge, 2014.

Brian Bell, Jack Blundell, and Stephen Machin. Where is the land of hope and glory? the geography of intergenerational mobility in england and wales. *The Scandinavian Journal of Economics*, 125(1):73–106, 2023.

Hoyt Bleakley and Jeffrey Lin. Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644, 2012.

Richard H Britnell and Bruce MS Campbell. *A commercialising economy: England 1086 to c. 1300.* Manchester University Press, 1995.

Jack Britton, Laura van der Erve, Ben Waltmann, and Xiaowei Xu. London calling? higher education, geographical mobility and early-career earnings. *Institute for Fiscal Studies Report, https://ifs. org. uk/publications/london-calling-higher-education-geographical-mobility-and-early career-earnings*, 2021.

Social Mobility Commission. State of the nation 2018-19: social mobility in great britain. apr 2019, 2019.

Timothy G Conley. Gmm estimation with cross sectional dependence. *Journal of econometrics*, 92(1):1–45, 1999.

Carl-Johan Dalgaard, Nicolai Kaarsen, Ola Olsson, and Pablo Selaya. Roman roads to prosperity: Persistence and non-persistence of public infrastructure. *Journal of Comparative Economics*, 50(4):896–916, 2022.

Paul Dalton. *Conquest, anarchy and lordship: Yorkshire, 1066-1154*. Number 27. Cambridge University Press, 2002.

Henry Clifford Darby and Henry Clifford Darby. *Domesday England*. Cambridge University Press, 1986.

Donald R Davis and David E Weinstein. Bones, bombs, and break points: the geography of economic activity. *American economic review*, 92(5):1269–1289, 2002.

Vincent Delabastita and Sebastiaan Maes. The feudal origins of manorial prosperity: Social interactions in eleventh-century england. *The Journal of Economic History*, 83(2):464–500, 2023.

Melissa Dell. The persistent effects of peru's mining mita. *Econometrica*, 78(6):1863–1903, 2010.

Dave Donaldson and Richard Hornbeck. Railroads and american economic growth: A "market access" approach. *The Quarterly Journal of Economics*, 131(2):799–858, 2016.

Vivian H Galbraith. An episcopal land-grant of 1085. *The English Historical Review*, 44 (175):353–372, 1929.

Cecile Gaubert, Patrick M Kline, and Danny Yagan. Place-based redistribution. Technical report, National Bureau of Economic Research, 2021.

Stephen Gibbons, Stephan Heblich, and Ted Pinchbeck. The spatial impacts of a massive rail disinvestment program: the beeching axe. 2018.

Mark Hagger. *William: King and Conqueror*. Bloomsbury Publishing, 2021.

Luke Heath Milsom. Persistent spatial equilibria. evidence from a sudden river. Technical report, The Institute for Fiscal Studies, 2024.

J Vernon Henderson, Tim Squires, Adam Storeygard, and David Weil. The global distribution of economic activity: nature, history, and the role of trade. *The Quarterly Journal of Economics*, 133(1):357–406, 2018.

David Horspool. *The English Rebel: One Thousand Years of Trouble-making from the Normans to the Nineties*. Penguin UK, 2009.

Remi Jedwab, Edward Kerby, and Alexander Moradi. History, path dependence and development: Evidence from colonial railways, settlers and cities in kenya. *The Economic Journal*, 127(603):1467–1494, 2017.

Remi Jedwab, Noel D Johnson, and Mark Koyama. Pandemics, places, and populations: Evidence from the black death. 2019.

Morgan Kelly. The standard errors of persistence. 2019.

Morgan Kelly, Joel Mokyr, and Cormac Ó Gráda. The mechanics of the industrial revolution. *Journal of Political Economy*, 131(1):59–94, 2023.

Jeffrey Lin and Ferdinand Rauch. What future for history dependence in spatial economics? *Regional Science and Urban Economics*, 94:103628, 2022.

Philip McCann. Perceptions of regional inequality and the geography of discontent: Insights from the uk. *Regional Studies*, 54(2):256–267, 2020.

John McDonald and Graeme D Snooks. The determinants of manorial income in domesday england: evidence from essex. *The Journal of Economic History*, 45(3):541–556, 1985.

Guy Michaels and Ferdinand Rauch. Resetting the urban network: 117–2012. *The Economic Journal*, 128(608):378–412, 2018.

Edward Miguel and Gerard Roland. The long-run impact of bombing vietnam. *Journal of development Economics*, 96(1):1–15, 2011.

Luke Heath Milsom. Moving opportunitylocal connectivity and spatial inequality. Technical report, CEPREMAP, 2023.

Ulrich K Müller and Mark W Watson. Spatial correlation robust inference. *Econometrica*, 90(6):2901–2935, 2022.

Nathan Nunn and Diego Puga. Ruggedness: The blessing of bad geography in africa. *Review of Economics and Statistics*, 94(1):20–36, 2012.

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

J. Palmer. Electronic edition of domesday book: Translation, databases and scholarly commentary, 1086; second edition [computer file]. *http://dx.doi.org/10.5255/UKDA-SN-5694-1*, 2016.

Stephen J Redding and Daniel M Sturm. The costs of remoteness: Evidence from german division and reunification. *American Economic Review*, 98(5):1766–1797, 2008.

Peter Rex. *The English Resistance: The Underground War Againt the Normans*. Amberley Publishing Limited, 2014.

David Roffe. *Decoding Domesday*. Boydell & Brewer Ltd, 2015.

Josiah Cox Russell. British medieval population. *(No Title)*, 1948.

Pauline Stafford. Women in domesday. *Reading Medieval Studies*, 15:75–94, 1989.

Anna Stansbury, Daniel Turner, and Edward Balls. Tackling the uk's regional economic inequality: Binding constraints and avenues for policy intervention. *M-RCBG Associate Working Paper Series*, 2023.

MJ Strickland. *Armies, chivalry and warfare in medieval Britain and France*, volume 7. Paul Watkins, 1998.

Ordericus Vitalis. *The ecclesiastical history of England and Normandy*, volume 3. HG Bohn, 1854.

James T Walker. National income in domesday england. In *Money, prices and wages: Essays in honour of Professor Nicholas Mayhew*, pages 24–50. Springer, 2015.