European
Historical
Economics
Society

EHES Working Paper | No. 262 | August 2024

A Perfect Storm: First-Nature Geography and Economic Development

Christian Vedel,
University of Southern Denmark

EHES Working Paper | No. 262 | August 2024

A Perfect Storm:
First-Nature Geography and Economic Development*

Christian Vedel,
University of Southern Denmark

Abstract

Is geography destiny? What is the role of first-nature geography in determining prosperity? This paper estimates the effect of randomly removing and introducing favorable first-nature geography to a specific region using a difference in difference design. In 1825 a storm created a new natural navigable waterway, bringing trade and prosperity to the otherwise relatively isolated northwestern Denmark. 700 years prior, the same event happened in reverse, when a previous channel closed up between 1086 and 1208. The elasticity of geography-induced market access is estimated to be 1.6, corresponding to 26.7 percent population growth within a generation of the event. Demonstrated mechanisms include trade, fertility, fishing, and the rise of manufacturing. The central finding is replicated in reverse in a register of dated archaeological sites. The 1086-1208 closing caused fewer buildings and sites containing coins. The general insight is the same: First-nature geography determines the levels and location of prosperity.

JEL Codes: N01, N73, O18, R1

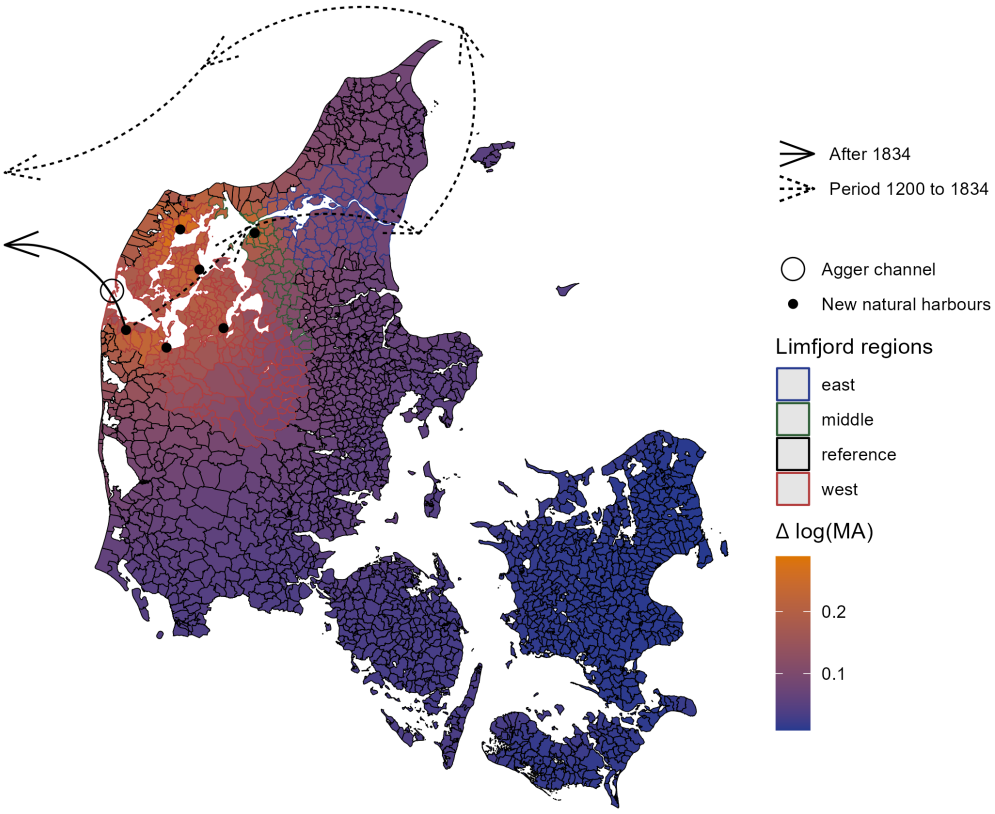
Keywords: First-nature, Trade, Geography, Infrastructure, Natural Experiment

[1]**Corresponding Authors: Christian Vedel (Christian-vs@sam.sdu.dk)**

* This project was made possible through generous funding by the Independent Research Fund Denmark (DFF – 6109-00123). I want to thank Paul Sharp, Christian Møller-Dahl, Casper Worm Hansen, Max Schulze, Gregory Clark, James Fenske, Mathias Barding, Torben Johansen, Nadja van't Hoff, Kerstin Enflo, Neil Cummins, and former PhD student-colleagues (too numerous to mention individually) both at SDU and at LSE for valuable feedback. This project also benefitted from the conscientious research assistance of Andreas Slot Ravnholt. This project was presented at the 2021 and 2022 DGPE workshops, the 2022 EHES conference (Groningen), the SSEH meeting in 2022 (Gothenburg), the LSE Economic History Graduate Seminar series in 2022, the EHS conference in 2023, the End-of-Semester workshop of the Department of Economics 2023 (University of Copenhagen), the Annual CITP conference (Nottingham), the Danish Society of Transport Economics and multiple times at the PhD seminar series at the Department of Economics (University of Southern Denmark). I am grateful for the insightful comments and suggestions by participants at these events. This paper also benefitted from improvements suggested by ChatGPT. Important early inspiration for this paper came from the podcast 'Kongerækken' and Poulsen (2019). Finally, I want to thank my fiancee, Maria Fay Courtney Bohr. On our vacation during the COVID-summer of 2020, we stumbled over some fascinating local history resulting in the present paper. All mistakes are my own. The paper was previously called "A perfect storm and the natural endowments of trade-enabling infrastructure". All code to reproduce all the steps from raw data to the final regressions and the paper itself is available at https://github.com/christianvedels/A_perfect_storm

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

Figure 1: Map of Denmark and the event in 1834



Notes: The effect of the 1825 Agger Isthmus breach on shipping routes. The map shows improved market access (fill color) and Limfjord regions (border color). The arrows indicate shipping routes before (dashed) and after (solid) the breach.
Source: Parish borders from www.digdag.dk

1 Introduction

In the broadest sense, there are two types of long-run explanations for why some places prosper while others struggle: Institutions (Rodrik, Subramanian, & Trebbi, 2004) and geography (Henderson, Squires, Storeygard, & Weil, 2018). A predominant consensus suggests that institutions play a primary role, with geography exerting its influence through its impact on the formation of institutions (Acemoglu, Johnson, & Robinson, 2001; Easterly & Levine, 2003; Ketterer & Rodríguez-Pose, 2018; Rodrik et al., 2004). But even among geographical explanations, the evidence to suggest the importance of path-dependence and second-nature geography (Ager, Eriksson, Hansen, & Lønstrup, 2020; Bleakley & Lin, 2012; Krugman, 1991) and relegates the role of first-nature geographical fundamentals as being merely important as seeds of future growth and development (Allen, Bertazzini, & Heldring, 2023; Bosker & Buringh, 2017; Davis & Weinstein, 2002). This paper demonstrates the importance of first-nature geography in its own right. This is done by using a natural experiment that suddenly changed first-nature geography. This in turn caused an increase in trade, population density, and both

manufacturing and fishing: A large permanent causal effect of first-nature geography on the level and location of economic activity.

In 1825 a violent storm breached the Agger Isthmus in Northwestern Denmark. An erosion process was initiated and in 1834 a navigable channel had formed. The isthmus had previously separated the Limfjord from the North Sea. But now the region was endowed with convenient access to the open ocean and world markets (see Figure 1). Market towns of the West Limfjord had previously been small and isolated. Now they grew to become relatively important ports of international trade. Similarly, a channel in approximately the same location closed between 1085 and 1208. After the last ice age, Northwestern Denmark was left with a convenient westward channel, used by traders and Viking fleets alike. But it would close due to gradual land-rises - an ongoing process since the last ice age (J. T. Christensen, Cedhagen, & Hylleberg, 2004). What followed was a long-run stagnation of the West Limfjord region until the geographic circumstances were reversed in 1825.

This paper provides evidence of the effects of these shocks in a difference-in-difference framework. This is done using historical census data, a large database of Northern European historical trade (the Sound Toll Register), and the densities of archaeological findings. It is shown that the 1825 breach led to an increase in population of 26.7 percent in 1901. To document the mechanisms, I have collected new data on all Danish occupations (13.5 million observations) in the period 1787-1901. The Danish censuses contain a description of each person's occupation. A new machine learning procedure, *OccCANINE*, is applied to these descriptions to convert them into standardized HISCO codes with high accuracy (Dahl, Johansen, & Vedel, 2024; Leeuwen, Maas, & Miles, 2002). The HISCO-coded census data is then used to estimate the effect on occupational structure. The key finding is that the channel caused both fishing and manufacturing to grow. An important question is whether the improved geography caused growth or simply reallocation of prosperity. By using age, gender, and birthplace information it is shown that the population growth came primarily from fertility - not migration. Denmark in the 19th century was in a post-Malthusian regime, where improved living standards cause fertility (Galor, 2011; Jensen, Pedersen, Radu, & Sharp, 2022). Taken together, the evidence of this paper suggests that intrinsic growth is the most likely explanation of the observed effect. That is, improved first-nature geography seemingly causes intrinsic growth - not just a reallocation of growth that would have occurred elsewhere.

For the 11th-century closing, the analysis relies on archaeological data and the closing of the previous channel as an exogenous shock. The time variation comes from the dating of archaeological sites. The geolocation of sites containing coins and buildings is used to construct panels of the probable location of economic activity. A branch within archaeology, Settlement Scaling Theory (Ortman & Lobo, 2020), argues for such approaches to infer historical economic activity, and archaeological evidence has also seen recent application within economics (Allen et al., 2023; Bakker, Maurer, Pischke, & Rauch, 2021; Barjamovic, Chaney, Coşar, & Hortaçsu, 2019; Davis & Weinstein, 2002). This paper presents a novel approach to estimating the location of economic activity using archaeological data. The method involves estimating the probability that archaeological sites containing coins or buildings were active within specific time periods. The findings indicate that a one percent decrease in market access results

in a 13 percent reduction in the probability of economic activity by this measure. Denmark of the 11th century is very different from its 19th-century counterpart. But despite differences in technology, culture, institutions, religion, and most other aspects thought important for economic development, geomorphology played a similar role in the 12th century as it did in the 19th century, providing evidence of (temporal) external validity of the effect of first-nature geography on prosperity.

A separate but related contribution of this paper is in understanding the benefits of infrastructure. Some of the most expensive and extensive projects that a society can undertake are infrastructure projects. Infrastructure changes distances between people. Infrastructure alters first-nature geography. However, infrastructure investments are often endogenous to outcomes of interest. If a dramatic increase in e.g. population density is observed following infrastructure, it is difficult to identify whether this is because it was just constructed in an already thriving place. There are different standardized identification strategies commonly applied in the literature (Redding & Turner, 2015). This paper can be seen as an entirely new way of studying infrastructure change by relying on changes in the natural endowments of it.

Although events like the ones studied in this paper are rare, there are still plenty of opportunities to investigate similar phenomena in the vast expanse of history. For instance, both the histories of Bruges and Königsberg/Kaliningrad are marked by the consequences of an unstable channel (Britannica, 2018; Charlier, 2011; Houtte, 1966). And climate change threatens to cause many similar events as a consequence of natural disasters (Portner, Roberts, Trisos, Simpson, & Möller, 2022) and the melting ice might cause a new northern sea route (Bekkers, Francois, & Rojas-Romagosa, 2018). The approach presented in this paper offers a feasible method to extract new knowledge from these events and many others like them. This yields and enables new insights into both current and historical developments.

This paper proceeds as follows. Section 2 outlines this paper's relation to the literature. Section 3 presents the historical background. Section 4 presents the overall empirical strategy. Section 5 presents a general overview of the data used. The remaining sections demonstrate empirical results. Section 6 demonstrates how the shock caused an increase in trade. Section 7 shows how the channel caused the population to grow. Section 8 covers mechanisms. Section 9 demonstrates the effect of the closed-up channel in the 1100s. Section 10 concludes.

2 Literature

This paper enters a debate about the fundamental determinants of prosperity. Is geography or are institutions the fundamental cause? Acemoglu et al. (2001) argue that extractive institutions set up by colonizing Europeans are partially to blame for the current low-income status of many African nations. Rodrik et al. (2004) argue that geography plays a limited role if institutions are kept constant. Henderson et al. (2018) use satellite image data to show that geography through trade and agriculture can explain 47 percent of the worldwide variation in economic activity. Ketterer and Rodríguez-Pose (2018) find that first nature does play a role but institutions are far more important even when a comparison is made within European regions. It would be natural to get the impression, that

geography is simply unimportant in its own right. What I demonstrate in this paper is the contrary.

Before introducing this in more detail, it is useful to provide some clarity regarding some terms already used and provide a simple taxonomy of geography and economics. Cronon (1992) introduced the distinction between first- and second-nature geography, where first-nature geography is everything determined independently of humans, and second-nature geography is all of the geography caused by humans and their interactions. This has since become the standard framework for understanding the interplay of economics and geography in broad terms (Caruana-Galizia, Okubo, & Wolf, 2021). Among first-nature geography, we can distinguish between:

1. *Things in/on the soil/land*: Soiltype, natural resources, rainfall, etc.
2. *Geomorphology*:[1] Waterways, lakes, mountains, etc.

A huge literature is dedicated to second-nature geography in terms of infrastructure (covered below). Among first-nature explanations of the 'what's in the soil'-type, Andersen, Jensen, and Skovsgaard (2016); Dalgaard, Knudsen, and Selaya (2020); Nunn and Qian (2011) demonstrate that variation in the productivity for food (the potato, the heavy plow, and fish stock respectively) of a particular place caused increased population density. But 'what's in/on the soil'-first-nature also affects other outcomes. Decet and Marcucci (2023) show how a lack of rainfall is associated with an increased level of conflict in modern-day Africa. Boberg-Fazlić, Lampe, Martinelli Lasheras, and Sharp (2022) demonstrates how land productivity caused inequality, and Clark and Gray (2014) demonstrate the influence on literacy. Of course, the ground might be endowed with other determinants of growth like coal (Fernihough & O'Rourke, 2020) or peat (Henriques, Sharp, Tsoukli, & Vedel, 2021).

On the other hand, first-nature *geomorphology* explanations are more sparsely covered with Allen et al. (2023); Diamond (1997); Matranga (2024) - all related to the neolithic revolution and the emergence of the first civilizations - as the most notable exceptions. Furthermore Henderson et al. (2018) and Bosker and Buringh (2017) demonstrate that *geomorphology* is correlated with prosperity. Of course, this is often based on the comparison between places with and without certain geomorphological features, which makes the causality contingent on the degree to which you can trust their method for controlling other differences between these places. Achieving a clean identification remains a formidable challenge. Consequently, a survey by Bosker (2022) calls for the exploration of natural experiments to gain insights into the causal effects of first-nature *geomorphology* on population growth. The present paper tries to fill that gap.

All of this is related to a central result by Davis and Weinstein (2002): Even very large disruptions of city populations (the bombs over Hiroshima and Nagasaki at the end of WWII) do not change the long-run outcomes. Brakman, Garretsen, and Schramm (2004) replicate a similar result for the bombing campaigns in Germany. This result has two types of potentially overlapping explanations: It is either because *first-nature geography* is the *fundamental* determinant of city population for a

[1] Also often called 'physical geography'. But this term is confusing. A bridge is very physical. But it is not what is meant by 'physical geography'. A land-bridge on the other hand is 'physical geography'. Using the term Geomorphology circumvents this confusion.

given state of the economy, or it is because of path dependence.[2] In both cases, city sizes will tend to gravitate towards their long-run trend despite disruptions. But it is hard to disentangle the causes. However, if the location of prosperity is fundamentally determined by first-nature geography, then a shock to this should have a permanent rather than temporary effect. This is what I demonstrate: First-nature geography does indeed have a permanent impact on the location of prosperity.

That is not to say, that path dependence does not play a role. Bleakley and Lin (2012) convincingly document a case of path dependence. Places that used to be in favorable geography continued to be urban centers even when the geographical features that made them so became irrelevant. An implication of Krugman (1991) is that small initial first-nature geographical differences can accumulate into very different outcomes via second-nature geographical interactions (Caruana-Galizia et al., 2021), which is consistent with the observation of Bleakley and Lin (2012). The present paper documents fairly large and permanent effects within a generation when first nature changes, which firmly demonstrates an important role of the *first-nature fundamentals* explanation of the location of economic activity. When favorable geomorphology in the Limfjord region deteriorated, so did the local economy, and when it reverted 700 years later, the economy (and population) increased dramatically. It is not *just* path dependence.[3]

The central natural experiment used in this paper is related to Seror (2020). He uses the randomly altered courses of the Yellow River in China to measure the effects of waterways on population growth. The paper finds that market access led to increased population, increased taxes, and in turn reallocation of elites - presumably to avoid the taxes. Allen et al. (2023) use river shifts and archaeological data in what is today Iraq to demonstrate that river shifts are associated with state formation. In comparison to these two papers, the relatively rich Danish data and the unique natural experiment, allow us to exploit both the 11th- and 19th-century effects of the same event, and especially the 19th-century data, combined with detailed historical accounts, makes it possible to explore the mechanisms.

Geomorphology affects market access. That is the fundamental mechanism here. As such, we can also compare the present paper to the literature on market access and infrastructure. Infrastructure is second-nature geography, but a special case: It is a deliberate change of the first-nature *geomorphology*. Railways, roads, bridges, canals - even the internet - all change the effective distance between people, which used to be different because of the geomorphology. In general, a large number of papers show how infrastructure shapes everything from land value, health, industrialization, manufacturing and the persistent location of modern-day economic activity (Berger, 2019; Berger & Enflo, 2017; Bogart, Lefors, & Satchell, 2019; Bogart, You, Alvarez-Palau, Satchell, & Shaw-Taylor, 2022; Donaldson & Hornbeck, 2016; Hornbeck & Rotemberg, 2019; Hornung, 2015; Richard A. Hornbeck, 2024; Zimran,

[2]Crafts and Wolf (2014) show a case for this: Cotton mills grew where the first-nature geography, in the form of available water power, was prevalent. When this was replaced with steam power, the location was persistent due to an interaction of first and second-nature geography - some first-nature geography (water access) is useful for both water power and steam-powered mills and the investments in buildings and surrounding infrastructure (second-nature), might as well be made to good use rather than reallocating it.

[3]Two more related publications are Ager et al. (2020) and Cermeño and Enflo (2019). Ager et al. (2020) show how a natural disaster can change relative city sizes. And on the other hand Cermeño and Enflo (2019) show that institutions can create towns in otherwise unfavorable geography.

2020). Of course, a central worry is the endogeneity of infrastructure. Infrastructure is often built because a place is already prospering or the opposite: Because prosperity is desired. In either case, this selection on the expected growth trajectory causes bias. Redding and Turner (2015) classify identification strategies for the effect of infrastructure into three broad classes: Planned routes instruments (Baum-Snow, 2007), historical routes instruments (Duranton & Turner, 2012) or inconsequential units instruments (Berger & Enflo, 2017). The present paper is an innovation in this regard as a *fourth* approach: The emergence of natural infrastructure.

The search for time-variation in effective distances is also motivated by the simultaneity of trade and income. If there is trade, then there is growth (in theory) and if there is growth, then that also promotes trade. As a result, there have been discussions on the relationship between trade and growth at least since Ricardo (1817). This paper builds on a strand of this debate arising from Frankel and Romer (1999). The insight of Frankel and Romer (1999) is that geography (distances specifically) is the source of variation in trade. Using changes in trade predicted by distances can then be used to estimate this otherwise simultaneous relationship. Rodriguez and Rodrik (2001) caution that distance between countries is correlated with other things, which also determine growth, e.g. distance to the equator, institutions, etc. When this is taken into account it is no longer possible to demonstrate an effect. Yet the work of Frankel and Romer (1999) still stands as an important inspiration, which motivates the seemingly paradoxical hunt for time variation in distances. The present paper is one such case of time variation in distances.

Using time-variation in distances is also the approach of Feyrer (2021) who uses the closing of the Suez Canal in 1967-1975 to show that pair-wise distances between countries affect trade and trade affects income. Maurer and Yu (2008) estimate the social savings introduced by the Panama Canal in the tradition of Fogel (1964). Rauch and Maurer (2022) measured the increased market access from the Panama Canal and use it to measure the elasticity of market access to population growth. Bakker et al. (2021) demonstrate that trade-inducing geography even in the time of the Phoenicians seemingly caused prosperity as measured by the density of archaeological findings.

In the interpretation of results, I rely on the theoretical nexus between population growth and economic growth in Malthusian and post-Malthusian economies, which is well accounted for in Unified Growth Theory (Galor, 2005, 2011). Importantly, evidence suggests, that Denmark was in a post-Malthusian regime in the 19th century. This implies, that a positive shock to productivity will cause both raised living standards and higher population density (Jensen et al., 2022; Klemp & Møller, 2016). And this is indeed what the present work demonstrates to be the effect of improved market access.

3 Historical background

A fjord is a navigable inlet of water - typically found in Northern Europe and in particular Scandinavia. Fjords act as important natural infrastructure by connecting otherwise remote regions with convenient waterways. The word 'fjord' takes its origin in the Proto-Indo-European 'pertu-', which evolved into many of the words describing roads, access points, or infrastructure more broadly. E.g. 'pertu-' is

also the origin of 'port', 'per' as in per capita (Harper, 2022). This demonstrates how fjords even in prehistory were perceived as important, primarily in terms of their transportation capabilities. The fjords acted as the Scandinavian highways in a time before modern transportation systems. Looking at a map, it is noticeable that many Scandinavian towns and cities, even today, are located on the shore of a fjord.[4]

The story of the present paper is centered on the so-called Lim-fjord, named after the high concentration of subterranean lime (Meesenburg, 2017). It is located in Northwestern Denmark, more specifically the Northern part of Jutland (the peninsula 'jutting' out of Germany). The region is located conveniently at the heart of Scandinavia, between modern-day Norway and Sweden. Before the 1100s, the Limfjord had both an eastern and a western opening, making it a safer shortcut for ships wanting to avoid the rough sea of Skagerak - the strait between Norway and Denmark. It is no surprise that the Limfjord developed into the main hub for Vikings to gather before sailing west to trade (and pillage) (Rasmussen, 1966). In the center of the Limfjord, the settlement of Aggersborg grew large and would become the location of the largest of a series of so-called ring castles constructed in the 10th century when the Jelling dynasty (of which Harald Bluetooth is probably the most well-known) consolidated power over an emerging Danish state (A. Pedersen & Wilson, 2014). But this would not last: The western opening would close and the local economy would stagnate, as shown by the analysis of the present paper.

The last historical evidence of the use of a western opening was in 1085. King Canute IV gathered his fleet in the Limfjord, to sail west and uphold his 'claim' to the English throne after his great uncle Canute II ('the great') who had been King of the so-called North Sea empire consisting of England, Denmark, Norway and parts of Sweden (Spejlborg, 2012). The sailors and soldiers rebelled against Canute IV, the fleet never left Denmark and he was killed when seeking refuge in Saint Alban's church in the town of Odense in 1085.[5] He was later sainted for this and earned the nickname 'the holy' (Pajung, 2012). This event marks the last attempt at gathering a Viking fleet in Denmark to sail west. Never again would Vikings be able to use the western channel. Never again would Danish Vikings try to conquer England (Roesdahl, 2009).

The next time this channel shows up in history is in the writings of the early historian Saxo Grammaticus (1208). Saxo reports in the past tense that there 'used to' be a convenient Western channel.[6] There is some uncertainty surrounding the timing of the writings of Saxo. But it is known that it was around the year 1200, and the historian, Mortensen (2018), notes that the year 1208 is most likely. Therefore, it can be determined that the western channel probably closed sometime between 1085 and 1208. This approximate timing is supported by geological surveys. By drilling through the layers of soil and deposits, geologists have demonstrated a change in the levels of salinity consistent with a closing channel around this time (J. T. Christensen et al., 2004).

What followed was no less than a complete reorientation of the prosperity of the Limfjord region.

[4]Stockholm, Oslo, Gothenburg, Bergen, Stavanger, Odense, Aalborg and Roskilde just to name a few important cities from around Scandinavia, which fit this description.

[5]The very place where I wrote this paper.

[6]Grammaticus (1208) book XIII, section 5

The Eastern Limfjord – particularly the town of Aalborg - became an important center of the herring trade between Hanseatic Lübeck and the rest of the Baltic (Ørnbjerg, 2022). By the earliest counts of market town populations in 1672, Aalborg had become the largest Danish market town after the capital, Copenhagen. The Western Limfjord towns lagged far behind the Eastern Limfjord (Degn, 1989). The regional socio-economic and legal structure that developed would mirror the geography. The west coast was dangerous, difficult, and costly for ships to approach (Hald, 1833, p. 176). In practice, Western Limfjord market towns mainly accessed the market by traveling eastwards through the difficult-to-navigate waters of Løgstør in the middle Limfjord and then further on through the town of Aalborg (see Figure 1). In several court cases in the 16th and 17th centuries, the Western Limfjord market towns fought (unsuccessfully) for the right to trade independently of the whims of the Aalborg merchants (Balle et al., 1974, pp. 78-89), but the geography made it difficult. Ships would have to pass both Løgstør and Aalborg and only limited trade occurred (Poulsen, 2019).

It is important to note that there was *some* trade to and from the region via the west coast. But this was under direct or indirect Aalborg authority. Only small boats could directly approach Jutland's west coast beach (under perilous danger). But that was enough for a small consistent trade between Norway and the Western Limfjord to be sustained. Aagard (1802), p. 30, reports that Thisted (in the West Limfjord) shipped off 6993 barrels of barley and 6832 barrels of oats in 1800. Of this 31 percent and 47 percent respectively were shipped directly from the dangerous west coast beaches, whereas the rest was shipped via Aalborg. Of this, we can assume that most went to Norway. In return, Norway would then ship timber to the region used for construction (C. Christensen, 1735, p. 234). However, this isolation would not persist.

On the night between February 3rd and 4th, 1825, a storm flooded the narrow strip of land called 'Agger tange' (Agger Isthmus). The Agger Isthmus was located in a remote corner of the Limfjord and was sparsely populated. The few inhabitants (388 in the 1801 census) of the Agger parish were offered money to relocate to less dangerous places to live. Many of them settled further into the Limfjord (Poulsen, 2019, 2022). The more remarkable effect was that on the local geography. The Limfjord and the North Sea were now connected. The geographical morphology had changed fundamentally; Shrubs and other plant material protecting the eroding coast were washed away and a shallow Agger channel had formed.

The main effects of the breach happened in three stages. *First* there was a windfall gain in fish caught within the Limfjord (1825-1828) after which fishing within the Limfjord stagnated, while open ocean fishing grew. *Secondly* the channel became navigable and caused a large influx of trade (1834-). *Thirdly* the influx of trade caused institutions and secondary infrastructure to adapt, which boosted local prosperity even more (1840-).[7]

As for the *first effect*, the breach caused the salinity of the water to change and thus the fish (that otherwise had lived in brackish waters) fled. In this process, they were easy to catch and a few 'golden years' of fishing followed (Poulsen, Holm, & MacKenzie, 2007). This peaked around 1828, and what followed was a period of very low yields. Several species went extinct or near extinct (Poulsen, 2019)

[7]For a thorough historiographical account see Poulsen (2022).

and local fishermen had to adapt to fishing for other species in the open ocean instead of the fjord. Whether the benefit of trade outweighs the loss of food, was a topic of discussion even in contemporary times (P. B. Petersen, 1853, p. 1). However, the results presented in the present work indicate that it was on average a benefit.

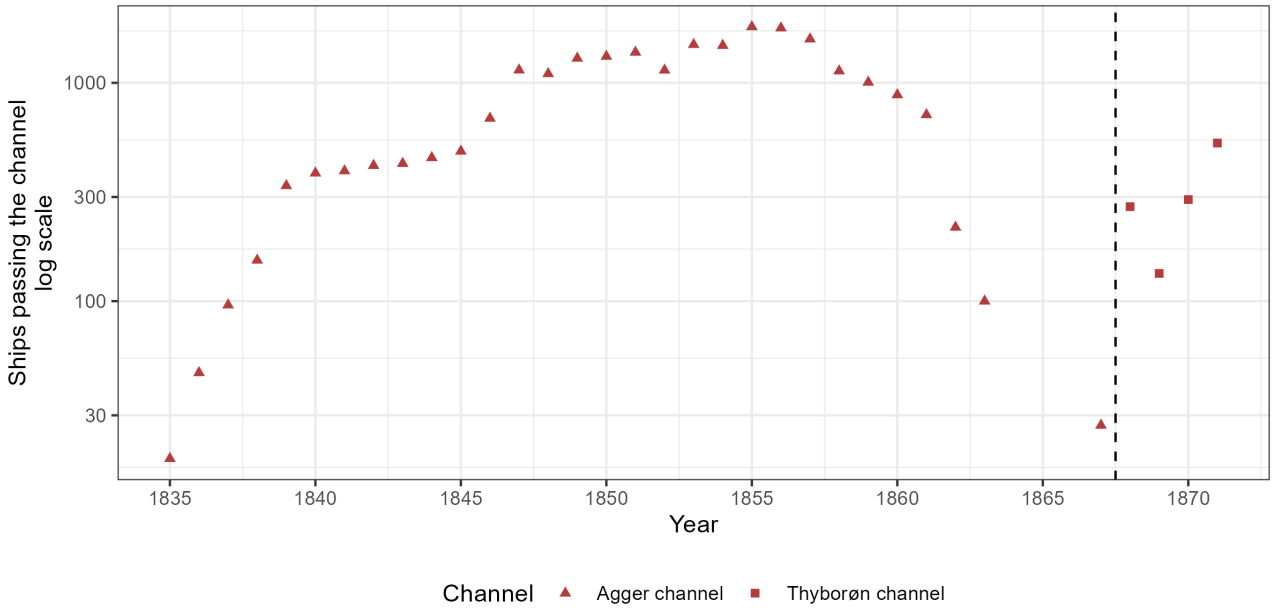
The *second effect* was the increase in trade of the region. In 1834 ships started consistently travelling through the channel. Already in 1831, experiments were conducted. But it would not be until later, in 1834 that ships would pass consistently. The simple before/after impact on trade was large. The data from the sound toll register is used to document this econometrically later in this paper, but for now, it is also useful to note that some indicative direct but scattered quantitative evidence exists for this in the historical sources. Taking the case of the market town of Thisted, (after the channel) ideally located in the West Limfjord: In 1834, 6 ships were registered in the Thisted port, while in 1876 the number was 62. The impact in terms of products traded was also large. In 1853 Thisted exported 24735 barrels of barley and 68091 barrels of oats (Balle et al., 1974, p. 153-159). Comparing these figures, to those reported by Aagard (1802) in 1800 (mentioned earlier), this is an increase of 354 and 997 percent respectively. For the entire Limfjord, the effect was also substantial. In 1835, 19 ships used the Agger channel and in 1855 almost 2000 ships used the channel (Svalgaard, 1977). Newspapers report of ships arriving from England from 1834 and forward with things like coal (Amtsavis, 1834; Roskilde Amts og Advertissementstidende, 1836; Viborg Stiftstidende, 1852). This has led historians to conclude that a golden age of trade in the region had begun (Poulsen, 2019; Ravn, 1993). It quickly became the job of someone to count and register ships that passed the channel. Figure 2 shows the remarkable growth in traffic as reported from this (Svalgaard, 1977).

The *third effect* was a process of institutional adaptation and more infrastructure construction to accommodate this new trade-economy under rapid growth. Redding and Turner (2015) argue that an initial infrastructure investment might attract even more infrastructure to accommodate the growth. They argue that this will inflate the initial effect of infrastructure investment and this might be the reason why the empirical papers find a larger effect of infrastructure than theoretical predictions. This is exactly what happened in this case. In 1841, the King granted international trading rights — independent of Aalborg — to all the Limfjord market towns. In practice, this meant customs offices in all these market towns. Previously, only Aalborg had a customs office. We can speculate that this was a matter of practicality given the new geography. With all the trade now coming through the Agger channel, it seemed self-defeating to maintain the administration of all the trade in Aalborg. The traffic itself was also a logistical problem for the ports. The market towns on the Limfjord coast only had very limited ports, and capacity was overwhelmed. This attracted investment in a series of new ports.[8].

Another large infrastructure investment, the Frederik VII canal (finished in 1861), was attracted

[8]Thisted got a port in 1841 (Diørup, 1842, p. 384-386); Struer got a port in 1856 which was further expanded in 1864 (Trap, Falbe-Hansen, Westergaard, & Weitemeyer, 1906, vol V, p. 467); In 1857 a port was constructed in Lemvig (Trap et al., 1906, vol V, p. 474); Nykøbing Mors had a port from 1788, but in 1843 it was purchased by the city to support trade (Trap et al., 1906, vol IV, p. 214); Løgstør had a port from 1820 but it was expanded in 1852 (Trap et al., 1906, vol IV, pp. 399-400)

Figure 2: Number of ships passing the Agger channel



Notes: This shows the number of ships passing the Agger channel and later Thyborøn channel.
Source: Svalgaard (1977). The single observation from 1867 is from Ravn (1993).

because of fears that the Agger channel would silt up. The sand would travel with the currents, and only a true local would know exactly when and where to approach the channel with ships. The depth of the Agger channel varied a lot. In 1845 it was 5 feet deep - in 1846 it was $9\frac{1}{2}$ feet deep (J. C. Petersen, 1877) and the traffic passing would vary in size and count together with this. Around 2000 ships used the channel in 1855. In 1867 it was only 28 (Ravn, 1993). Political pressure arose to permanently improve connectivity and market access within the prospering Limfjord region (P. B. Petersen, 1853). This would eventually involve, coastal protection of the new western waterway. But initially, for the time and the technology, this was most practical to achieve by doing something for the connection between the East and West Limfjord at Løgstør (around the black dot in the middle Limfjord region in the map of Figure 1). At first, attempts were made in 1843 at clearing the silt out of the shallow waters of Løgstør (p. 311 Bergsøe, 1844; P. B. Petersen, 1853, p. 4). This would fail and eventually, a more expensive but feasible solution was found. It was decided to build a canal along the coast near Løgstør to circumvent these waters.

A part of the story of adaptation also revolves around the introduction of steamships. In 1842 a permanent route by steam was established between Aalborg and Copenhagen and soon after the towns of the Limfjord were also interconnected via steamships (Klem, 1967). Steamships would be the vessel for the future important butter trade with origin in Esbjerg (Lampe & Sharp, 2015), but the first attempt at establishing a Danish route - the steamship "Jylland" - took its origin in the Limfjord and the Agger channel (Schovelin, 1891, p. 62). However, by then the Agger channel had become more shallow and this put a natural limit on the size of the ship, which merely had a draught

of five feet (Lassen, 1883). The Agger channel, would eventually completely silt up, but before this happened another channel (approximately in the same location) would open in 1867 (J. C. Petersen, 1877). This channel - called the Thyborøn channel - still exists today.

The problems with the navigability of the West Limfjord channels probably contributed to the choice of Esbjerg as the location of a large port for Danish/British trade - rather than somewhere in the Limfjord. After the success of artificially establishing a safe port in Esbjerg (Lampe & Sharp, 2015), and due to increasing concerns that the Thyborøn channel would also silt up, it was decided to establish a series of groins along the northwestern coast of Denmark in 1875. This final bit of adaptation would protect the area from coastal erosion and secure the existence of a western waterway the Limfjord to this day.

4 Empirical Strategy

This paper identifies three main effects: The effect of the channel on population size, the effect of the channel on trade, and the effect when a similar channel closed in the 1100s on economic activity. However, the empirical strategy is similar across all. The events used here are by themselves isolated exogenous shocks limited to a specific time and a specific region. The effect of the channel can therefore simply be computed as the change in an outcome in the areas affected by the channel compared to the areas not affected by it. This generally takes the form of a classical event-study specification.

$$\log(y_{it}) = \alpha_t + \alpha_i + \sum_{j=1787}^{1901} 1[t = j] Affected_i \beta_j + \varepsilon_{it}. \tag{1}$$

Here y_{it} is the outcome of interest (e.g. population) in a parish i at time t. $t = 1801$ is used as the reference. $1[t = j]$ is an indicator taking the value 1, when $t = j$ and zero otherwise. $Affected_i$ measures how much parish i was affected by the new channel. Therefore β_j represents the time-varying effect of being affected by the channel. That is, how much larger is e.g. log(Population) in affected parishes than in non-affected parishes when the general population trends and initial sizes have been accounted for? Whether a specific parish is affected by the channel is measured in two distinct ways, one is atheoretical and based on a simple dummy, and another is inspired by theory and based on estimates of how the channel affected local market access. On one hand, it is beneficial to have a simple binary measure for locations affected or not affected. On the other hand, it is probably not realistic that the effect is homogeneous across various locations. However, modeling this requires assumptions on the spatial pattern of the effect of the channel. This motivates the use of both specifications.

The dummy measure of being affected is constructed by a simple heuristic of categorizing parishes into different parts of the Limfjord. This is used in the following regression setup:

$$\log(y_{it}) = \alpha_t + \alpha_i + \sum_{j=1787}^{1901} 1[t=j] \times 1[i \in west]\beta_{j,w} + \sum_{j=1787}^{1901} 1[t=j] \times 1[i \in middle]\beta_{j,m} + \sum_{j=1787}^{1901} 1[t=j] \times 1[i \in east]\beta_{j,e} + \varepsilon_{it}. \tag{2}$$

The region directly affected by the channel is the West Limfjord (see historical background) while the Middle and Eastern Limfjord were affected indirectly. For this reason, the parameters of interest are the $\beta_{j,w}$'s, which measures the average change in the outcome in logs - approximate percentages for small values. Parishes are categorized as being 'Limfjord' parishes if they are closer to the Limfjord than any coast. The West and East Limfjord were separated by difficult-to-navigate waters at Løgstør (see historical background). A roughly northwest/southeast line going through Løgstør is defined by the coordinates (57.044185, 9.186837) and (56.958951, 9.275585). This separates the eastern and western Limfjord in a conventional sense. A buffer zone of 20 km around this line is defined as the 'middle' Limfjord. Any parish west of this buffer zone is categorized as 'west' and any parish east of this is categorized as 'east'. Figure 1 illustrates all of this information.

Estimates of Market Access are motivated by the spatial heterogeneous effect of the channel. Some parishes had more benefit than others since they were closer (in cost distance) to the new waterway. The effect of the change in Market Access is estimated and used in the following regression setup:

$$\log(y_{it}) = \alpha_t + \alpha_i + \sum_{j=1787}^{1901} 1[t=j] \times \Delta log(MA_i)\beta_j + \varepsilon_{it}. \tag{3}$$

Here $\Delta log(MA_i) = log(MA_{i,after}) - log(MA_{i,before})$ is the change in log market access as defined below. The parameters of interest are the β_j's, which measure the elasticity of the outcome to market access. Market Access is computed based on the cost distance approach following Rauch and Maurer (2022) and inspired by a vast literature using similar methods (Ahlfeldt, Redding, Sturm, & Wolf, 2015; Donaldson & Hornbeck, 2016; Harris, 1954; Redding & Sturm, 2008). The change in market access comes from the change in the set of effective ports. After the 1834 channel opened, the West and Middle Limfjord became viable ports for international cargo.

The measure of change in market access is inspired by the process described in the historical background. Suddenly the West Limfjord ports became coastal. This change is what needs to be measured. The Sound Toll Register contains ports used for international cargo trade. Most of these are coastal market towns. Most trade went via the coastal market towns simply because it was legally required to do so (Degn, 1989). Goods would be produced in the countryside and carried to the local market town to be sold. Constructing any measure of this process requires a sensible reduction of all cost distances between H ports and P parishes. In this regard, Tobler's first law of geography is useful: *"Everything is related to everything else, but near things are more related than distant things."* - Tobler (1970). Estimating the Market Access is one solution to this problem, which is theoretically motivated (Eaton & Kortum, 2002), follows Tobler's law, and has seen wide empirical use. In this setting, the most

useful property of this approach is that for every location, it returns a single number, which expresses how close that location is to all other locations. Market access is computed as:

$$MA_p = \sum_{h \in H} [CostDist(p, h; \alpha) + 1]^{\theta} \tag{4}$$

Here H is the set of all available ports (H for harbour). h is a specific port location and p is a specific parish centroid. $CostDist(p, h; \alpha)$ is the relative cost of travelling from parish p to port h given the parameter α (defined below).

The core aim of this entire exercise is to gain insights into how this measure changes with the breach. That is, the aim is to estimate two quantities: a) $MA_p|No\,channel$ b) $MA_p|channel$. Quite simply, this comes from the change in the set of available ports $H \rightarrow H^*$. H is all the consistently available ports before 1834. H^* is those consistently available after 1834 as a consequence of the Agger Channel. The estimated conditional market access can then be expressed as

$$\begin{aligned} MA_p|No\ channel &= MA_p|H \\ MA_p|Channel &= MA_p|H^*. \end{aligned} \tag{5}$$

Here H contains all ports observed in the Sound Toll Register more than once but leaving out the ports in the Middle and West Limfjord. H^* includes these. To be accurate, these ports existed both before and after. But before the breach, they saw only limited use. As this paper documents this dramatically changed with the introduction of the channel. What $MA_p|channel$ measures is the market access after the five Limfjord ports become suitable for international cargo because of the channel. θ is the distance elasticity, which in the default specification is assumed to be -1 as in Harris (1954) and Rauch and Maurer (2022). CostDistance is computed based on $\alpha = 10$. That is, it is 10 times as costly to travel over land than sea. Additional details are described in Appendix A. The map in Figure 1 shows the change in Market Access across Denmark given the newly available ports following the breach.

5 Data

For detailed trade data, the analysis relies on the Sound Toll Registers from Elsinore (of Shakespeare's *Hamlet* fame), north of Copenhagen. Denmark taxed ships passing through from 1420 to 1857 and thorough accounts were kept of all ships (1.8 million in total) (Gøbel, 2010). This was digitized by Veluwenkamp and Woude (2009) and their team. This data contains information on the origin and destination of each ship. Conveniently, these ports have even been labeled with coordinates. A total of 126 ports in present-day Denmark are reported in the period 1750-1855. The sum of traffic to and from these ports are used as a measure of the level of trade to and from that destination.

Census data was obtained from Link Lives (Mathiesen, Robinson, Thomsen, & Revuelta-Eugercios, 2022). For the purposes here, it contains individual-level information such as occupation, age, gender,

and the parish of residence. In the main analysis, it is aggregated into parish-level population counts. Data is available for 1787, 1801, 1834, 1840, 1845, 1860, 1880 and 1901. I manually made a crosswalk between the census data parish names and the historical parish borders (which is available in this project's public repository). Only parishes that are consistently observed throughout all the censuses were used. This leaves 1589 of 1783 parishes that were used in the analysis.

This census data is also used to test plausible mechanisms. This is done using age, place of birth, gender, and occupation. All of the occupational descriptions in the census were transcribed into HISCO codes using *OccCANINE* described by Dahl et al. (2024). The HISCO-coded census data is publicly available from Vedel, Dahl, and Johansen (2024). The census data contains an occupational description for each person such as "he is a fisherman and a farmer". For the years 1787, 1801, 1845 and 1880 these already had HISCO codes assigned to them (Clausen, 2015). This was part of the training data for *OccCANINE*. 200 random HISCO codes from the algorithm were manually and diligently checked to be 94 percent accurate.

Table 1 contains summary statistics for the main variables used in the analysis. 'Population' is the number of people in each parish for that year. 'HISCO Agricultural' is the count of all individuals with HISCO codes starting with 6. 'HISCO Manufacturing' is the count of all those individuals with HISCO codes starting with 7, 8, or 9. 'Born in different county' counts the number of people who were born in a different county, than where they then lived later. 'Child-women ratio' is the ratio of children aged between 1 and 5 to women aged between 15 and 45, which is used as an indicator of fertility. Figure 3 shows the standardized distribution of all of these before the event with the obvious exception of 'Born in different county' which is only observed from 1845. From this figure, it should be noticed that there is a large overlap between variables of interest before the Agger Istmus breach. But the pre-event difference is also noticeable. This is addressed in Section 7 and Appendix D.3.

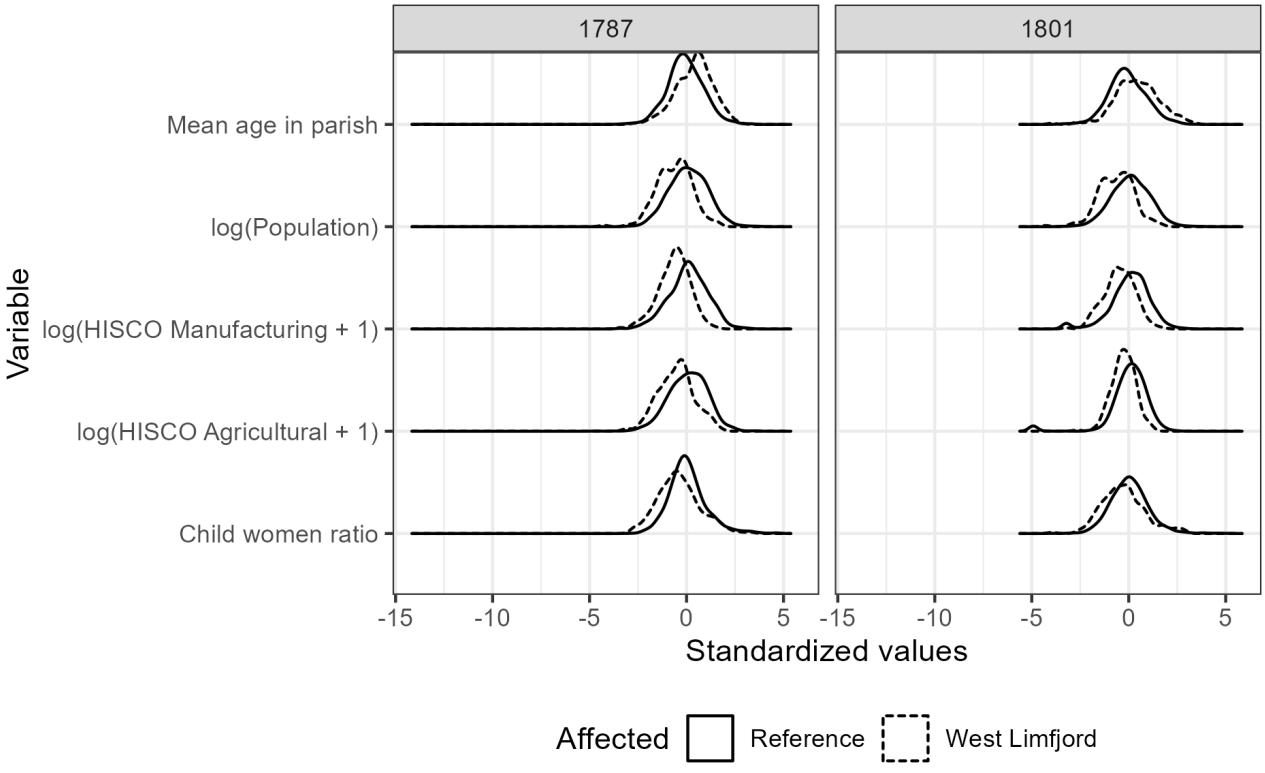
Archaeological data is obtained from a public registry of archaeological findings. All Danish archaeological sites are registered in a publicly available database. The data is geo-referenced, categorized by type, and dated to a certain interval of years. Every coordinate was matched to a parish using the parish borders described above. The archaeological dating of the site was then used to construct a panel of economic activity. Details of this are outlined in section 9.

Table 1: Summary statistics for parish level census data

	Observations	mean	sd	min	Q0.25	median	Q0.75	max
Population	14301	652.02	505.88	27.00	332.00	526.00	816.00	13087.00
Affected: West Limfjord	14301	0.12	0.33	0.00	0.00	0.00	0.00	1.00
Affected: $\Delta log(MA_i)$	14301	0.07	0.06	0.01	0.02	0.05	0.09	0.27
HISCO Agricultural	14301	181.07	120.03	0.00	100.00	152.00	230.00	1512.00
HISCO Manufacturing	14301	65.65	74.77	0.00	23.00	46.00	84.00	2809.00
Born in different county	7945	253.13	378.91	0.00	19.00	72.00	360.00	4299.00
Child-women ratio	14274	0.47	0.11	0.00	0.40	0.46	0.53	1.36

Notes: This table contains summary statistics for the variables used, which are ultimately sourced in the census data. For most variables, there are 14301 observations corresponding to 1589 parishes and 9 census years. The child-women ratio is the ratio of children age 1-5 to women of age 15-45. 27 observations are missing from this variable since no women in the the relevant age group are observed. 'Born in different county' is only observed from 1845 and forward.
Source: Danish census data

Figure 3: Variable distributions



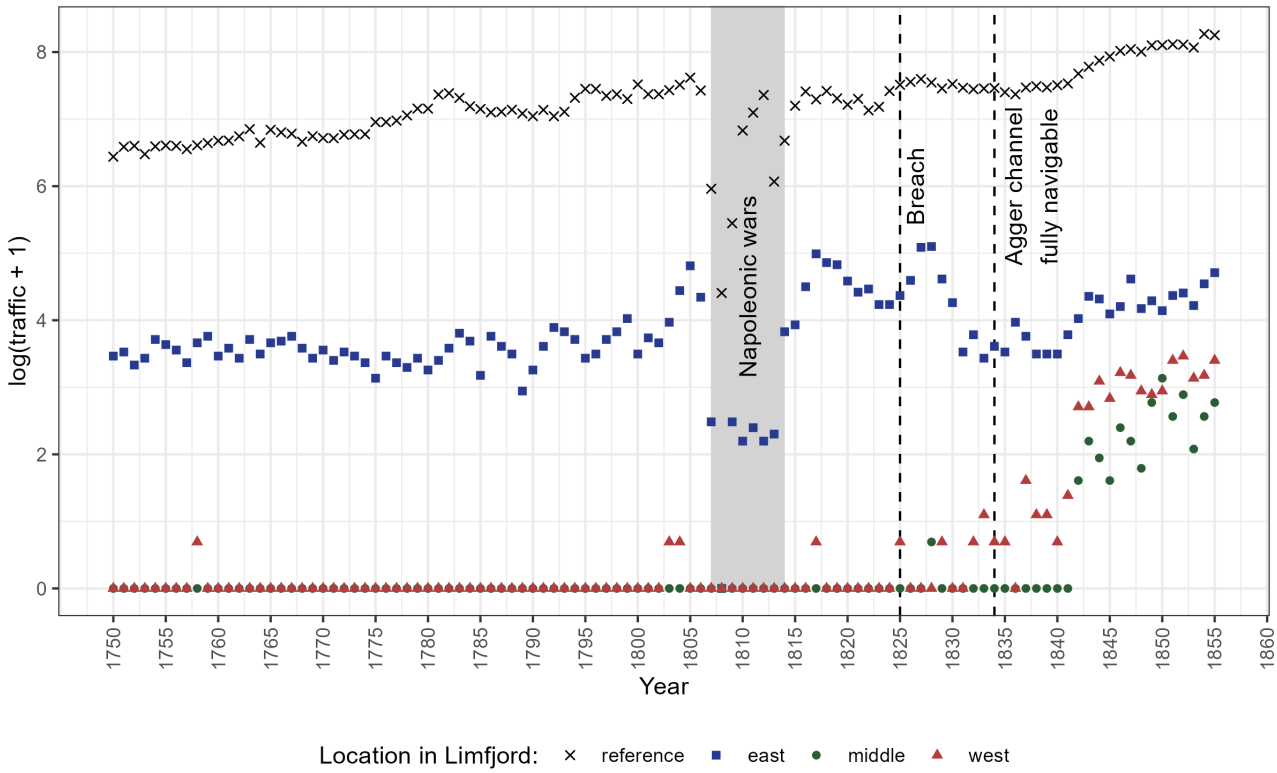
Notes: This shows the distribution of variables of interest before the 1825 breach in the West Limfjord and the rest of the country (excluding other Limfjord parishes). All these characteristics are broadly overlapping approximately similar, with a general trend that the West Limfjord is behind the rest of the country regarding population density and fertility.
Source: Danish census data

6 The effect on trade

After the Agger channel opened in 1834, there was a substantial increase in trade in the region. Figure 4 shows the detailed statistics from the Sound Toll Database. This illustrates the sum of traffic to and from ports on a log scale. An important caveat is that the Sound Toll Database only counts ships that would go eastwards and pass the Sound Toll at Elsinore. However, it should be noted how practically no traffic with origin or destination in the West Limfjord is replaced by a substantial amount afterwards. And this is no quirk of uncertain data. In fact, the amount of trade is relatively stable. This stability was only briefly interrupted when Denmark actively entered the Napoleonic War in 1807-1814 (Feldbæk, 2015).

Figure 4 (and Figure 2 from Section 3) unmistakably highlights the introduction of trade to the Middle and West Limfjord following the creation of the Agger channel. While there is a certain level of traffic prior to 1834, it lacks consistency, aligning with historical sources that confirm the channel becoming fully navigable in that year. Unsurprisingly, a regression version of this (which can be found in Appendix B) shows the same result across many common specifications for this type of estimation. The analysis reveals a notable immediate, lasting surge in trade, which - according to the historical descriptions - would have placed a strain on local infrastructure, prompting further investment. First nature caused trade.

Figure 4: Number of ships to different regions of the Limfjord



Notes: This shows the log-transformed sum of traffic to and from ports in Denmark as captured by those ships that passed Elsinore.

Source: Sound Toll Register Online.

7 The effect to population

The improved first-nature geography also caused the population to increase. Figure 5 shows regression results from estimating equation 1. Panel (a) shows the simplest estimate, where being affected by the Agger channel is defined as a dummy for being in the West Limfjord. This allows parameters to be interpreted as the approximate percentage increase in the affected region caused by the channel. By 1901 the parishes of the West Limfjord region had a 23.6 log-points (26.7 percent) larger population than what would have been expected by parishes from elsewhere and the population in 1801. Panel (b) shows the results when 'affected' is defined as estimated improved market access from the channel. This shows that by 1901, affected areas had experienced a population growth equivalent to an elasticity of market access of 1.59.[9]

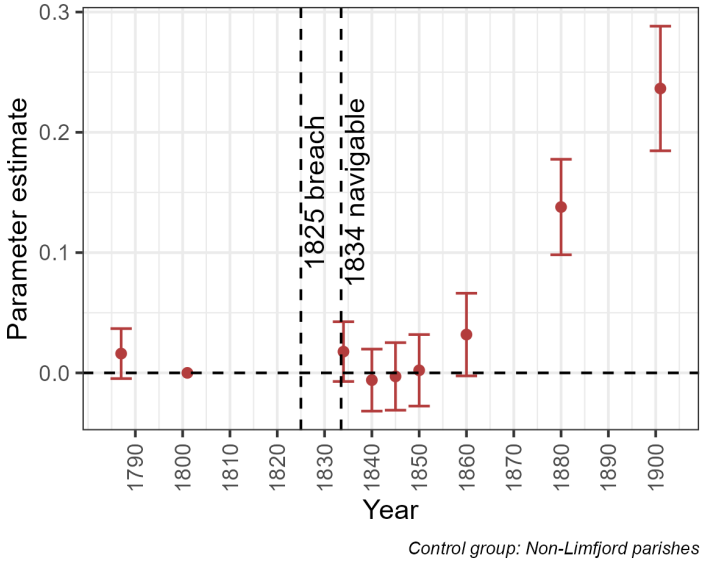
The relative delay of the effect should be noted. It is remarkable that a relatively large increase in ship traffic is not associated with any discernible effect on population size immediately. It is only by 1860, that a positive effect on the size of the affected population is observed. And for the market access approach, it even seems like there is a slight negative effect immediately after the channel became navigable. But a lot happens behind the scenes. This is addressed in Section 8.

From Figure 5 it can be observed that there are no pre-trends in 1787. It is also possible to robustify the results against the choice of comparison group by running the same regression using a set of alternative feasible subgroups of parishes. Figure A2 in Appendix C shows exactly this. It shows the multiverse of all feasible choices of subgroups (Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016) for which the analysis could have been carried out, and the last panel shows all the feasible parameters for market access computation, which could have been assumed. Alternative parameters for the market access function or alternative comparison groups would all have generated positive results, leading to the same qualitative conclusion. As for pre-trends, Appendix C also contains results for the parameter for 1787. None of the alternative specification show any pretrends. Another potential worry is balance. This is addressed by correcting for pre-treatment characteristics (age, occupation, fertility) using the doubly-robust difference-in-difference estimator suggested by Callaway and Sant'Anna (2021). The results of this are also shown in Appendix C. The effect of the channel is robust to all of this.

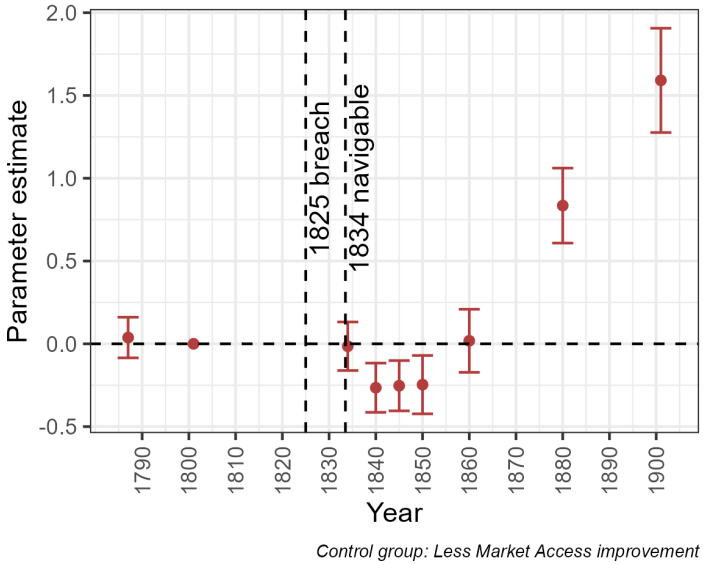
[9] This is similar to the magnitude found for the Panama Canal by Rauch and Maurer (2022).

Figure 5: Effect of the Agger channel on population size

(a) Dummy approach



(b) Market access approach



Notes: Effect of the channel on the size of the population in the affected area. The error bars represent 95 percent confidence intervals based on cluster-robust standard errors. In the first panel, being affected is defined as a dummy for being in the West Limfjord and the second panel it is defined on the estimated improved market access from the channel. 1801 is zero because it is the reference year for both regressions. All parameter estimates can be found in appendix C.1. *Source: Danish census data*

8 Mechanisms

For population growth to occur, the people must come from somewhere. This section tests whether the observed population growth plausibly relates to higher levels of prosperity. This is done to corroborate the connection between the shock to first-nature geography and economic development. It is demonstrated that both the manufacturing and fishing industries grew, indicating that the West Limfjord fundamentally became more prosperous. Furthermore, consistent with the post-Malthusian regime of the time (Jensen et al., 2022; Klemp & Møller, 2016), evidence is shown that population growth is explained by fertility rather than internal migration. This also indicates that the population growth was intrinsic rather than reallocated from elsewhere.

8.1 Occupations

To test the impact of the occupational channel, the HISCO system's structure is utilized (Leeuwen et al., 2002). The first digit of each HISCO code indicates a major category of which the occupation belongs. To estimate the effect on each category, the following regression setup is used:

$$f(y_{it}) = \alpha_t + \alpha_i + \sum_{j=1787}^{1901} 1[t = j] Affected_i \beta_j + \varepsilon_{it}. \tag{6}$$

Equation 6 is equivalent to equation 1 with the important difference that the outcome is not merely $log(y_{it})$ but a function f. Zero outcomes, which are frequent in this setting, imply a methodological trade-off between measuring the intensive and extensive margins separately or accepting that the measured coefficients are not unit-invariant (J. Chen & Roth, 2023). The function $f()$ addresses this issue in four distinct ways:

1. Extensive: $f(y_{it}) = 1[y_{it} > 0]$
2. Intensive: $f(y_{jt}) = log(y_{jt}|y_{it} > 0)$
3. $f(y_{it}) = log(y_{it} + 1)$
4. $f(y_{it}) = arcsinh(y_{it})$

The first function measures the extensive margin, which indicates the number of parishes that have at least one person with a specific occupation. The second function measures the intensive margin, showing the effect among parishes that already had at least one person with such an occupation. The third and fourth functions are two standard ways of measuring the combined effect, acknowledging that these estimates are not unit-invariant.

Another concern is interpreting whether the estimates are substantial enough to explain the overall increase in population. All the transformations yield coefficients, which should be interpreted as relative effects. However, it is not particularly interesting if, for example, the number of priests

increased by 50 percent from two to three. The primary focus is on whether the occupational structure shifted significantly: whether a large proportion of people adopted the professions for which the effect is tested, rather than just the relative effects. To address this, all estimates are converted to the *Average Partial Effect share* ($APE\,share$), which is the average partial effect (the implied number of people gaining that occupation in an average parish), divided by the average parish size in the West Limfjord in 1901. Intuitively, this answers changes to how common a profession is in a parish's population.[10]

Testing all seven occupational categories using both the market access and the dummy approach, and considering the four specifications of $f()$, and all of the census years involves a total of $9 \times 56 = 504$ separate parameters.[11]. Therefore, only the parameter estimates for 1901 are presented. To address concerns over multiple testing, the Bonferroni correction is applied to the standard error of the 56 coefficients in 1901.

Figure 6 illustrates the Bonferroni corrected results using the Dummy approach, displaying the estimated change in a specific occupational category from 1801 to 1901 in areas affected by the Agger channel as compared to the change in the rest of the country. Each panel represents different approaches to dealing with zero outcomes, and the x-axis corresponds to the major categories in the HISCO system. The plot reveals a significant impact on the intensive margin for two particular categories: Agriculture (6) and manufacturing (7/8/9). It is expected that with overall population growth (driven by fertility, as shown below), more individuals would enter agriculture as a default and common profession. However, these results also demonstrate a large absolute change in the occupational structure due to increased prosperity, leading to the rise of manufacturing, which was previously uncommon. Appendix D.1 includes coefficients from both the dummy and market access approach, while the event study plots from each of these estimates are available in the project's online repository.[12]

Figure 7 further breaks down this effect at a more detailed level, using the first three digits of the HISCO codes for agricultural occupations and the first two digits for manufacturing occupations. Notably, the growing manufacturing occupations include those traditionally associated with industrialization, such as textile production and "workers not elsewhere classified," which encompasses generic factory workers. This suggests a fundamentally increased level of economic activity resulting from the improved first-nature geography. Regarding agricultural occupations, the growth of fishing reflects how the population benefited from the sudden availability of fish in the North Sea.

Please note that due to the large number of regressions conducted, confidence intervals were not even estimated for Figure 7 as they would be uninformative after appropriate corrections. The figures should be understood as descriptive. The development of the fishing industry was gradual while spinning only

[10] Estimated from coefficients based on the formula $APE\,share_j = \overline{Occ_{j,1901}}\beta_j / \overline{Pop_{1901}}$, where $\overline{Occ_{j,1901}}$ is the average number of people with that profession in the West Limfjord in 1901 and $\overline{Pop_{1901}}$ is the average number of people in a West Limfjord parish in 1901.

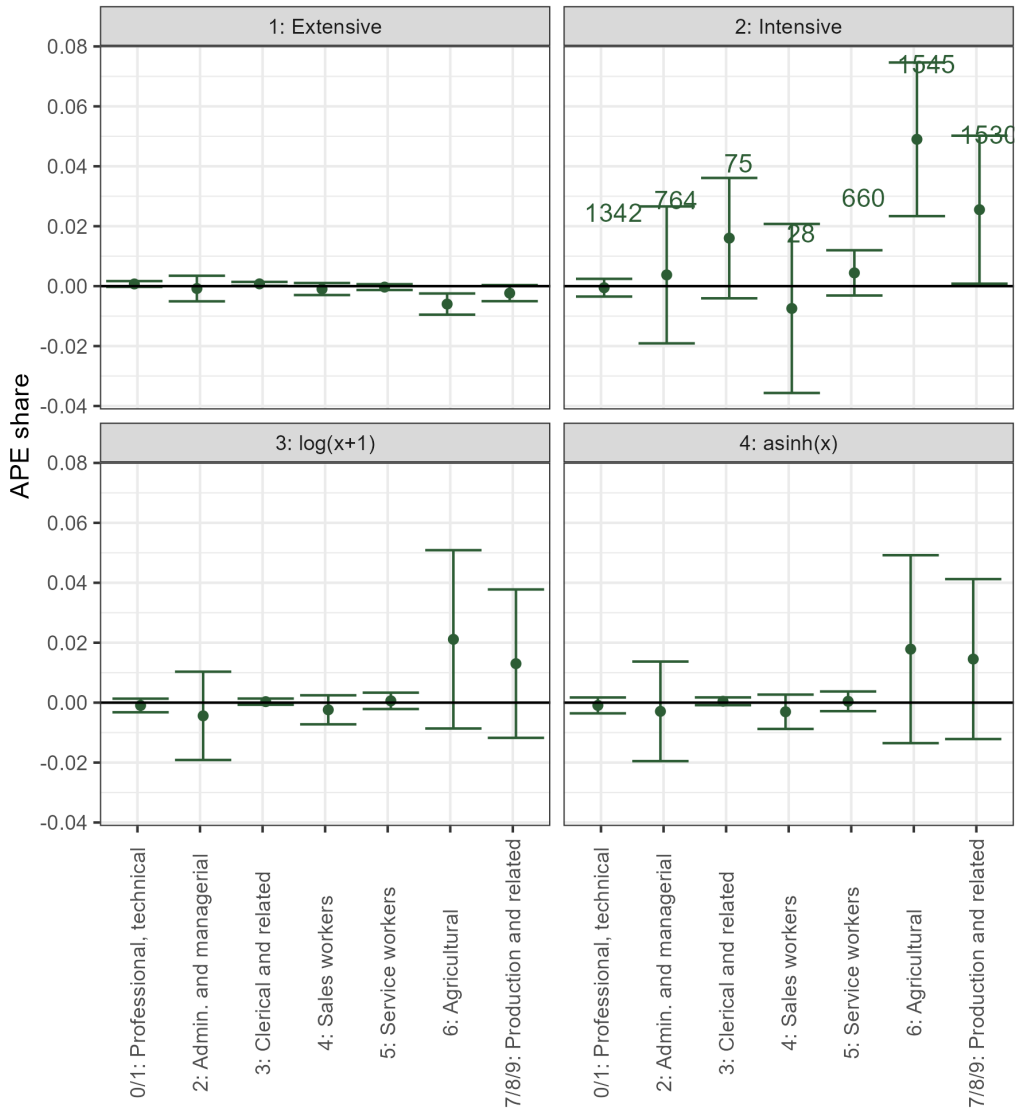
[11] A back-of-the-envelope calculation would predict that 25 of these coefficients would have a p-value below 0.05, purely by random chance

[12] `https://github.com/christianvedels/A_perfect_storm_replication/tree/main/Plots/Mechanism/Occupations`. Note that these plots are not corrected for multiple testing.

took off between 1880 and 1901. This can be seen in event plots of these two occupations, which are included in Appendix D.2. Poulsen et al. (2007) show how an effect of the channel was a collapse of Limfjord fishing due to the increased salinity when the saline Northsea broke through. A hitherto open question is whether open ocean fishing (which the channel enabled) was able to offset this negative effect. This was indeed the case: Fishing approximately became 15-20 percent more common as a profession in the area, laying the grounds for a future industry. Today, Thyborøn, located just at the mouth of the channel, is one of Denmark's main fishing ports, accounting for 25 percent of the Danish catch in 2021.[13] It is noticeable how there is essentially no development in spinning until something rapidly happens between 1880 and 1901, which is timed with the economic take-off in Denmark (Khaustova & Sharp, 2015). Taken together, these results are evidence to suggest that the improved first-nature geography caused increased levels of economic activity as a plausible mechanism for population growth.

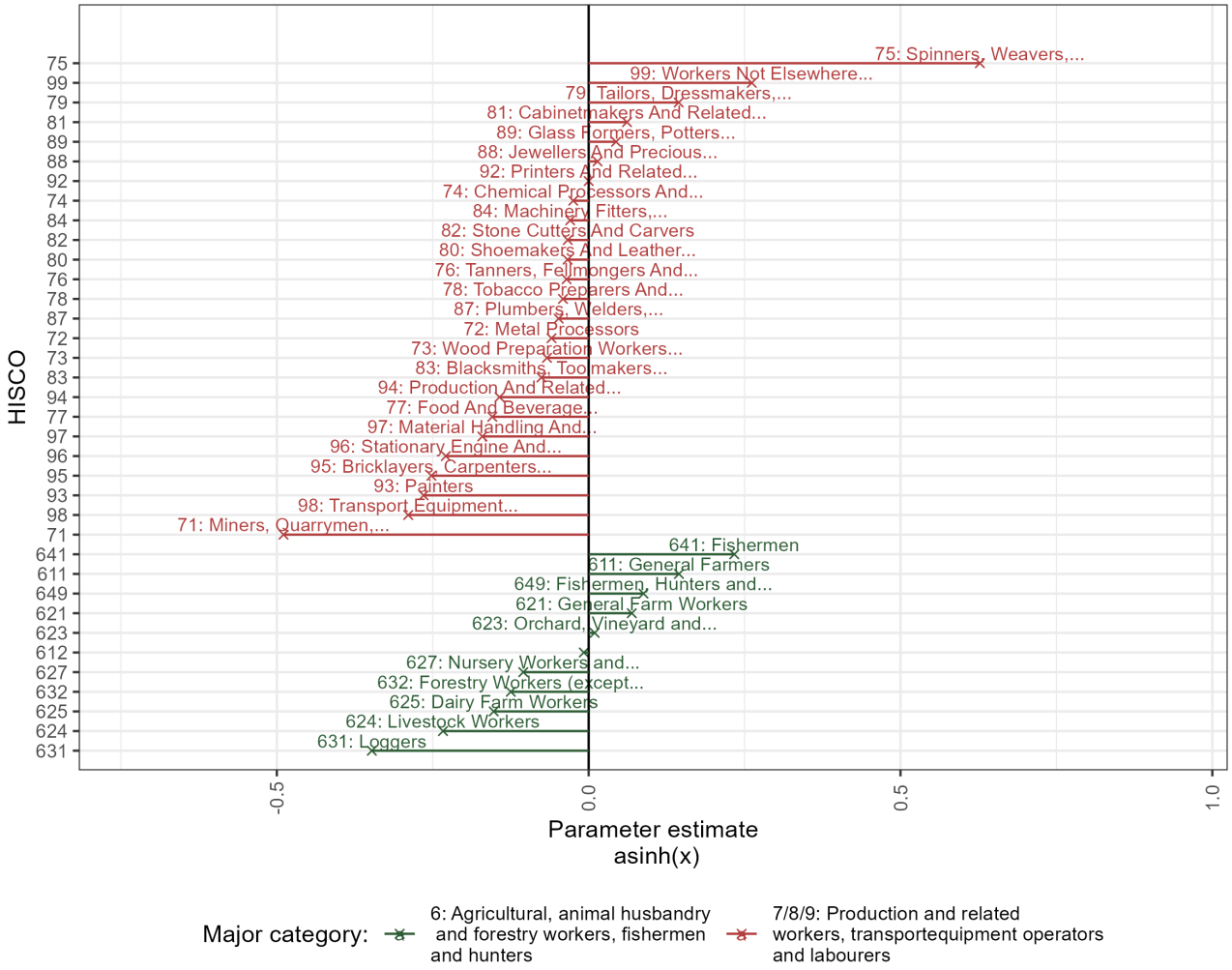
[13]Data extracted from Ministry of Food and Fisheries (2022).

Figure 6: Impact of the Agger Channel on Occupational Structure in 1901



Notes: This figure illustrates the average partial effect (as a share of parish size) on the occupational structure in 1901. The first panel shows the effects on the extensive margin. The second panel shows the effect on the intensive margin (with included number of parishes shown as well). Instead of choosing between intensive and extensive margins, the third panel uses the $log(x+1)$ transformation and the fourth panel uses the inverse hyperbolic sine transformation. The results are based on the Dummy definition of being affected by the channel. Appendix D.1 shows a table of all results including results based on market access. The error bars represent 95 percent confidence intervals corrected for multiple testing using the Bonferroni correction. This is based on standard errors, which are clustered at the parish level.

Figure 7: Effects on Detailed Occupational Structure



Notes: This figure depicts the relative changes in subcategories of agricultural and manufacturing occupations (HISCO codes starting with 6 and 7/8/9). The plot presents the results using the arcsinh transformation and the dummy-definition of being affected by the channel. Qualitatively similar results using alternative approaches are available in the online repository.

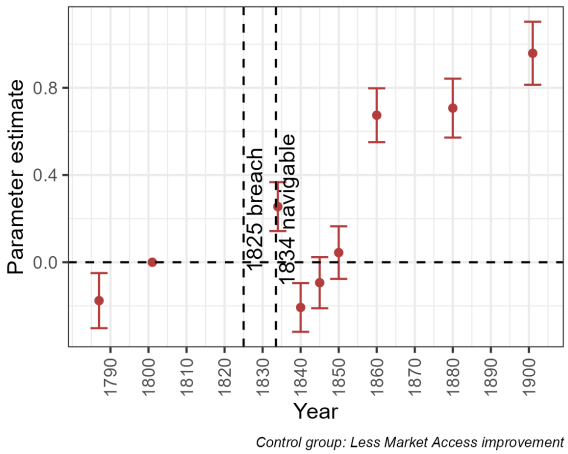
8.2 Fertility and migration

A key question is whether first-nature causes growth or simply reallocates it. Indicative evidence comes from migration rates. If the channel reallocated population from elsewhere, then it should be observed in higher (internal) migration rates to the affected area from elsewhere. On the other hand, if the increase in population is more intrinsic, then given the post-Malthusian dynamics of this time (Jensen et al., 2022; Klemp & Møller, 2016), it is expected that increased living standards would lead to higher fertility.

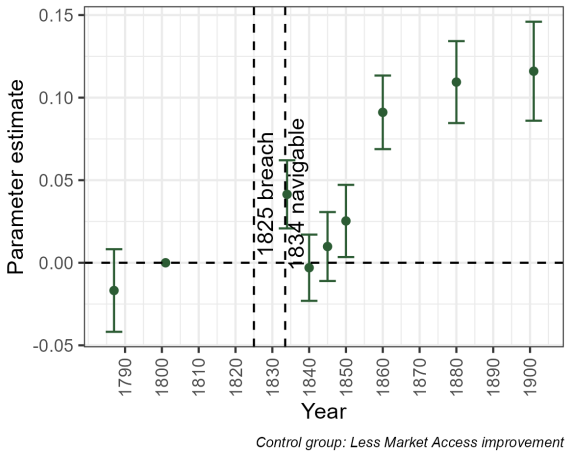
Figure 8 shows event plots for migration and fertility rates. Migration is measured as the number of people born in another county than their current county as a share of the total population in their parish of residence. Fertility is measured as the child-women ratio. Panel (a) and panel (b) show results for fertility. The channel resulted in 11.6 percent higher fertility in 1901 (panel b), which corresponds to 0.96 percent increase for every percent increase in market access (panel b). The results for migration indicate a negative effect or no effect depending on whether the market access approach is used (panel c) or the dummy approach is used (panel d). This indicates that the observed population increase is not the consequence of migration. In fact, there was out-migration in the period in terms of cost distance from the newly available ports, which is what is captured by the market access approach. The dynamics of the effects to fertility correspond to the historical background. The first effect of the channel was a windfall increase in the number of fish caught inside the Limfjord. This corresponds to the increased fertility already in 1834. What followed was a collapse of within-Limfjord fishing, which was observed in the following years until the local population started adapting to open ocean fishing and to the trade opportunities of better market access. After this, the fertility increased to higher levels (Poulsen, 2022; Poulsen et al., 2007). In Appendix D these results are broken down by age groups. The results align with what is presented here. The population composition became comparatively younger as more children were born and/or survived beyond infancy. Again, given Malthusian dynamics, this indicates that population growth came from an intrinsic improvement in the life quality of those who were affected.

Figure 8: Effects on fertility and internal migration

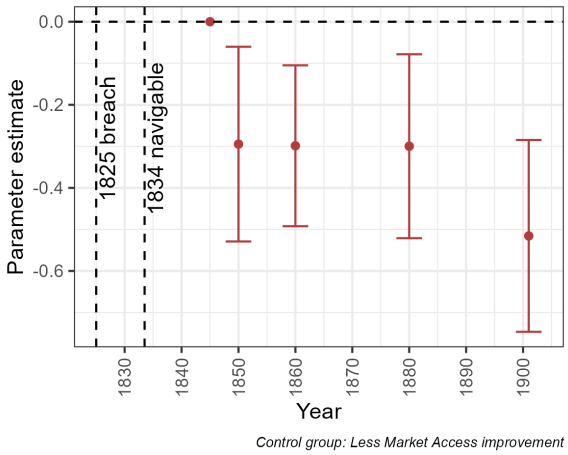
(a) Fertility (MA approach)



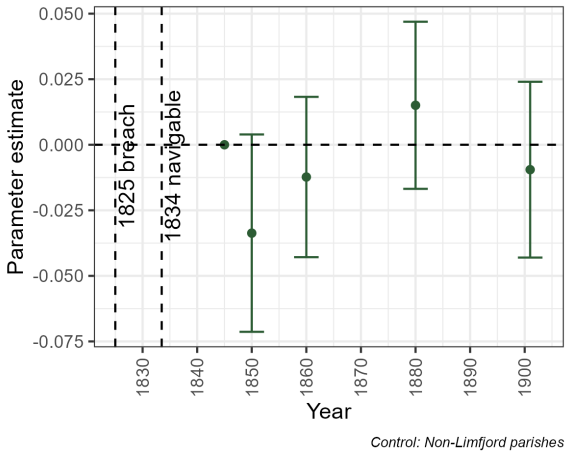
(b) Fertility (dummy approach)



(c) Internal migration (MA approach)



(d) Internal migration (dummy approach)



Notes: This shows the effect on indicators of fertility and migration. Panel (a) and panel (b) show the effect on the child-women ratio. Panels (c) and (d) show the effect on the number of people born in a different county, than where they usually live as a share of the total population in that parish. The results of panels (a) and (b) indicate that the channel caused fertility to increase. The results of panels (c) and (d) indicate, that migration either contributed negatively to the increase in population or did not contribute at all.
Source: Danish census data

9 The reverse natural experiment and external validity

9.1 Method and data

If a causal effect has been identified, one implication is that it should be replicable across heterogeneous environments (Peters, Bühlmann, & Meinshausen, 2016). This is the formal causal reasoning behind the classical arguments for replication studies and concerns of external validity. But it also gives a clue to a way of approaching it. What is truly sought after, when external validity is pursued, is a repetition of the same event in a completely different environment. If the effect is the same, this is evidence, that the effect is not just some quirk of the particular setup. The setting of this particular paper allows me to test exactly this.

The Limfjord region was endowed with favorable first-nature geography after 1834. But this was not the first time the region had been lucky in the geomorphological lottery. The region was endowed with a similar channel in a similar location before the 1100s (see section 3 Historical background). But sometime between 1085 and 1208, the channel closed. In other words, the same natural experiment occurred. Only this time in reverse. Denmark in the 1100s had recently converted from paganism to Christianity, was dominated by Viking traditions and technology, and did not have the benefit of modern science or enlightenment (Lampe & Sharp, 2018; Roesdahl, 2022). Denmark of the 1100s is different from Denmark of 1834 in terms of all the important factors that otherwise impact economic performance in the long run. It is almost like comparing two entirely separate countries. But instead of comparing two countries separated in space, they are separated in time by 700 years. This also implicitly keeps the geography constant with the channel availability as the obvious exception. That is, the only geography that varies in these two cases, is the exact same (mirror image) shock to first-nature geography.

A key problem in this is that the amount of available data from the 1100s is very sparse, to say the least. One available approach is to use archaeological evidence as an indication of economic activity or population density. The idea comes out of the positive archaeological tradition of Settlement Scaling Theory, which simply utilizes that a larger settlement will generate more things to be discovered by archaeologists (Ortman & Lobo, 2020). This approach has also seen application within economics (Allen et al., 2023; Bakker et al., 2021; Barjamovic et al., 2019; Davis & Weinstein, 2002). A methodological innovation of the present paper is to derive and use Bayes' formula to calculate the implied timing and location of economic activity based on the dating and location of archaeological sites.

The Danish Registry of Archaeological Sites 'Fund og Fortidsminder' contains data on all archaeological sites reported by Danish museums and the National Agency of Culture and Palaces. This data is publicly available.[14] The data contains a specific coordinate of the finding, a type (e.g. if it were coins, a former house, burial mount, etc.), and a range of years to which it is dated. The current version of the database contains 290.524 findings in total. But this includes all kinds of sites from the earliest Paleolithic until a few years before the present time. The focus is narrowed a bit on what likely contains useful information in this particular analysis. The data is limited to only include sites

[14] All the data can be downloaded here: https://www.kulturarv.dk/fundogfortidsminder/Download/

that are dated to somewhere between 750 and 1500 CE and only two types of findings: Coins and Buildings. Coins indicate that some amount of trade took place in and around this particular site. Buildings directly reflect population density. More people require more buildings, which increases the probability that remains will be left for archaeologists to discover. The database contains 3411 coin findings and 4396 findings of buildings. These observations are used to construct parish-level panels of economic activity.

An example is pertinent at this point: The finding with id 338 is a finding of coins, which can be dated to somewhere between 1300 and 1535. The site of the finding is at the decimal coordinate (14.92617, 55.02614). This is within the historic parish of Pedersker on the island of Bornholm. This means that some activity that caused coins to be used happened around this location in this time frame. A simple approach would then be to mark Pedersker with a '1', indicating that coins were found here for the years 1300-1535. Applying this principle to all parishes would then allow for the construction of an entire panel of the likely location of economic activity. The descriptive results from this procedure are shown in Figure 10 (details about this in the following). However, this measure causes places with more uncertain dating to be more likely to have economic activity associated with it. This is simply because the coins in Pedersker would show up multiple times, whereas a coin associated with e.g. a single year would only show up once, even though the coin with a smaller range arguably contains more information about the time and location of economic activity. Moreover, this simple method would reflect the inverse of what is of interest and this would make interpretation difficult.

It is useful to think about this in terms of the probabilities the data represents. What is observed is information on the probability of years being associated with a finding: $P(t|finding)$ (e.g. 1300 to 1535 given coin in Pedersker parish). However, what is of interest is the opposite - $P(finding|t)$ - the probability that e.g. 1350 or 1400 is the actual year for which a finding was generated. Moreover, the amount of findings is of little use. It is no surprise that archaeologists find more things in places, where they have already found something. One finding will generate more excavations. For this reason, the intensive margin is not of interest, only the extensive margin is. $P(\{findings\}|t)$ expresses the probability of *any* finding being generated at a specific time. Notice $\{findings\}$ with the curly brackets added. The curly brackets indicate the event that one or more findings are associated with a particular parish. However, this adds further complication to the analytical expression. But it can be derived,[15] the formula can be expressed as

$$P_i(\{findings\}|t) = \left[1 - \prod_{c=1}^{K_i} (1 - P(t|finding_c))\right] P(\{findings\}) \tag{7}$$

where, assuming $P(t|finding)$ is a uniform distribution,

$$P(t|finding_c) = 1[t \in [Y^c_{min}; Y^c_{max}]] \left(\frac{1}{Y^c_{max} - Y^c_{min}}\right). \tag{8}$$

[15]See more details in Appendix E.1.

Here $P_i(\{findings\}|t)$ is the probability of any findings (coins or buildings) having been generated at time t for later discovery in parish i. This is calculated over all the archaeological findings in that parish $c \in [1, K_i]$ and the particular dating distribution of each finding (given by a range of years Y_{min}, Y_{max}). All of this is multiplied by the prior probability that any findings would have been generated.[16] Estimating this probability for every year and for every parish generates a panel of economic activity across all parishes.

This can be calculated by drawing random samples from this distribution. This is quite simply a matter of drawing a random year in the range of years, that a particular finding has been dated to. This is repeated 1000 times. Each time it is checked whether any particular year and parish pair has any finding associated with it in this random draw. These are all draws from the probability distribution of interest and with enough samples an adequate representation of it comes from simply counting the frequency with which draws are successful.

This data enters the general approach presented in section 4. The outcome in this case is the probability that a coin finding was generated in a specific parish in a specific year (or ± 25 within a specific year to be more exact).

$$P_i(Coin|t) = \alpha_t + \alpha_i + \sum_{j=750, j\neq 1000}^{1500} 1[t = j] \times Affected_i \beta_j + \varepsilon_{it} \tag{9}$$

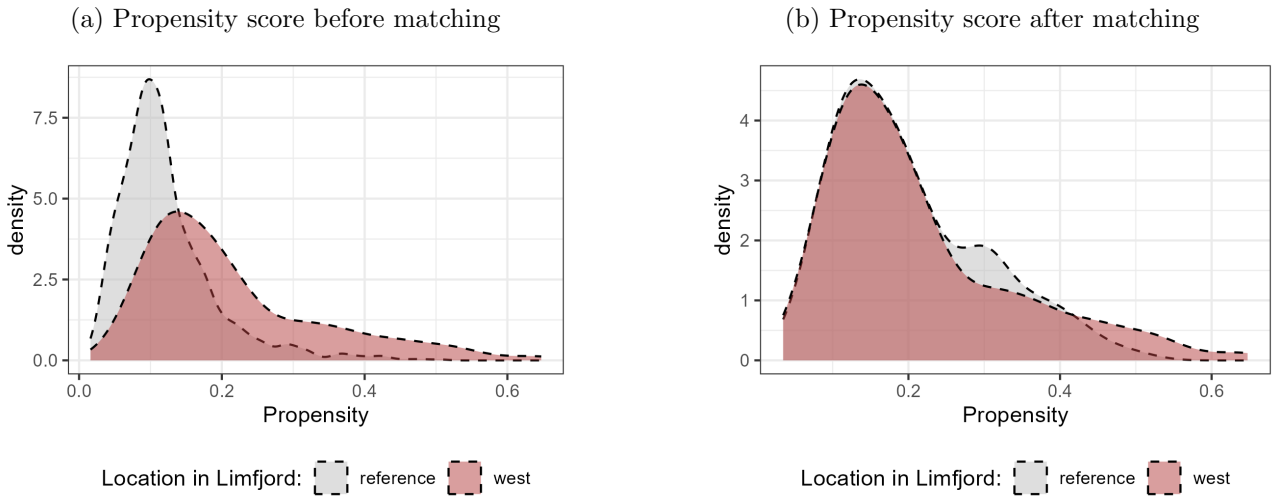
As before, $Affected$ is measured both in terms of a dummy and the estimated change in market access. Inference is made with a clustered bootstrap, which resamples the panel from the 1000 samples drawn from the distribution of equation 7. A further concern is that soil types might cause endogeneity. It might be that places with more fertile soil types had a differential developmental path (Andersen et al., 2016; Boberg-Fazlić et al., 2022). The western part of the Limfjord is rich in fertile soil. At the same time, the soil type might also influence the rate of survival of archaeological sites.

To counter this, propensity score matching is used. The aim of this is to ensure that parishes affected by the channel closing are compared to parishes, which in terms of soil type, are similar parishes to those not affected. The propensity score is estimated using extreme gradient boosting trees (T. Chen & Guestrin, 2016), which is a machine learning technique suitable for classification using tabular data, which can be thought of as a generalization of a random forest (Breiman, 2001). Data on soil types are obtained from S. Pedersen, Hermansen, Nathan, and Tougaard (2019). Only the soil types found in at least 10 percent of parishes are used, otherwise the model overfits by utilizing the fingerprint of allocation of more unique soil types. The procedure estimates a model, which predicts the probability (propensity), that a given parish is a West Limfjord parish based on the soil types in the parish

$$P_i(Affected|\mathbf{X}_i) = f(\mathbf{X}_i) + \varepsilon_i. \tag{10}$$

[16] Whether archaeologists' dating ranges should be interpreted as a uniform distribution is not given. An obvious alternative is to assume a normal distribution, where the uncertainty represents a 95 percent confidence interval from the distribution. Results using this approach are shown in Appendix E.3. These results are qualitatively equivalent.

Figure 9: Soil type propensity scores

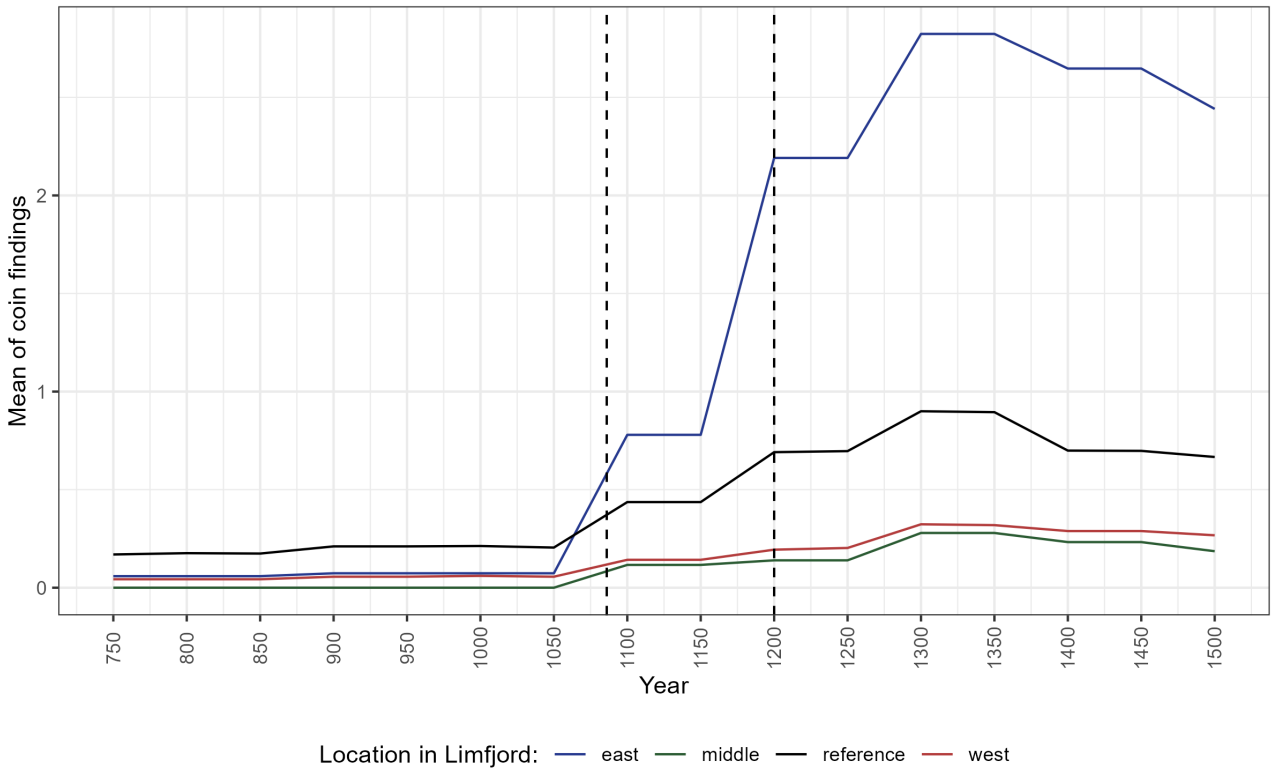


Notes: Propensity scores between the West Limfjord and the reference group before and after matching. Propensity scores are estimated using extreme gradient boost. Matching is done with a *greedy* matching procedure in random order.

Each parish in the West Limfjord is then matched with a counterpart, which has the most similar propensity score, outside the Limfjord region. This represents the match with the most similar soil type. This is done in a *greedy* fashion - without replacement - ensuring that each parish is matched with *one* other parish which at the same time is not the match of any other affected parish. Figure 9 shows the propensity score of parishes included in the sample before and after matching. Notice that after matching, the distributions of propensity scores are almost identical across the West Limfjord and the reference group.

Before showing anything using the estimated probabilities, Figure 10 simply displays the mean number of coin findings in various regions of the Limfjord as well as in the remainder of the country, spanning from 750 to 1500. This is suggestive of the effect of the closing of a channel between 1085 and 1208 and its impact on the West Limfjord region. The plot demonstrates a gradual rise in coin findings, especially in the East Limfjord and the rest of the country, following the channel closure, whereas the west and middle Limfjord regions remain at lower levels. This finding suggests that the west and middle Limfjord regions may have experienced an economic and commercial decline during this period. This reveals that an effect can be observed even with a straightforward approach. However, more uncertain observations have more weight in this very simple descriptive plot. Regression results, that account for this, are presented in Figure 11 and 12 and table 2.

Figure 10: Simple count of coin findings



Notes: This plot shows the average number of coin findings, where coin findings are simply counted if they are potentially attributable to a particular year. This is done for the West, Middle, and East Limfjord region as well as for the rest of the country as reference.
Source: Danish registry of archaeological findings

9.2 Results

Figure 11 and table 2 provide more reliable estimates of the probability of archaeological findings being generated in the affected regions of the Limfjord. Figure 12 shows the parameter distribution in 1350 yielded by the sampling procedure presented. The regression table presents the results of archaeological regressions for the full sample of all of Denmark and a matched sample, using both the dummy definition of being affected and the change in market access approach. The outcome is the probability that a coin finding or building was generated in the area covered by that parish within ± 25 years of the reported year.

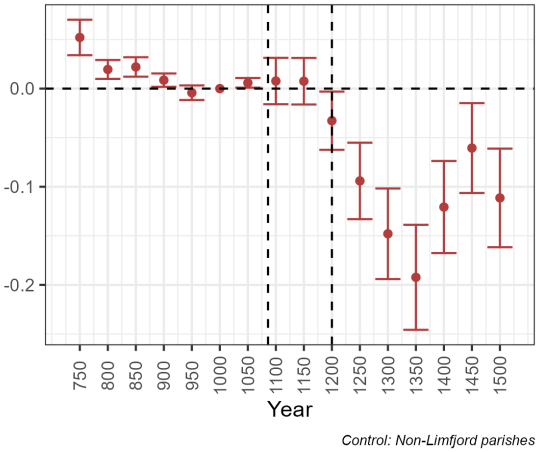
The parameters from the Market Access approach should be interpreted as semi-elasticities, i.e. how does the probability of a coin being generated change with a relative change in market access? These are approximately interpretable as percentage changes. I.e. a 1 percent decrease in Market Access led to a 13-19 percent change in the probability of a coin finding being generated in 1350 - 150 years after the channel closed (columns 1 and 5). Equivalently, it led to a reduction in the chance of buildings being generated as archaeological findings by 6 to 8 percent (columns 3 and 7). The parameters for the dummy approach can be interpreted as the change in the probability that findings were generated from a West Limfjord parish compared to other parishes. In 1350, the chance of coin findings being generated had fallen by around 1.5 or 1.9 percent (columns 2 and 6), and the probability of buildings to become archaeological findings had decreased by 0.6 to 1 percent (columns 4 and 8).

For all these estimates it is seen, that there is a rather stable effect size before 1085-1208 of approximately zero, but after this the probability sharply declines. Appendix E.2. provides tables of all parameters. It is worth considering whether these results can be spurious because of the added layer of uncertainty added by the Monte Carlo method. However, the standard errors used take this added uncertainty into account. These standard errors are constructed by resampling the Monte Carlo samples in a clustered fashion. Figure 12 shows the parameter distribution in 1350 given by the bootstrap samples.

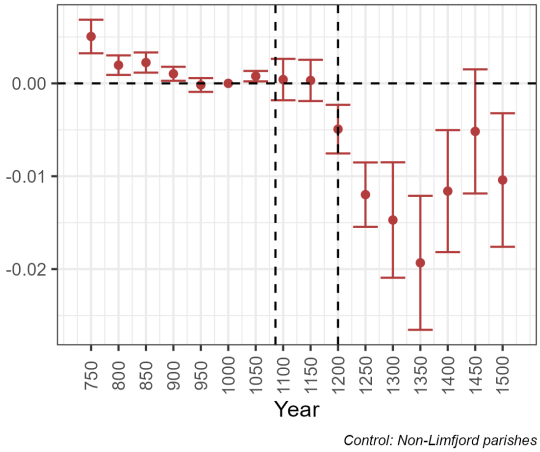
Whichever methodological alternatives are employed, the results show a large negative effect of the closure of the channel for the affected parishes. The results suggest that the closure of the channel between 1085 and 1208 had a significant impact on the West Limfjord regions, causing a decline in the probability of coin findings and buildings being generated in those areas. These findings are consistent with the hypothesis that the channel closure caused economic and commercial decline in the region. It is not possible to do an apples-to-apples comparison to the census data for the Agger channel of 1834. But it is certainly noteworthy that the parameters are of similar sign and magnitudes for these two channels in the same location. This corroborates the external validity of the main result: First-nature geography is the cause of the location of prosperity across very different societies - two very different versions of historical Denmark.

Figure 11: Archaeological results

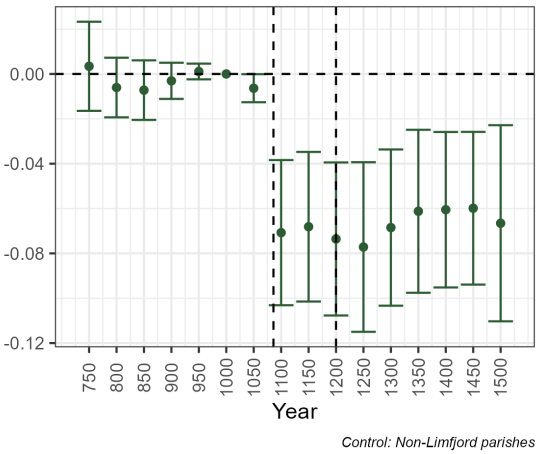
(a) Coins: Market access approach



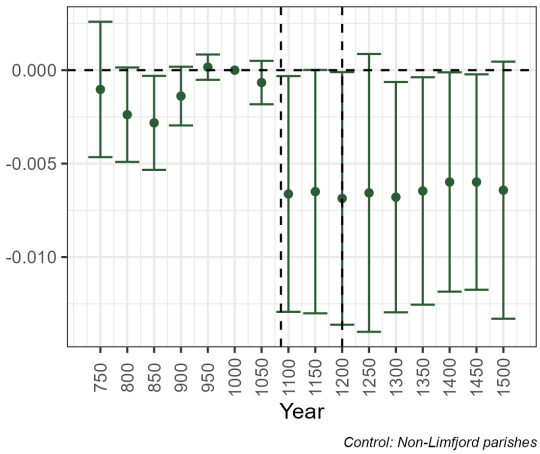
(b) Coins: Dummy approach



(c) Buildings: Market access approach



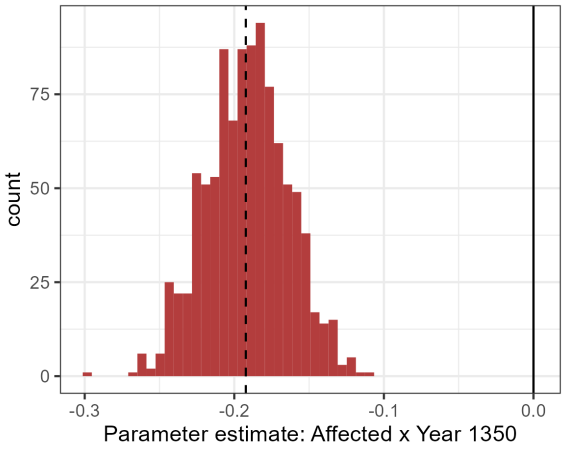
(d) Buildings: Dummy approach



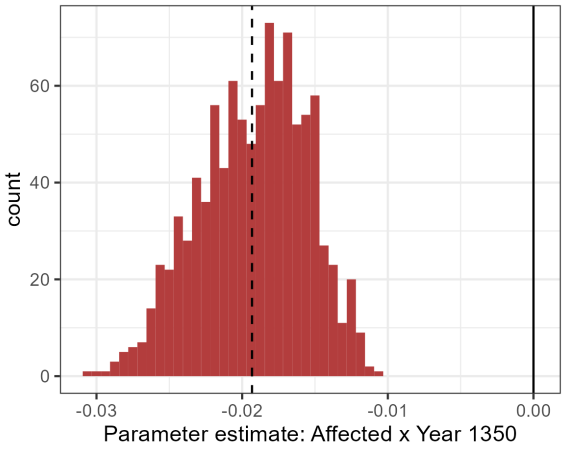
Notes: The effect of the closing of the channel on the probability of archaeological findings. Error bars represent 95 percent confidence intervals based on clustered bootstrap errors. Panel (a) shows results for coin findings relying on the market access approach. Panel (b) also shows results for coins but using the dummy approach. Panel (c) shows results for buildings using the market access approach. Panel (d) shows the results for buildings using the dummy approach.
Source: Danish registry of archaeological findings

Figure 12: Distribution of parameter estimates in 1350

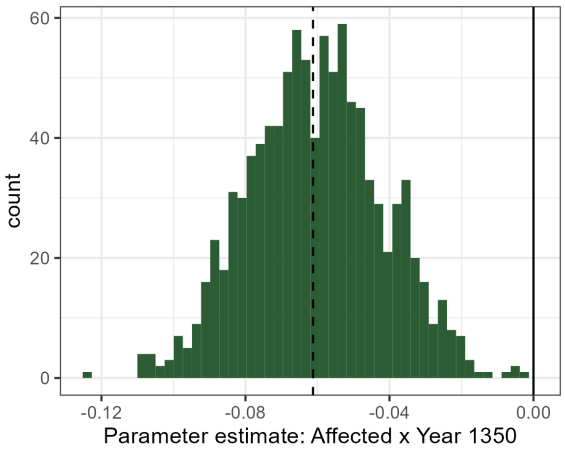
(a) Coins: Market access approach



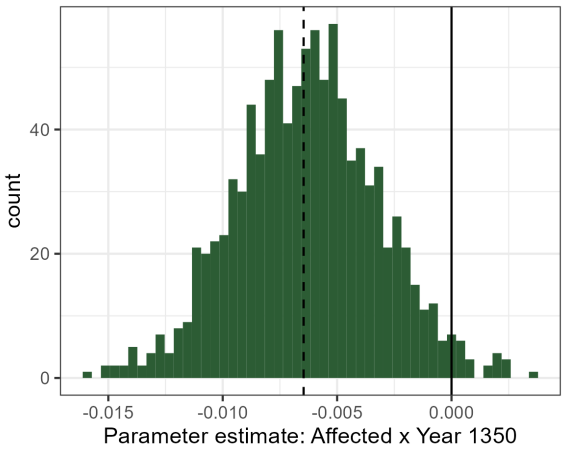
(b) Coins: Dummy approach



(c) Buildings: Market access approach



(d) Buildings: Dummy approach



Notes: This shows the archaeological regression results. Clustered bootstrapped standard errors are shown in the parentheses. Clustered at the parish level and sampled from the Monte Carlo procedure samples. Columns 1-4 show results using the full sample of all of Denmark. Columns 5-8 show results for a matched sample. All the even columns show results using the dummy definition of being affected. All the uneven columns show results using the change in market access approach. The outcome is the probability that a given finding type was generated in the area covered by that parish within ±25 years of the reported year. A parameter estimate for all years 750, 800, ..., 1500 can be found in the Appendix E.2. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Source: Danish registry of archaeological findings

Table 2: Archaeological regression results

	Full sample				Matched sample			
Outcome:	Coin findings		Buildings		Coin findings		Buildings	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	MA	Dummy	MA	Dummy	MA	Dummy	MA	Dummy
Year950 × Affected	-0.0042	-0.0002	0.0011	0.0002	-0.0177*	-0.0021*	0.0008	-8.62×10^{-6}
	(0.0038)	(0.0004)	(0.0018)	(0.0003)	(0.0096)	(0.0011)	(0.0039)	(0.0006)
Year1050 × Affected	0.0059**	0.0008***	-0.0063**	-0.0007	0.0051	0.0006	-0.0083	-0.0012
	(0.0025)	(0.0003)	(0.0032)	(0.0006)	(0.0039)	(0.0005)	(0.0064)	(0.0010)
Year1150 × Affected	0.0075	0.0003	-0.0681***	-0.0065*	0.0320	0.0036	-0.0705**	-0.0078
	(0.0121)	(0.0011)	(0.0170)	(0.0033)	(0.0245)	(0.0029)	(0.0318)	(0.0051)
Year1250 × Affected	-0.0941***	-0.0120***	-0.0772***	-0.0066*	-0.1216***	-0.0150***	-0.0829***	-0.0093*
	(0.0199)	(0.0018)	(0.0193)	(0.0038)	(0.0350)	(0.0045)	(0.0318)	(0.0054)
Year1350 × Affected	-0.1923***	-0.0193***	-0.0612***	-0.0065**	-0.1408***	-0.0159***	-0.0669**	-0.0076
	(0.0273)	(0.0037)	(0.0185)	(0.0031)	(0.0414)	(0.0058)	(0.0297)	(0.0046)
Observations	29,568	29,568	29,568	29,568	7,424	7,424	7,424	7,424
Parishes	1848	1848	1848	1848	464	464	464	464
Parish FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Archaeological regression results. Clustered bootstrapped standard errors are shown in the parentheses. Clustered at the parish level and sampled from the Monte Carlo procedure samples. Columns 1-4 show results using the full sample of all Denmark. Columns 5-8 show results for a matched sample. All the even columns show results using the dummy definition of being affected. All the uneven columns show results using the change in market access approach. The outcome is the probability that a given finding type was generated in the area covered by that parish within ±25 years of the reported year. A parameter estimate for all years 750, 800, ..., 1500 can be found in the Appendix E.2. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Source: Danish registry of archaeological findings

10 Conclusion

First nature in the form of waterways has a key role in determining the location of economic activity by determining market access. We are all prisoners of *geomorphology*, but what happens when the shackles are loosened? The unexpected emergence of the Agger channel in Denmark's West Limfjord region in the early 19th century serves as an important example of the contemporaneous importance of first-nature geography. The channel brought new trade opportunities and revitalized a region that had been lagging behind its neighbors for almost a millennium. This paper demonstrates that geographical fundamentals matter in their own right, suggesting that first-nature geography determines the location of prosperity beyond what is eventually driven by path dependence.

The empirical findings of this study reveal that the sudden emergence of the Agger channel in 1834 had a profound and lasting impact on the economic landscape of Denmark's West Limfjord region. Using a difference-in-differences design, it is shown that the channel's creation led to a 26.7 percent increase in the population of affected parishes by 1901 compared to unaffected areas. This population growth was driven primarily by enhanced market access, which facilitated a significant rise in trade, fishing and manufacturing activities. Moreover, the findings indicate that this growth was intrinsic, with higher fertility rates contributing to the increased population rather than migration. Additionally, a reverse natural experiment from the 12th century, when a similar channel closed, resulted in a substantial decline in economic activity, as evidenced by archaeological data. These results underscore the critical role of first-nature geography in driving economic prosperity.

In a world of climate change and space innovation, the first-nature characteristics of the world around us are changing. The opening of the northwest passage in the Arctic, the colonization of space, and the flooding of large land masses are all possible scenarios that involve a significant change to the first-nature environment in which our economy exists. It is hard to know exactly what might happen, but a part of the answer relies on a better understanding of the universal relationship that humans have with the distances between us. And this is determined by geomorphology. The present work is a contribution to that understanding. In two distinctly different societies - 700 years apart - changing the relative distances between humans with the same channel opening/closing played a similar role.

Events like the Agger channel are rare, but the expanse of history offers many opportunities for similar studies. The breaking of a narrow isthmus in modern-day Belgium in 1134 created market opportunities for Bruges, which declined when the channel silted up in the 14th century (Charlier, 2011; Houtte, 1966). Similarly, the history of Königsberg/Kaliningrad is filled with consequences of the unstable channel around Pillau (Britannica, 2018). These events - and many more - can now also be studied with the approach introduced in this paper.

References

Aagard, K. (1802). *Physisk, oeconomisk og topographisk Beskrivelse over Thye.* Viborg: Author's own publisher. Retrieved from https://bibliotek.slaegt.dk/cgi-bin/koha/opac-detail.pl?biblionumber=28303

Acemoglu, D., Johnson, S., & Robinson, J. A. (2001). The Colonial Origins of Comparative Development: An Empirical Investigation. *American Economic Review*, *91*(5), 1369–1401.

Ager, P., Eriksson, K., Hansen, C. W., & Lønstrup, L. (2020). How the 1906 San Francisco earthquake shaped economic activity in the American West. *Explorations in Economic History*, *77*(October 2019), 101342. Retrieved from https://doi.org/10.1016/j.eeh.2020.101342 doi: 10.1016/j.eeh.2020.101342

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M., & Wolf, N. (2015). The economics of density: Evidence from the berlin wall. *Econometrica*, *83*(6), 2127-2189. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA10876 doi: https://doi.org/10.3982/ECTA10876

Allen, R. C., Bertazzini, M. C., & Heldring, L. (2023, October). The economic origins of government. *American Economic Review*, *113*(10), 2507-45. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20201919 doi: 10.1257/aer.20201919

Amtsavis, T. (1834). *Thisted Amtsavis.* Retrieved from http://hdl.handle.net/109.3.1/uuid:bef130f1-8d10-4151-a8cf-fb6d80caa0e6

Andersen, T. B., Jensen, P. S., & Skovsgaard, C. V. (2016). The heavy plow and the agricultural revolution in Medieval Europe. *Journal of Development Economics*, *118*, 133–149. Retrieved from http://dx.doi.org/10.1016/j.jdeveco.2015.08.006 doi: 10.1016/j.jdeveco.2015.08.006

Bakker, J. D., Maurer, S., Pischke, J.-S., & Rauch, F. (2021). Of Mice and Merchants: Connectedness and the Location of Economic Activity in the Iron Age. *The Review of Economics and Statistics*, *103*(October), 1–14. doi: 10.1162/rest_a_00902

Balle, T., Søndberg, B., Bjerregaard, M., Svalgaard, R., Miltersen, J., & Jørgensen, M. (1974). *Thisted købstads historie - Historisk Årbog for Thy og Mors og Vester Hanherred.* Thisted: Historisk Samfund for Thy og Mors, Thisted Bogtrykkeri. Retrieved from https://slaegtsbibliotek.dk/911456.pdf

Barjamovic, G., Chaney, T., Coşar, K., & Hortaçsu, A. (2019). Trade, Merchants, and the Lost Cities of the Bronze Age*. *The Quarterly Journal of Economics*, *134*(3), 1455–1503. Retrieved from https://doi.org/10.1093/qje/qjz009 doi: 10.1093/qje/qjz009

Baum-Snow, N. (2007). Did Highways Cause Suburbanization? *The Quarterly Journal of Economics*, *122*(2), 775–805. Retrieved from http://www.jstor.org/stable/25098858

Bekkers, E., Francois, J. F., & Rojas-Romagosa, H. (2018). Melting ice caps and the economic impact of opening the northern sea route. *The Economic Journal*, *128*(610), 1095-1127. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12460 doi: https://doi.org/10.1111/ecoj.12460

Berger, T. (2019). Railroads and Rural Industrialization: evidence from a Historical Policy Experiment. *Explorations in Economic History*, *74*(October 2018), 101277. Retrieved from

https://doi.org/10.1016/j.eeh.2019.06.002 doi: 10.1016/j.eeh.2019.06.002

Berger, T., & Enflo, K. (2017). Locomotives of local growth : The short- and long-term impact of railroads in Sweden R. *Journal of Urban Economics*, *98*, 124–138. Retrieved from http://dx.doi.org/10.1016/j.jue.2015.09.001 doi: 10.1016/j.jue.2015.09.001

Bergsøe, A. F. (1844). *Den danske Stats Statistik, vol I.* København: Author.

Bleakley, H., & Lin, J. (2012). Portage and Path Dependence. *The Quarterly Journal of Economics*, *127*(2), 587–644. Retrieved from https://doi.org/10.1093/qje/qjs011 doi: 10.1093/qje/qjs011

Boberg-Fazlić, N., Lampe, M., Martinelli Lasheras, P., & Sharp, P. (2022). Winners and losers from agrarian reform: Evidence from Danish land inequality 1682–1895. *Journal of Development Economics*, *155*, 102813. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304387821001656 doi: https://doi.org/10.1016/j.jdeveco.2021.102813

Bogart, D., Lefors, M., & Satchell, A. E. M. (2019). Canal carriers and creative destruction in English transport. *Explorations in Economic History*, *71*, 1–24. Retrieved from https://www.sciencedirect.com/science/article/pii/S0014498318300263 doi: https://doi.org/10.1016/j.eeh.2018.08.005

Bogart, D., You, X., Alvarez-Palau, E. J., Satchell, M., & Shaw-Taylor, L. (2022). Railways, divergence, and structural change in 19th century England and Wales. *Journal of Urban Economics*, *128*, 103390. Retrieved from https://www.sciencedirect.com/science/article/pii/S0094119021000723 doi: https://doi.org/10.1016/j.jue.2021.103390

Bosker, M. (2022). City origins. *Regional Science and Urban Economics*, *94*, 103677. Retrieved from https://www.sciencedirect.com/science/article/pii/S0166046221000375 doi: https://doi.org/10.1016/j.regsciurbeco.2021.103677

Bosker, M., & Buringh, E. (2017). City seeds: Geography and the origins of the european city system. *Journal of Urban Economics*, *98*, 139-157. Retrieved from https://www.sciencedirect.com/science/article/pii/S0094119015000613 (Urbanization in Developing Countries: Past and Present) doi: https://doi.org/10.1016/j.jue.2015.09.003

Brakman, S., Garretsen, H., & Schramm, M. (2004). The strategic bombing of German cities during World War II and its impact on city growth. *Journal of Economic Geography*, *4*(2), 201–218. Retrieved from https://doi.org/10.1093/jeg/4.2.201 doi: 10.1093/jeg/4.2.201

Breiman, L. (2001). Random Forests. *Machine Learning*, *45*(1), 5–32. Retrieved from https://doi.org/10.1023/A:1010933404324 doi: 10.1023/A:1010933404324

Britannica. (2018). *Kaliningrad.* Retrieved from https://www.britannica.com/place/Kaliningrad

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407620303948 doi: https://doi.org/10.1016/j.jeconom.2020.12.001

Caruana-Galizia, P., Okubo, T., & Wolf, N. (2021). Underlying sources of growth: First and second nature geography. In S. Broadberry & K. Fukao (Eds.), *The cambridge economic history of the*

modern world (p. 382–417). Cambridge University Press.

Cermeño, A. L., & Enflo, K. (2019). Can kings create towns that thrive? The long-term implications of new town foundations. *Journal of Urban Economics*, *112*, 50–69. Retrieved from https://www.sciencedirect.com/science/article/pii/S0094119019300488 doi: https://doi.org/10.1016/j.jue.2019.06.001

Charlier, R. H. (2011). The Zwin: From golden inlet to nature reserve. *Journal of Coastal Research*, *27*(4), 746–756. doi: 10.2112/10A-00003.1

Chen, J., & Roth, J. (2023). *Log-like? Identified ATEs defined with zero-valued outcomes are (arbitrarily) scale-dependent.* Retrieved from http://www.jonathandroth.com

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *CoRR*, *abs/1603.0*. Retrieved from http://arxiv.org/abs/1603.02754

Christensen, C. (1735). Tilstanden i Thy og paa Mors ved Aar 1735. *Historie/Jyske Samlinger*, *4. række,*. Retrieved from http://img.kb.dk/tidsskriftdk/pdf/ho/ho_4rk_0003-PDF/ho_4rk_0003_75584.pdf

Christensen, J. T., Cedhagen, T., & Hylleberg, J. (2004). Late-Holocene salinity changes in Limfjorden, Denmark. *Sarsia*, *89*(6), 379–387. doi: 10.1080/00364820410002640

Clark, G., & Gray, R. (2014). Geography is not destiny: Geography, institutions and literacy in England, 1837-63. *Oxford Economic Papers*, *66*(4), 1042–1069. doi: 10.1093/oep/gpu020

Clausen, N. F. (2015). The Danish Demographic Database—Principles and Methods for Cleaning and Standardisation of Data. *Population Reconstruction*, 1–22. doi: 10.1007/978-3-319-19884-2_1

Crafts, N., & Wolf, N. (2014). The location of the uk cotton textiles industry in 1838: A quantitative analysis. *The Journal of Economic History*, *74*(4), 1103–1139. Retrieved 2024-03-17, from http://www.jstor.org/stable/24550695

Cronon, W. (1992). *Nature's metropolis: Chicago and the great west* (1st pbk. ed. ed.). New York: W. W. Norton.

Dahl, C. M., Johansen, T. S. D., & Vedel, C. (2024). *Breaking the hisco barrier: Automatic occupational standardization with occcanine.* Retrieved from https://arxiv.org/abs/2402.13604

Dalgaard, C. J., Knudsen, A. S. B., & Selaya, P. (2020). *The bounty of the sea and long-run development* (Vol. 25) (No. 3). Springer US. Retrieved from https://doi.org/10.1007/s10887-020-09181-8 doi: 10.1007/s10887-020-09181-8

Davis, D. R., & Weinstein, D. E. (2002, dec). Bones, Bombs, and Break Points: The Geography of Economic Activity. *American Economic Review*, *92*(5), 1269–1289. Retrieved from https://www.aeaweb.org/articles?id=10.1257/000282802762024502 doi: 10.1257/000282802762024502

Decet, D., & Marcucci, A. (2023, 12 21). Water wars. *SSRN Electronic Journal*. (Available at SSRN: https://ssrn.com/abstract=4672234 or http://dx.doi.org/10.2139/ssrn.4672234)

Degn, O. (1989). Byer, byhierarkier og byudvikling. *Historie/Jyske Samlinger*. Retrieved from https://tidsskrift.dk/historiejyskesamling/article/view/40197

Diamond, J. (1997). *Guns, Germs and Steel - The Fates of Human Societies* (1st ed.). New York, London: W. W. Norton & Company.

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*(1), 269–271. Retrieved from https://doi.org/10.1007/BF01386390 doi: 10.1007/BF01386390

Diørup, C. (1842). *Thisted Amt.* Tr. Retrieved from https://books.google.co.uk/books?id=FOdRAAAAcAAJ

Donaldson, D., & Hornbeck, R. (2016). Railroad and American Economic Growth: A 'Market Access' Approach. *The Quarterly Journal of Economics*, *131*(2), 799–858. doi: 10.1093/qje/qjw002.Advance

Duranton, G., & Turner, M. A. (2012). Urban Growth and Transportation. *The Review of Economic Studies*, *79*(4), 1407–1440. Retrieved from https://doi.org/10.1093/restud/rds010 doi: 10.1093/restud/rds010

Easterly, W., & Levine, R. (2003). Tropics, germs, and crops: How endowments influence economic development. *Journal of Monetary Economics*, *50*(1), 3–39. doi: 10.1016/S0304-3932(02)00200-3

Eaton, J., & Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, *70*(5), 1741–1779. doi: 10.1111/1468-0262.00352

Feldbæk, O. (2015). *Englænderkrigene ('The English Wars').* Retrieved from https://denstoredanske.lex.dk/Engl{\ae}nderkrigene

Fernihough, A., & O'Rourke, K. H. (2020). Coal and the European Industrial Revolution. *The Economic Journal*, *131*(635), 1135–1149. Retrieved from https://doi.org/10.1093/ej/ueaa117 doi: 10.1093/ej/ueaa117

Feyrer, J. (2021). Distance, trade, and income — The 1967 to 1975 closing of the Suez canal as a natural experiment. *Journal of Development Economics*, *153*(May), 102708. Retrieved from https://doi.org/10.1016/j.jdeveco.2021.102708 doi: 10.1016/j.jdeveco.2021.102708

Fogel, R. (1964). *Railroads and American Economic Growth: Essays in Econometric History* (1st ed.). The Johns Hopkins University Press.

Frankel, J. A., & Romer, D. (1999). Does trade cause growth? *American Economic Review*, *89*(3), 379–399. doi: 10.1257/aer.89.3.379

Galor, O. (2005). Chapter 4 From Stagnation to Growth: Unified Growth Theory. *Handbook of Economic Growth*, *1*(SUPPL. PART A), 171–293. doi: 10.1016/S1574-0684(05)01004-X

Galor, O. (2011). *Unified Growth Theory* (1st ed.). Princeton & Oxford: Princeton University Press.

Gøbel, E. (2010). øresundstolden og dens regnskaber 1497-1857. *M/S Museet for Søfarts årbog*, *69*, 41–72.

Grammaticus, S. (1208). *Gesta Danorum (Translation by Grundtvig 1818).* Retrieved from http://runeberg.org/saxo/

Hald, J. C. (1833). *Ringkjøbing Amt.* Landhuusholdningsselskabet.

Harper, D. (2022). *Etymology of fjord.* Retrieved 2022-05-29, from https://www.etymonline.com/word/fiord

Harris, C. D. (1954). The Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers*, *44*(4), 315–348.

Henderson, J. V., Squires, T., Storeygard, A., & Weil, D. (2018). THE GLOBAL DISTRIBUTION OF ECONOMIC ACTIVITY: NATURE HISTORY, AND THE ROLE OF TRADE. *The Quarterly Journal of Economics*, 357–406. doi: 10.1093/qje/qjx030

Henriques, S., Sharp, P., Tsoukli, X., & Vedel, C. (2021). Resilience, Adaptability and Transformability: Danish Butter Factories in the Face of Coal Shortages. *EHES Working Paper No. 220*. Retrieved from http://ehes.org/EHES_220.pdf

Hornbeck, R., & Rotemberg, M. (2019). Railroads, Reallocation, and the Rise of American Manufacturing. *NBER Working Paper*(26594). Retrieved from https://www.nber.org/papers/w26594

Hornung, E. (2015). Railroads and growth in prussia. *Journal of the European Economic Association*, *13*(4), 699–736. Retrieved 2024-03-17, from http://www.jstor.org/stable/24539267

Houtte, J. A. V. (1966). The Rise and Decline of the Market of Bruges. *The Economic History Review*, *19*(1), 29–47. Retrieved from http://www.jstor.org/stable/2592791

Jensen, P. S., Pedersen, M. U., Radu, C. V., & Sharp, P. R. (2022). Arresting the Sword of Damocles: The transition to the post-Malthusian era in Denmark. *Explorations in Economic History*, *84*, 101437. Retrieved from https://www.sciencedirect.com/science/article/pii/S0014498321000644 doi: https://doi.org/10.1016/j.eeh.2021.101437

Ketterer, T. D., & Rodríguez-Pose, A. (2018). Institutions vs. 'first-nature' geography: What drives economic growth in Europe's regions? *Papers in Regional Science*, *97*(269868), S25–S62. doi: 10.1111/pirs.12237

Khaustova, E., & Sharp, P. R. (2015). A Note on Danish Living Standards using Historical Wage Series, 1731-1913. *Journal of European Economic History*, *44*(3), 143–172.

Klem, K. (1967). Hjuldamperen IRIS og Limfjordsfarten. *M/S Museet for Søfarts årbog*, *26*, 125–152. Retrieved from https://tidsskrift.dk/mfs_aarbog/article/view/96199

Klemp, M., & Møller, N. F. (2016). Post-Malthusian Dynamics in Pre-Industrial Scandinavia. *The Scandinavian Journal of Economics*, *118*(4), 841–867. Retrieved from http://www.jstor.org/stable/45097668

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, *99*(3), 483–499. Retrieved from https://www.jstor.org/stable/2937739IncreasingReturnsandEconomicGeography doi: 10.1086/261763

Lampe, M., & Sharp, P. (2015). How the Danes discovered Britain: the international integration of the Danish dairy industry before 1880. *European Review of Economic History*, *19*(4), 432–453. Retrieved from https://doi.org/10.1093/ereh/hev013 doi: 10.1093/ereh/hev013

Lampe, M., & Sharp, P. (2018). *A Land of Milk and Butter*. University of Chicago Press. Retrieved from https://ideas.repec.org/b/ucp/bkecon/9780226549507.html

Lassen, K. C. (1883). Udviklingen af Dampskibstrafikken fra Danmark til. *Tidsskrift for Landøkonomi*(65). Retrieved from https://tidsskrift.dk/tidsskriftlandoekonomi/issue/view/6869

Leeuwen, M., Maas, I., & Miles, A. (2002). *Hisco: Historical international standard classification of occupations*. Leuven University Press. Retrieved from https://books.google.dk/books?id=

EMPtAAAAIAAJ

Marczinek, M., Maurer, S. E., & Rauch, F. (2022). Trade Persistence and Trader Identity - Evidence from the Demise of the Hanseatic League. *CEPR Discussion paper*(DP16957). Retrieved from https://cepr.org/active/publications/discussion_papers/dp.php?dpno=16957

Mathiesen, N., Robinson, O., Thomsen, A., & Revuelta-Eugercios, B. (2022). *Link-Lives Data v.1.2.1.* Denmark: Danish National Archives/University of Copenhagen. Retrieved from https://www.rigsarkivet.dk/udforsk/link-lives-data/

Matranga, A. (2024, 04). The Ant and the Grasshopper: Seasonality and the Invention of Agriculture. *The Quarterly Journal of Economics*, qjae012. Retrieved from https://doi.org/10.1093/qje/qjae012 doi: 10.1093/qje/qjae012

Maurer, N., & Yu, C. (2008). What T. R. took: The economic impact of the Panama Canal, 1903-1937. *Journal of Economic History*, *68*(3), 686–721. doi: 10.1017/S0022050708000612

Meesenburg, H. (2017). *Limfjorden.* Retrieved from lex.dk

Ministry of Food, A., & Fisheries. (2022). *Fiskeristatistik - dynamisk landingstabel.* Retrieved from https://fiskeristyrelsen.dk/fiskeristatistik

Mortensen, L. B. (2018). *Saxo.* Aarhus: Aarhus Universitetsforlag.

Nunn, N., & Qian, N. (2011). The potato's contribution to population and urbanization: Evidence from a historical experiment. *Quarterly Journal of Economics*, *126*(2), 593–650. doi: 10.1093/qje/qjr009

Ørnbjerg, J. (2022). *Det vilde sildeboom 1703* (T. Oldrup, Ed.). Aarhus Universitetsforlag. Retrieved from https://unipress.dk/udgivelser/d/det-vilde-sildeboom/

Ortman, S., & Lobo, J. (2020). Smithian growth in a nonindustrial society. *Science Advances*, *6*(25). doi: 10.1126/sciadv.aba5694

Pajung, S. (2012). Knud den Hellige, ca. 1042-1086. *danmarkshistorien.dk*. Retrieved from https://danmarkshistorien.dk/vis/materiale/knud-den-hellige-ca-1042-1086/

Pedersen, A., & Wilson, D. M. (2014). *Aggersborg - The Viking-Age settlement and fortress.* National Museum of Denmark, Jutland Archaeological Society.

Pedersen, S., Hermansen, B., Nathan, C., & Tougaard, L. (2019). Surface geology map of Denmark 1:200 000, version 2. *GEUS*. Retrieved from https://eng.geus.dk/products-services-facilities/data-and-maps/maps-of-denmark

Peters, J., Bühlmann, P., & Meinshausen, N. (2016). Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *78*(5), 947–1012. doi: 10.1111/rssb.12167

Petersen, J. C. (1877). Om Aggertangen før og nu. *Geografisk Tidskrift*, *1*.

Petersen, P. B. (1853). *Oplysende Bemærkninger angaaende Limfjorden, Løgstør og dens Omegn, især med Hensyn til de paatænkte Forbedringer og de forandrede Forhold i Fremtiden.* Retrieved from https://books.google.co.uk/books?id=JtCeGL0Nm-4C

Portner, H.-O., Roberts, D., Trisos, C., Simpson, N., & Möller, V. (2022). Summary for Policymakers: Climate Change 2022: Impacts, Adaptation, and Vulnerability. *IPCC*. doi: 10.1017/9781009325844.001

Poulsen, B. (2019). *Stormflod* (1st ed.; P. Bejder, Ed.). Aarhus: Aarhus Universitetsforlag. Retrieved from https://unipress.dk/bogserier/100-danmarkshistorier/

Poulsen, B. (2022). Between Adaptation and Mitigation. The Nineteenth-century North Sea Storm Surges and the Entangled Socio-Natural Transformation of the Limfjord Region, Denmark. *Journal for the History of Environment and Society*, *6*(1), 129–158.

Poulsen, B., Holm, P., & MacKenzie, B. R. (2007). A long-term (1667-1860) perspective on impacts of fishing and environmental variability on fisheries for herring, eel, and whitefish in the Limfjord, Denmark. *Fisheries Research*, *87*(2-3), 181–195. doi: 10.1016/j.fishres.2007.07.014

Rasmussen, H. (1966). Vel kaldes det tilsammen Limfjorden. *Skalk*(5), 18–27.

Rauch, F., & Maurer, S. (2022). Economic geography aspects of the Panama Canal. *Oxford Economic Papers*. doi: https://doi.org/10.1093/oep/gpac009

Ravn, T. B. (1993). Gennembrud. *Skalk*(3), 18–24.

Redding, S. J., & Sturm, D. M. (2008). The Costs of Remoteness: Evidence from German Division and Reunification. *American Economic Review*, *98*(5), 1766–1797. Retrieved from https://www.jstor.org/stable/29730152

Redding, S. J., & Turner, M. A. (2015). *Transportation Costs and the Spatial Organization of Economic Activity* (1st ed., Vol. 5). Elsevier B.V. Retrieved from http://dx.doi.org/10.1016/B978-0-444-59531-7.00020-X doi: 10.1016/B978-0-444-59531-7.00020-X

Ricardo, D. (1817). *On the Principles of Political Economy and Taxation.* J. Murray. Retrieved from https://books.google.co.uk/books?id=cUBKAAAAYAAJ

Richard A. Hornbeck, M. R. (2024). Growth off the rails: Aggregate productivity growth in distorted economies.. Retrieved from https://bpb-us-w2.wpmucdn.com/voices.uchicago.edu/dist/b/2217/files/2024/02/Railroads_HR_Feb2024-59d1fb5846b8aeff.pdf

Rodriguez, F., & Rodrik, D. (2001). *to the Cross-National Evidence Policy Skeptic ' s and to Growth : the Cross-National Evidence* (Vol. 15) (No. January).

Rodrik, D., Subramanian, A., & Trebbi, F. (2004). Institutions rule: the primacy of institutions over geography and integration in economic development. *Journal of economic growth*, *9*(2), 131–165.

Roesdahl, E. (2009). De sidste årtier af vikingetiden. *danmarkshistorien.dk*. Retrieved from https://danmarkshistorien.dk/perioder/vikingetiden-ca-800-1050/de-sidste-aartier-af-vikingetiden/

Roesdahl, E. (2022). Vikingetiden, ca. 800-1050. *Danmarkshistorien.dk*. Retrieved from https://danmarkshistorien.dk/vis/materiale/vikingetiden-ca-800-1050

Roskilde Amts og Advertissementstidende. (1836). *Roskilde Amts og Advertissementstidende.* Retrieved from http://hdl.handle.net/109.3.1/uuid:ae3932f7-2868-4c8e-9524-d3115a891af3

Sant'Anna, P. H. C., & Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, *219*(1), 101–122. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304407620301901 doi: https://doi.org/10.1016/j.jeconom.2020.06.003

Santos Silva, J. M. C., & Tenreyro, S. (2022). The Log of Gravity at 15. *Portuguese Economic*

Journal. Retrieved from https://doi.org/10.1007/s10258-021-00203-w doi: 10.1007/s10258-021-00203-w

Schovelin, J. (1891). *Blade af den Danske Dampskibsfartshistorie.* Retrieved from https://slaegtsbibliotek.dk/905231.pdf

Seror, M. (2020). *Random river: Trade and rent extraction in imperial china* (Document de travail No. 2020-17). Montréal. Retrieved from http://hdl.handle.net/10419/234817

Silva, J. S., & Tenreyro, S. (2006). The Log of Gravity. *The Review of Economics and Statistics*, *88*(4), 641–658. Retrieved from https://www.jstor.org/stable/40043025

Spejlborg, M. B. (2012). Knud den Store ca. 995-1035. *Danmarkshistorien.dk*. Retrieved from https://danmarkshistorien.dk/vis/materiale/knud-den-store-ca-995-1035/

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016, sep). Increasing Transparency Through a Multiverse Analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. Retrieved from https://doi.org/10.1177/1745691616658637 doi: 10.1177/1745691616658637

Svalgaard, R. (1977). Toldvagtskibe ved Agger Kanal samt toldvagtskibet "Thyborøn". *Årbog for Thy, Mors og Vester Hanherred*, 35–57. Retrieved from https://www.arkivthy.dk/images/Aarbog/1977/Svalgaard,RobertToldvagtskibetvedAggerkanalsamttold.pdfhttps://www.arkivthy.dk/

Tobler, W. R. (1970). A Computer Movie Simulating Urban Growth in the Detroit Region. *Economic Geography*, *46*, 234–240. doi: 10.1126/science.ns-13.332.462

Trap, J. P., Falbe-Hansen, V., Westergaard, H., & Weitemeyer, H. (1906). *Kongeriget Danmark (Trap Danmark 3rd edition)* (3rd ed.). Copenhagen: Universitetsboghandler G. E. C. Gad. Retrieved from http://runeberg.org/trap/

van Etten, J. (2017). R Package gdistance: Distances and Routes on Geographical Grids. *Journal of Statistical Software*, *76*(13), 1–21. Retrieved from https://www.jstatsoft.org/index.php/jss/article/view/v076i13 doi: 10.18637/jss.v076.i13

Vedel, C., Dahl, C. M., & Johansen, T. S. D. (2024). HISCO codes for Danish Census data. *Harvard Dataverse*. doi: 10.7910/DVN/WZILNI

Veluwenkamp, J., & Woude, S. (2009). SoundToll Registers Online. *Database*. Retrieved from www.soundtoll.nl/index.php/en/welkom

Viborg Stiftstidende. (1852). *Viborg Stiftstidende.* Retrieved from http://hdl.handle.net/109.3.1/uuid:2980a160-6418-42a7-a227-df25503f6134

Zimran, A. (2020). Transportation and health in the antebellum United States, 1820-1847. *Journal of Economic History*, *80*(3), 670–709. doi: 10.1017/S0022050720000315

Appendix

A Perfect Storm: First-Nature Geography and Economic Development

Christian Vedel, University of Southern Denmark,

christian-vs@sam.sdu.dk; https://github.com/christianvedels/A_perfect_storm

Contents

A Details of market access computation **ii**

B Additional trade results **ii**

C Population results **v**

C.1 Full regression table v

C.2 Population multiverse vi

C.3 Doubly robust estimates vii

D Mechanims **viii**

D.1 All occupational major categories estimates viii

D.2 Event plot fishing and spinning xi

D.3 Effects by age group xiii

E Archaeological findings **xiv**

E.1 Math note xiv

E.1.1 Probability of a single coin xiv

E.1.2 At the parish level xv

E.1.3 Reverse probability xv

E.1.4 Estimation xvi

E.2 All parameter estimates xvii

E.3 Normal distribution xix

A Details of market access computation

The effect of the channel on market access is computed based on

$$MA_p = \sum_{h \in H} [CostDist(p, h; \alpha) + 1]^{\theta} \tag{11}$$

Most of this is defined in the main paper, but here follows details on θ and CostDist(): θ determines the spread of the market potential function. It reflects the elasticity to distance. A large absolute value of θ corresponds to a very localized effect of a change to market potential. The standard $\theta = -1$ is used in the main specification. This is the original value suggested by Harris (1954), but is also used in Rauch and Maurer (2022) and is very close to what is empirically estimated by Redding and Sturm (2008). However, it is plausible that other values are more appropriate e.g. $\theta = -8$ as estimated by Donaldson and Hornbeck (2016). Robustness checks with $\theta \in (-1, -2, -4, -8, -16)$ can be found in this appendix. It makes no qualitative difference in the conclusions.

The function $CostDist()$ is the result of the following optimization:

$$CostDist(x, y) := \min_{r \in R} \left[Dist_r^{water}(x, y) + \alpha \, Dist_r^{land}(x, y) \right] \tag{12}$$

$CostDist(x, y)$ represents the cost of the shortest route r^* between x and y in the set of all possible routes R given that land travel is α times more expensive than ocean travel. $\alpha = 10$ is used following Marczinek, Maurer, and Rauch (2022) and Rauch and Maurer (2022). However robustness checks are carried out with both $\alpha = 5$, $\alpha = 20$ and $\alpha = 50$.

Computing this optimized distance is a computationally hard problem, which was solved via the 'gdistance' R package (van Etten, 2017). This implements Dijkstra's algorithm (Dijkstra, 1959) which is the standard method for calculating cost distances. The algorithm takes a series of nodes with a given cost between them and finds the least cost path. The nodes in this case are a grid representing Denmark with two node types 'land' and 'water' and the corresponding relative cost to traverse it is α of equation 12. From each grid cell, it is possible to travel to each of the 8 surrounding nodes (neighbouring grid cells) with the minimum cost of either node. The grid has a resolution of 500x500 m. This is the highest resolution which was computationally feasible.

Around Løgstør (see historical background) shallow water forced traders to reload goods onto prams and pay the locals for transport. This area is encoded as having the same cost as land transportation. This is an upper bound on the market powers of the locals. Principally they could charge up to this cost before it would be more profitable to transport goods via land instead.

B Additional trade results

Table A1 contains regression results for the sound toll data. The problem with the regression is all the zero outcomes. This contains estimates of the form

$$y_{it} = \exp\left(\beta_0 + Location_{it}\,\beta_{1r} + I[t \geq 1834]\,\beta_2 + Location_{it}\,\times I[t \geq 1834]\,\beta_{3r} + \varepsilon_{it}\right) \tag{13}$$

Here y_{it} is the amount of trade. In the standard specification measured as the sum of all traffic to or from port i at time t. Each port belongs to a location, $r \in (west, middle, east)$. This is similar to a classical difference in difference estimate but given three distinct treatments of being in each of the potentially affected parts of the Limfjord. This enables the estimator to capture both the effect on the region we expect to be affected (the western and middle Limfjord) and any spill-over effect of trade relocating from east to west. Silva and Tenreyro (2006) demonstrate that the Poisson Pseudo Maximum Likelihood (PPML) estimator is usually more appropriate for log-linear relationships like this. This has become standard for gravity model estimations (Marczinek et al., 2022; Santos Silva & Tenreyro, 2022), and the application presented here is similar. To guard against potential problems from the Napoleonic wars, the effect is estimated with and without data from 1807 to 1814 (the Napoleonic war) and the period 1825 to 1833 (before the channel was fully navigable). Results from the classical $log(y + 1)$ estimator is also presented. Recent work by J. Chen and Roth (2023) outline a number of alternatives, which motivates the additional estimators presented.

Column 1 shows poisson estimates of the the amount of trade using the Poisson estimator. Column 2 exclude years from 1807 to 1814 and 1825 to 1833. Column 3 uses the log(x+1) transformation to deal with zero outcomes. Column 4 uses the inverse hyperbolic sine. Column 5 estimates the intensive margin (whether any trade went to the port or not rather than the amount). Whichever method employed this demonstrates the same qualitiatve result as the plot in the main paper: The west and middle Limfjord received much more traffic after 1834.

Table A1: Channel introduction and trade

Dependent Variables:	traffic		log(traffic+1)	arcsinh(traffic)	$1[traffic > 0]$
	(1)	(2)	(3)	(4)	(5)
	Poisson	Poisson	OLS	OLS	OLS
Post×east	-0.4896**	-0.4817*	-0.0029	0.0115	0.0115
	(0.2172)	(0.2489)	(0.0915)	(0.1741)	(0.1741)
Post×middle	5.557***	11.76***	0.9672***	0.3811***	0.3811***
	(0.2151)	(0.2471)	(0.0584)	(0.0233)	(0.0233)
Post×west	4.129***	4.698***	0.3821	0.3194**	0.3194**
	(0.5090)	(0.4714)	(0.2635)	(0.1263)	(0.1263)
Observations	13,356	11,214	13,356	13,356	13,356
1807-1814 excl.	No	Yes	No	No	No
1825-1833 excl.	No	Yes	No	No	No

Notes: Cluster-robust standard errors in the parenthesis. Clustered at the port level. Coloumn (1) shows Poisson results. Column (2) shows Poisson results but where the years of the Napoleonic wars are removed. Column (3) and (4) show OLS results using the $log(x+1)$ or $archsinh(x)$ transformation to deal with zero outcomes. Column (5) shows extensive margin results. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.

Source: Danish census data

C Population results

C.1 Full regression table

Table A2: Regression results for population size

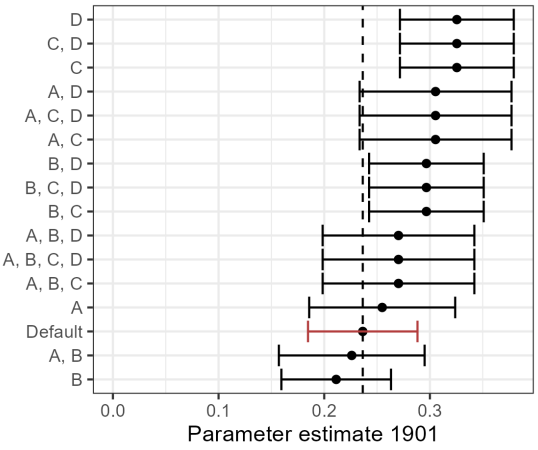
Outcome:	log(Population)	
	(1)	(2)
	Dummy approach	Market access approach
Year 1787 × Affected	0.0160	0.0379
	(0.0106)	(0.0626)
Year 1834 × Affected	0.0177	-0.0146
	(0.0127)	(0.0748)
Year 1840 × Affected	-0.0060	-0.2651***
	(0.0132)	(0.0759)
Year 1845 × Affected	-0.0029	-0.2531***
	(0.0143)	(0.0775)
Year 1850 × Affected	0.0021	-0.2466***
	(0.0152)	(0.0900)
Year 1860 × Affected	0.0319*	0.0183
	(0.0175)	(0.0972)
Year 1880 × Affected	0.1379***	0.8349***
	(0.0203)	(0.1155)
Year 1901 × Affected	0.2364***	1.591***
	(0.0264)	(0.1607)
Observations	14,301	14,301
Parishes	1,589	1,589
Parish FE	Yes	Yes
Year FE	Yes	Yes

Notes: Cluster-robust standard errors in the parenthesis. Clustered at the parish level. Affected is either a dummy for being in the West Limfjord or improvement in market access, which is indicated by the headers of the results. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$.
Source: Danish census data

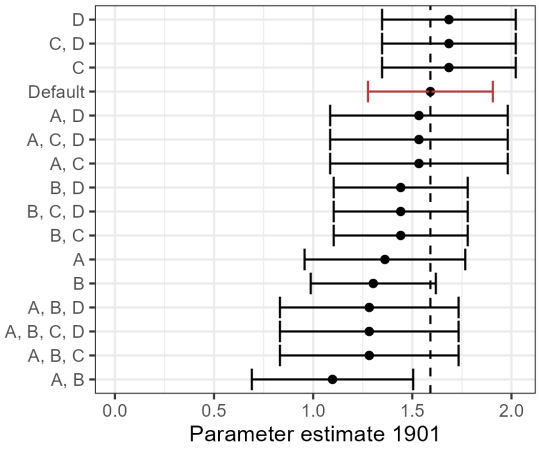
C.2 Population multiverse

Figure A1: Multiverse of the effect in different comparison groups and parameter choices, 1787

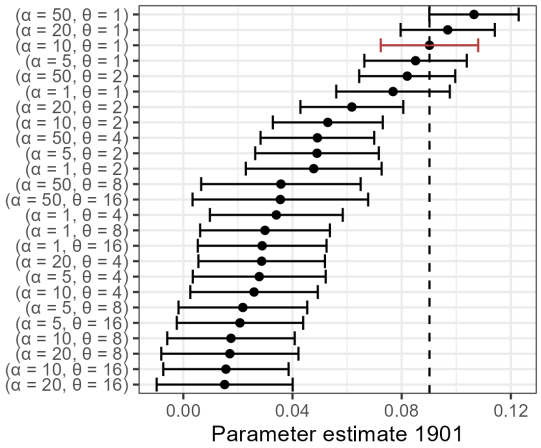
(a) Multiverse of control groups
Dummy approach



(b) Multiverse of control groups
Market access approach



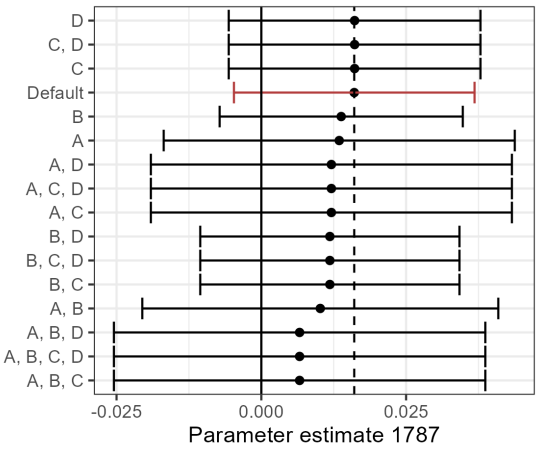
(c) Multiverse of feasible parameters
Market access approach



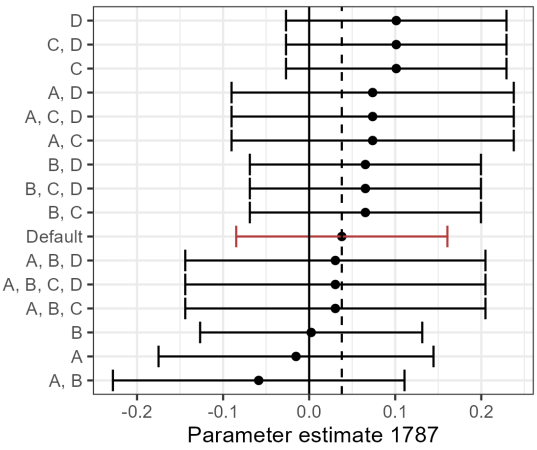
Notes: This is the multiverse of parameter estimates of the effect in 1901 given different feasible choices that could have been made for how to run the analysis. In panel a and panel b, 'A', 'B', and 'C', represents subgroups of the data. 'A' is the result, when the regression is computed using only parishes with a centroid less than 5 km from the coast. 'B' is the subgroups, where all parishes in around Copenhagen are excluded, 'C' is the subgroup of parishes, where the control group does not contain any parishes within 100 km of the Limfjord. 'D' represents the result when using only parishes located within 5 km of a market town. Panel (c) represents the effect given different market access parameters. For enhanced comparability, the log change in market access is standardized to unit variance and zero mean.
Source: Danish census data

Figure A2: Multiverse of the effect in different comparison groups and parameter choices

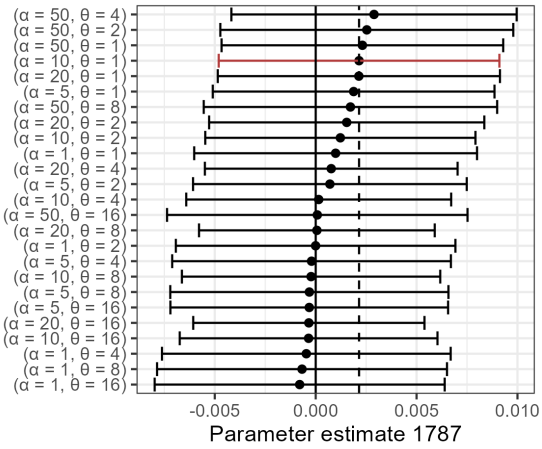
(a) Multiverse of control groups
Dummy approach



(b) Multiverse of control groups
Market access approach



(c) Multiverse of feasible parameters
Market access approach



Notes: Multiverse results for 1787. If any of these were different from zero it would indicate the existence of pretends. The dotted line indicate the default parameter estimate. The solid line is at zero.

C.3 Doubly robust estimates

This section contains estimates based on the doubly-robust multi-period did estimator from Callaway and Sant'Anna (2021). Table A3 shows the results. As is default in their implementation, the first period is the reference. As consequence, the reference year here is 1787 rather than 1801 from the rest of my paper. To address concern over balance, covariates are included. The doubly robust method combines a propensity score and an outcome regression and is consistent if either of these are correctly specified (Sant'Anna & Zhao, 2020). Column 1 shows results without covariates. Column 2 shows results including covariates (Occupation, Young children per woman and number of people in different

age groups). Column 3 also includes the pre-event population as a covariate.

Table A3: Callaway and Sant'Anna estimates

Outcome:	log(Population)		
	(1)	(2)	(3)
1801	-0.0152	-0.0273	-0.0273
	(0.0114)	(0.0141)	(0.0133)
1834	0.0221	0.0041	0.0041
	(0.0133)	(0.0148)	(0.0145)
1840	0.0004	-0.0097	-0.0097
	(0.0132)	(0.0154)	(0.0145)
1845	0.0031	-0.0058	-0.0058
	(0.0137)	(0.0156)	(0.0149)
1850	0.0076	-0.0040	-0.0040
	(0.0146)	(0.0174)	(0.0160)
1860	0.0343	0.0087	0.0087
	(0.0166)	(0.0187)	(0.0194)
1880	0.1336*	0.0873*	0.0873*
	(0.0198)	(0.0229)	(0.0233)
1901	0.2262*	0.1590*	0.1590*
	(0.0263)	(0.0276)	(0.0310)
Observations	14,2741	14,274	14,274

'' confidence band (95 percent) does not cover 0*

Notes: Effect using the estimator proposed by Callaway & Sant'Anna (2021). Column (1) includes no covariates. Column (2) adjusts for pre-event covariates. Column (3) adjusts for pre-event covariates including pre-event population.
Source: Danish census data

D Mechanims

D.1 All occupational major categories estimates

Table A4: Effect on occupation in 1901 (HISCO first digit 1 to 3)

hisco	Affected	Approach	Estimate (1901)	n parishes
0/1	MA	3: log(x+1)	-0.642 (0.337)	1589
0/1	MA	4: asinh(x)	-0.595 (0.397)	1589
0/1	MA	1: Extensive	0.414 (0.139)	1589
0/1	MA	2: Intensive	-0.41 (0.445)	1342
0/1	Dummy	3: log(x+1)	-0.083 (0.06)	1589
0/1	Dummy	4: asinh(x)	-0.081 (0.07)	1589
0/1	Dummy	1: Extensive	0.064 (0.025)	1589
0/1	Dummy	2: Intensive	-0.047 (0.079)	1342
2	MA	3: log(x+1)	-0.871 (0.463)	1589
2	MA	4: asinh(x)	-0.561 (0.521)	1589
2	MA	1: Extensive	0.182 (0.148)	1589
2	Dummy	3: log(x+1)	-0.085 (0.085)	1589
2	Dummy	2: Intensive	0.081 (0.148)	764
2	Dummy	4: asinh(x)	-0.056 (0.096)	1589
2	MA	2: Intensive	0.04 (0.761)	764
2	Dummy	1: Extensive	-0.015 (0.025)	1589
3	MA	2: Intensive	1.962 (4.096)	75
3	Dummy	2: Intensive	1.252 (0.472)	75
3	MA	1: Extensive	1.128*** (0.259)	1589
3	MA	4: asinh(x)	0.697 (0.479)	1589
3	MA	3: log(x+1)	0.486 (0.381)	1589
3	Dummy	1: Extensive	0.169*** (0.044)	1589
3	Dummy	4: asinh(x)	0.101 (0.087)	1589
3	Dummy	3: log(x+1)	0.073 (0.07)	1589

Notes: Parameter estimate of the effect on occupational structure of the channel in 1901. Each row corresponds to a sepperate regression with all individuals with hisco codes starting with 0/1, 2 or 3 as outcome. The last column shows the number of parishes included in the regression, which is different from the full sample (1589) in the intensive margin estimates. As a rule of thumb, results with fewer than 100 observations should be entirely disregarded. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$. Standard errors clustered on the parish level in parenthesis. All p-values are Bonferroni-corrected.
Source: Danish census data.

Table A5: Effect on occupation in 1901 (HISCO first digit 4 to 9)

hisco	Affected	Approach	Estimate (1901)	n parishes
4	MA	2: Intensive	-4.772 (3.826)	28
4	MA	4: asinh(x)	-1.342 (0.576)	1589
4	MA	3: log(x+1)	-1.121 (0.489)	1589
4	Dummy	2: Intensive	-0.353 (0.402)	28
4	Dummy	4: asinh(x)	-0.182 (0.103)	1589
4	Dummy	3: log(x+1)	-0.143 (0.087)	1589
4	MA	1: Extensive	-0.123 (0.202)	1589
4	Dummy	1: Extensive	-0.058 (0.036)	1589
5	MA	4: asinh(x)	-0.943 (0.554)	1589
5	MA	1: Extensive	-0.662*** (0.163)	1589
5	MA	3: log(x+1)	-0.625 (0.462)	1589
5	Dummy	2: Intensive	0.309 (0.159)	660
5	MA	2: Intensive	0.302 (0.881)	660
5	Dummy	3: log(x+1)	0.057 (0.079)	1589
5	Dummy	4: asinh(x)	0.042 (0.094)	1589
5	Dummy	1: Extensive	-0.031 (0.028)	1589
6	MA	2: Intensive	1.173*** (0.196)	1545
6	MA	1: Extensive	-0.243*** (0.049)	1589
6	Dummy	2: Intensive	0.197*** (0.031)	1545
6	MA	4: asinh(x)	-0.123 (0.322)	1589
6	Dummy	3: log(x+1)	0.085 (0.036)	1589
6	Dummy	4: asinh(x)	0.072 (0.038)	1589
6	MA	3: log(x+1)	0.025 (0.294)	1589
6	Dummy	1: Extensive	-0.024*** (0.004)	1589
7/8/9	MA	2: Intensive	1.763*** (0.339)	1530
7/8/9	MA	4: asinh(x)	0.813 (0.39)	1589
7/8/9	MA	3: log(x+1)	0.709 (0.358)	1589
7/8/9	MA	1: Extensive	-0.218*** (0.05)	1589
7/8/9	Dummy	2: Intensive	0.215** (0.063)	1530
7/8/9	Dummy	4: asinh(x)	0.123 (0.068)	1589
7/8/9	Dummy	3: log(x+1)	0.11 (0.063)	1589
7/8/9	Dummy	1: Extensive	-0.02 (0.007)	1589

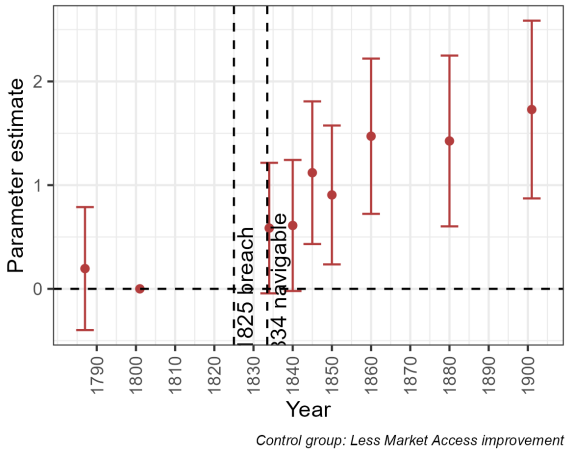
Notes: Parameter estimate of the effect on occupational structure of the channel in 1901. Each row corresponds to a sepperate regression with all individuals with hisco codes starting with 4, 5, or 6/8/9 as outcome. The last column shows the number of parishes included in the regression, which is different from the full sample (1589) in the intensive margin estimates. As a rule of thumb, results with fewer than 100 observations should be entirely disregarded. *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$. Standard errors clustered on the parish level in parenthesis. All p-values are Bonferroni-corrected.

Source: Danish census data.

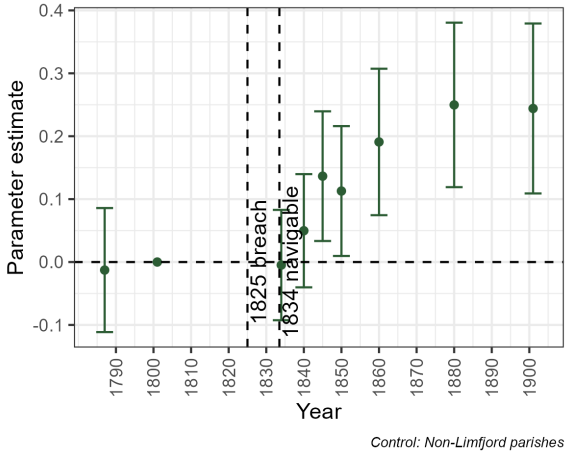
D.2 Event plot fishing and spinning

Figure A3: Fishermen and Spinners, Weavers, Knitters, Dyers And Related Workers

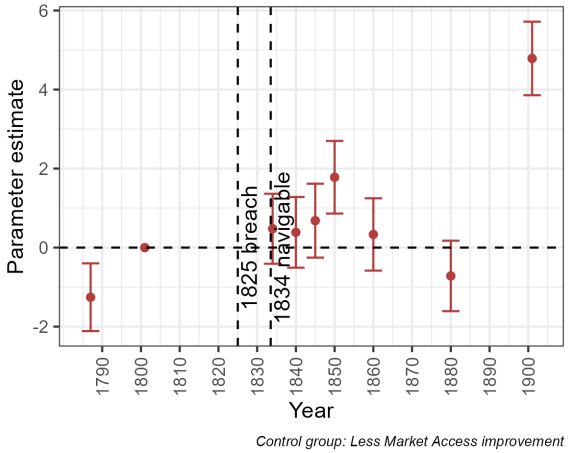
(a) Fishermen (MA approach)



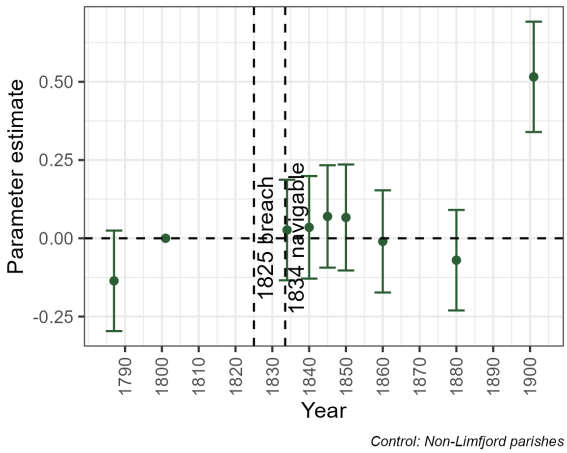
(b) Fishermen (Dummy approach)



(c) Spinners, weavers, knitters, dyers and related workers (MA approach)



(d) Spinners, weavers, knitters, dyers and related workers (Dummy approach)



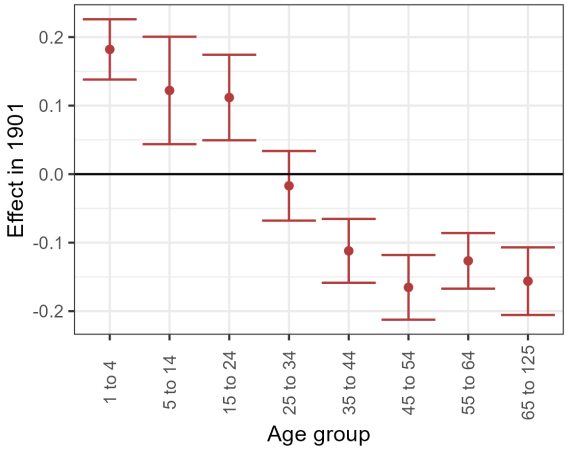
Notes: Panel (a) and panel (b) shows event plots for the effect of the channel on the number of fishermen. Panel (c) and (d) shows the effect to the number of spinners, weavers, knitters, dyers and related workers (HISCO codes starting with 75).

Source: Danish census data

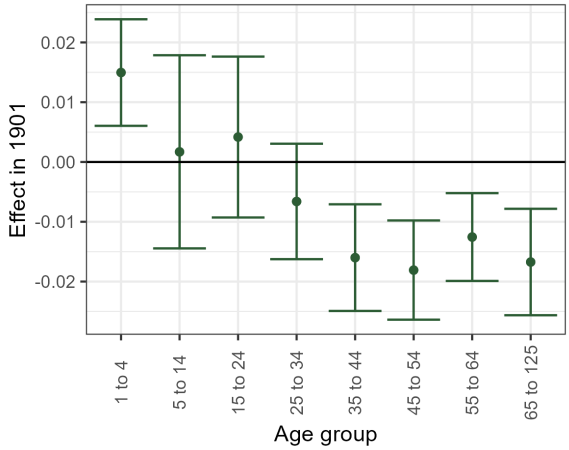
D.3 Effects by age group

Figure A4: Age group composition

(a) Effect by age group (MA approach)



(b) Effect by age group (dummy approach)



Notes: Regression parameter in 1901 given the market access approach (panel a) and the dummy approach (panel b). The outcome of each regression is the size of the particular age group as a share of the total population. What this shows, is that the population in the affected parishes became comparatively younger.
Source: Danish census data

E Archaeological findings

E.1 Math note

The derivation is based on a single parish i. To construct a panel this is simply repeated for all parishes. The derivation is based on coin findings. But it generalises to any kind of finding.

The data is of the form: $Coin = c$ was generated in time interval $t \in [Y^c_{min}; Y^c_{max}]$ (Archaeologists report a coin finding and date it to a range.)

That is, for each coin we observe $P(t|c) \sim [Y^c_{min}; Y^c_{max}]$

We want to know the probability, that any coin, $c \in \{1, ..., K\}$, finding was generated at any particular point in time. This event is referred to as $\{coins\}|t$

E.1.1 Probability of a single coin

We are interested in the probability that a any coin comes from a specific point in time. What is observed is:

$$P(t|c) = \begin{cases} \frac{1}{Y^c_{max} - Y^c_{min}}, & \text{if } Y^c_{min} \leq t \leq Y^c_{max} \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

Or writing the same with indicator function:

$$P(t|c) = 1[t \in [Y^c_{min}; Y^c_{max}]] \frac{1}{Y^c_{max} - Y^c_{min}} \tag{15}$$

I.e. it is equally likely that a coin truly originates at any particular point in time the range offered by the archaeologists.

Note: This is an assumption. The archaeologists specify a range but not a distribution. How should this range be interpreted? A straightforward alternative is to intepret it as a 95 percent confidence interval of the normal distribution. This is also tested.

$$\begin{aligned} &P(t|c) = \mathcal{N}(\mu_c, \sigma_c) \\ &\text{where:} \\ &\quad \mu_c = 0.5 \times (Y^c_{max} + Y^c_{min}) \\ &\quad \sigma_c \approx (Y^c_{max} - Y^c_{min})/1.96 \end{aligned} \tag{16}$$

E.1.2 At the parish level

The probability that *any* coin was generated as a finding in the parish (at least one of the draws from $P(t|c)$ was succesful) is a mixture distribution, where at least one component needs to be sucesful:

$$\begin{aligned} P(t|\{coins\}) &= 1 - \prod_{c=1}^{K} (1 - P(t|c)) \\ P(t|\{coins\}) &= 1 - \prod_{c=1}^{K} \left(1 - 1[t \in [Y^c_{min}; Y^c_{max}]] \left(\frac{1}{Y^c_{max} - Y^c_{min}}\right)\right) \end{aligned} \tag{17}$$

where $\{coins\}$ is the event of at least one coin $\{coins\} = \{1, ..., c, ..., K\}$.

Intuition The inner part of the expression $(1 - P(t|c))$ is the probability that a coin is *not* associated with that particular point in time. Taking the product over all coins, generates the combined probability, that *no coins at all* are associated with that particular point in time. The compliment of this is the probability we are interested in, in this step. It is the probability that *any* coin is associated with a particular point in time.

E.1.3 Reverse probability

We are interested in the reverse probability: The probability of any coins given a particular point in time. The formula for this is Bayes formula:

$$P(\{coins\}|t) = \frac{P(t|\{coins\})P(\{coins\})}{P(t)} \tag{18}$$

The prior, $P(\{coins\})$, is simply assumed to be a constant, $0 < c < 1$ and $P(t)$ can be found by marginalisation $P(t) = P(t|\{coins\}) + P(t|\neg\{coins\}) = 1$. It follows that

$$P(\{coins\}|t) = c \times P(t|\{coins\}) \tag{19}$$

E.1.4 Estimation

The following loop (pseudocode) produces samples from $P(\{coins\}|t)$ and estimates the probability of interest:

```
‘‘‘{pseudo code}
B = 1000 # Number of Monte Carlo samples

for b in 1 to B: # Loop of MC samples
... # Generate samples from Y_max_c Y_max_c:
... for c in 1 to C:
... ... t_c = sample_uniform(1, Y_min_c, Y_max_c)
... ... # Is t_c equal to t?
... ... coins_t[c] = t_c == t

... # Were there any coins associated with this time in this draw?
... number_of_coins = sum(coins_t)
... succces_t[b] = number_of_coins > 0

# Estimating the probability by analogy for each t
P_of_t_given_coins = sum(succces_t) / B

# Converting to P({coins}|t)
P_of_coins_given_t = c_hat * P_of_t_given_coins
‘‘‘
```

This code is then repeated for every parish and every point in time t. This gives a panel of size $N \times T$ containing the estimated probability that a coin finding was generated at a particular point in time. This in turn can be used in econometric applications.

E.2 All parameter estimates

Table A6 and A7 contain parameter estimates for all years for the regressions using archaeological findings.

Table A6: All parameters of table 3 columns 1-4

	Archaeological findings			
	(1)	(2)	(3)	(4)
	MA	Dummy	MA	Dummy
Year750 × Affected	0.0521***	0.0050***	0.0034	-0.0010
	(0.0092)	(0.0009)	(0.0101)	(0.0018)
Year800 × Affected	0.0195***	0.0020***	-0.0060	-0.0024*
	(0.0049)	(0.0005)	(0.0068)	(0.0013)
Year850 × Affected	0.0220***	0.0022***	-0.0072	-0.0028**
	(0.0051)	(0.0006)	(0.0068)	(0.0013)
Year900 × Affected	0.0086**	0.0010***	-0.0031	-0.0014*
	(0.0035)	(0.0004)	(0.0041)	(0.0008)
Year950 × Affected	-0.0042	-0.0002	0.0011	0.0002
	(0.0038)	(0.0004)	(0.0018)	(0.0003)
Year1050 × Affected	0.0059**	0.0008***	-0.0063**	-0.0007
	(0.0025)	(0.0003)	(0.0032)	(0.0006)
Year1100 × Affected	0.0077	0.0004	-0.0708***	-0.0066**
	(0.0121)	(0.0011)	(0.0165)	(0.0032)
Year1150 × Affected	0.0075	0.0003	-0.0681***	-0.0065*
	(0.0121)	(0.0011)	(0.0170)	(0.0033)
Year1200 × Affected	-0.0327**	-0.0049***	-0.0736***	-0.0069**
	(0.0151)	(0.0013)	(0.0174)	(0.0034)
Year1250 × Affected	-0.0941***	-0.0120***	-0.0772***	-0.0066*
	(0.0199)	(0.0018)	(0.0193)	(0.0038)
Year1300 × Affected	-0.1479***	-0.0147***	-0.0685***	-0.0068**
	(0.0236)	(0.0032)	(0.0178)	(0.0031)
Year1350 × Affected	-0.1923***	-0.0193***	-0.0612***	-0.0065**
	(0.0273)	(0.0037)	(0.0185)	(0.0031)
Year1400 × Affected	-0.1207***	-0.0116***	-0.0605***	-0.0060**
	(0.0239)	(0.0034)	(0.0177)	(0.0030)
Year1450 × Affected	-0.0606***	-0.0052	-0.0599***	-0.0060**
	(0.0234)	(0.0034)	(0.0174)	(0.0029)
Year1500 × Affected	-0.1114***	-0.0104***	-0.0666***	-0.0064*
	(0.0256)	(0.0037)	(0.0223)	(0.0035)

Custom standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Table A7: All parameters of table 3 columns 5-8

	Archaeological findings			
	(5)	(6)	(7)	(8)
	MA	Dummy	MA	Dummy
Year750 × Affected	0.0547**	0.0063**	-0.0043	-0.0003
	(0.0216)	(0.0025)	(0.0168)	(0.0026)
Year800 × Affected	0.0261**	0.0031**	-0.0248*	-0.0034
	(0.0122)	(0.0015)	(0.0136)	(0.0021)
Year850 × Affected	0.0231*	0.0027*	-0.0306**	-0.0049**
	(0.0118)	(0.0015)	(0.0137)	(0.0021)
Year900 × Affected	0.0071	0.0008	-0.0199**	-0.0033**
	(0.0069)	(0.0009)	(0.0097)	(0.0016)
Year950 × Affected	-0.0177*	-0.0021*	0.0008	-8.62×10^{-6}
	(0.0096)	(0.0011)	(0.0039)	(0.0006)
Year1050 × Affected	0.0051	0.0006	-0.0083	-0.0012
	(0.0039)	(0.0005)	(0.0064)	(0.0010)
Year1100 × Affected	0.0309	0.0036	-0.0700**	-0.0077
	(0.0245)	(0.0028)	(0.0310)	(0.0050)
Year1150 × Affected	0.0320	0.0036	-0.0705**	-0.0078
	(0.0245)	(0.0029)	(0.0318)	(0.0051)
Year1200 × Affected	-0.0338	-0.0044	-0.0736**	-0.0082
	(0.0254)	(0.0031)	(0.0293)	(0.0050)
Year1250 × Affected	-0.1216***	-0.0150***	-0.0829***	-0.0093*
	(0.0350)	(0.0045)	(0.0318)	(0.0054)
Year1300 × Affected	-0.1280***	-0.0131**	-0.0828***	-0.0099**
	(0.0379)	(0.0052)	(0.0294)	(0.0047)
Year1350 × Affected	-0.1408***	-0.0159***	-0.0669**	-0.0076
	(0.0414)	(0.0058)	(0.0297)	(0.0046)
Year1400 × Affected	-0.0962**	-0.0104*	-0.0682**	-0.0077*
	(0.0422)	(0.0058)	(0.0287)	(0.0046)
Year1450 × Affected	-0.0693	-0.0068	-0.0717**	-0.0082*
	(0.0466)	(0.0064)	(0.0293)	(0.0047)
Year1500 × Affected	-0.1069**	-0.0121*	-0.0511	-0.0066
	(0.0513)	(0.0071)	(0.0362)	(0.0052)

Custom standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

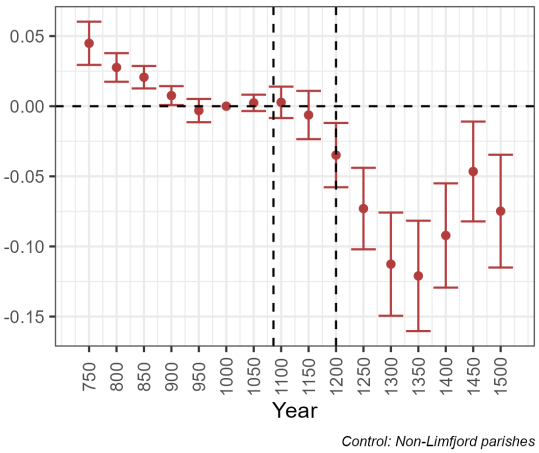
E.3 Normal distribution

Figure A5, A6, A7 and A8 show equivalent results to those presented in the paper. However, these are results based on assuming that the archaeological datings (e.g. coin finding dated to the years 1300-1495) represent a 95 percent confidence interval from a normal distribution rather than an uniform distribution. Figure A5 shows results confidence intervals for all parameters using the full sample. Figure A7 shows the same results using the matched sample. Figure A6 and A8 show all the bootstrap draws for 1350.

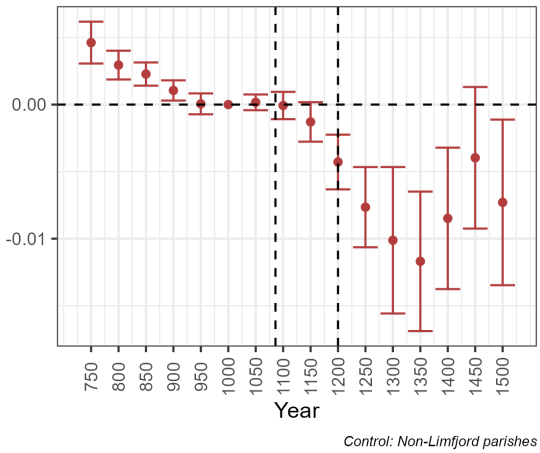
Note that the results are qualitatively the same as the main results.

Figure A5: Archaelogical results (full sample)

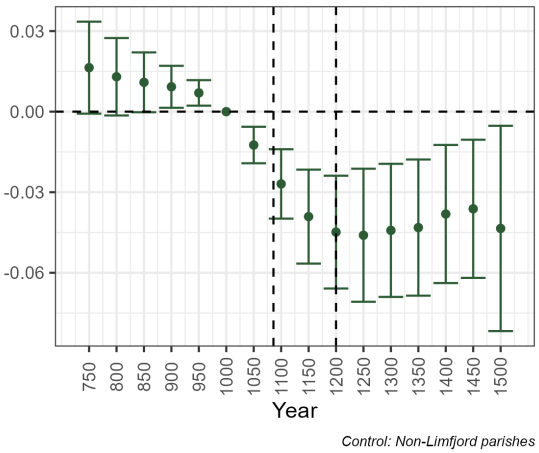
(a) Coins: Market access approach



(b) Coins: Dummy approach



(c) Buildings: Market access approach



(d) Buildings: Dummy approach

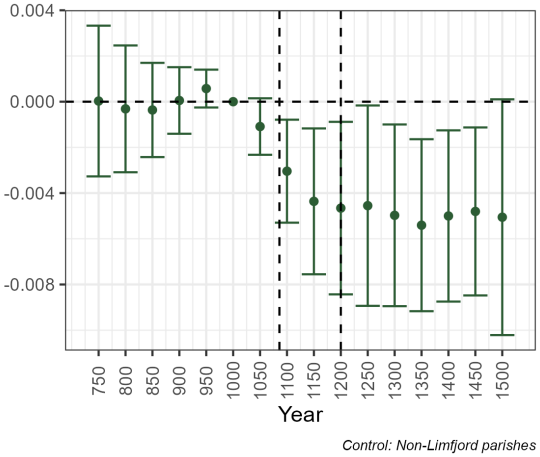
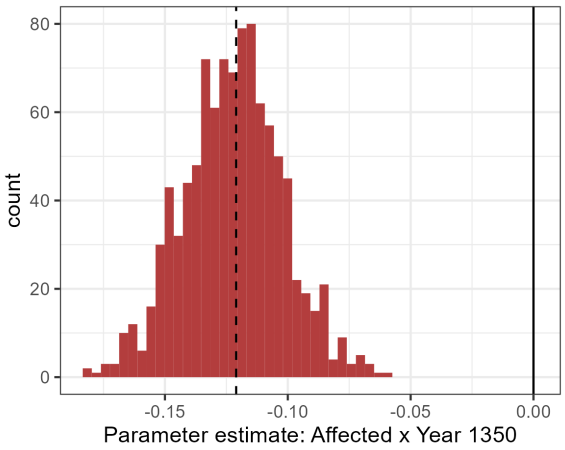
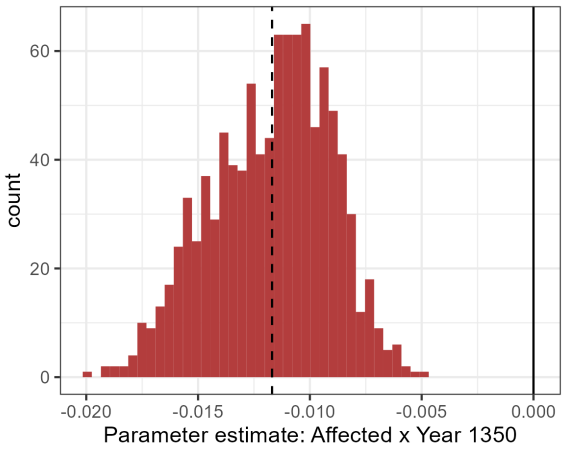


Figure A6: Distribution of parameter estimates in 1350 (full sample)

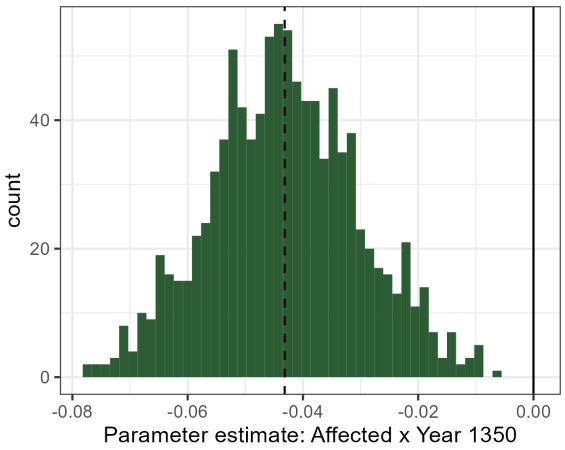
(a) Coins: Market access approach



(b) Coins: Dummy approach



(c) Buildings: Market access approach



(d) Buildings: Dummy approach

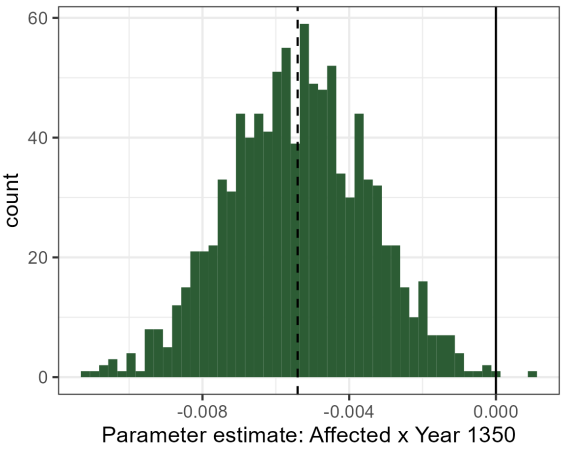
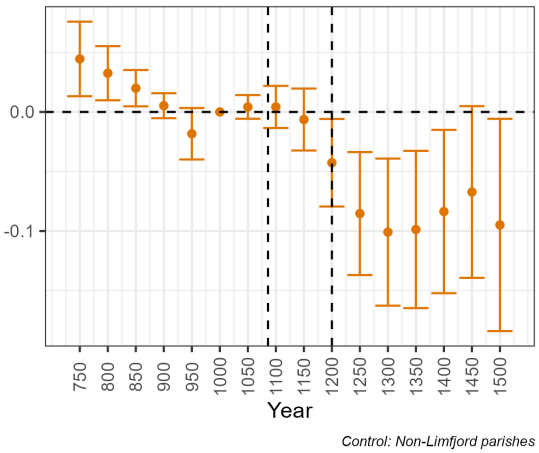
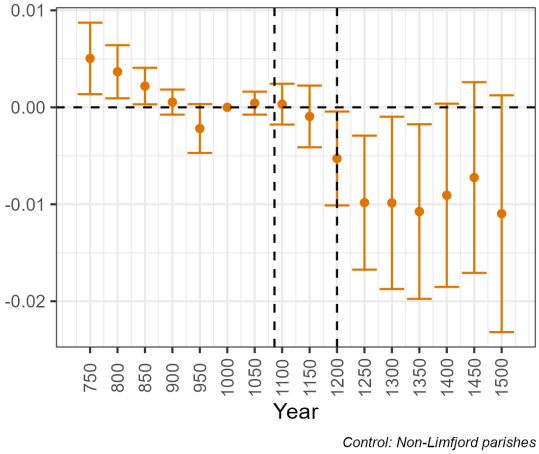


Figure A7: Archaelogical results (matched sample)

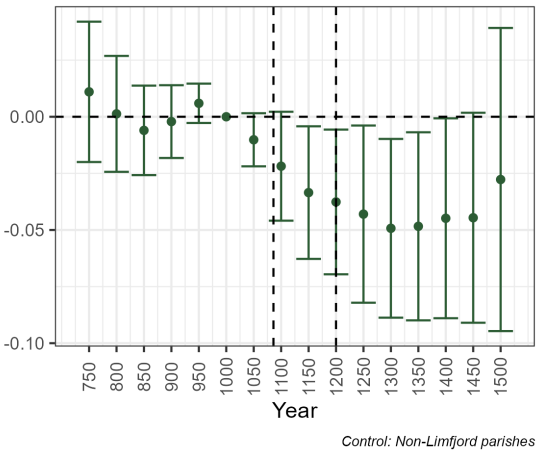
(a) Coins: Market access approach



(b) Coins: Dummy approach



(c) Buildings: Market access approach



(d) Buildings: Dummy approach

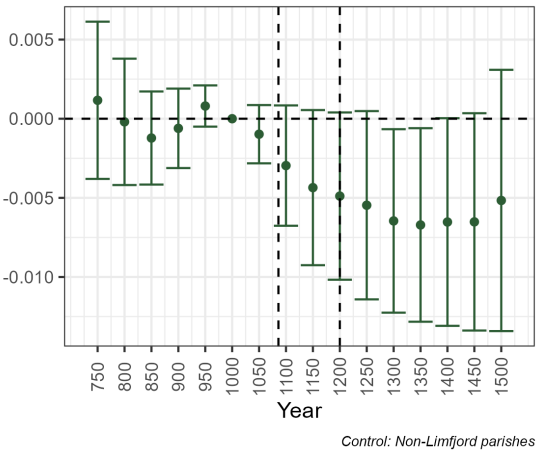
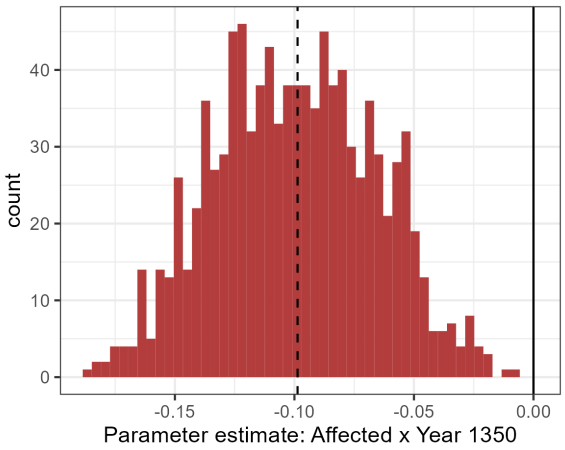
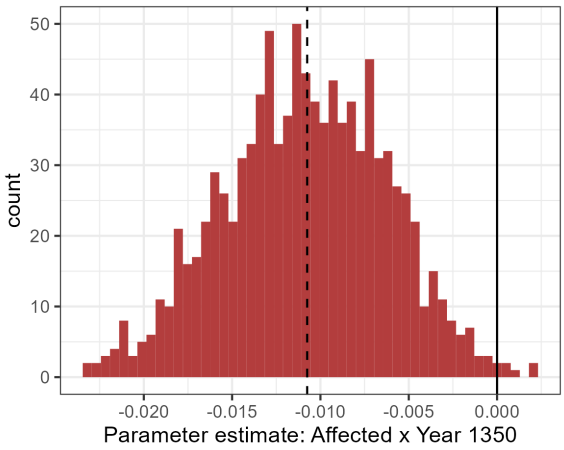


Figure A8: Distribution of parameter estimates in 1350 (matched sample)

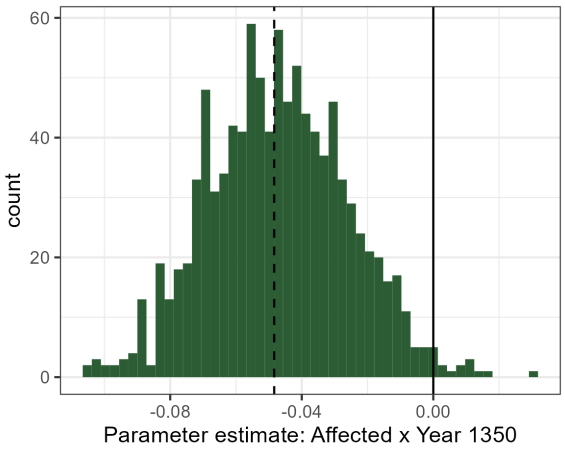
(a) Coins: Market access approach



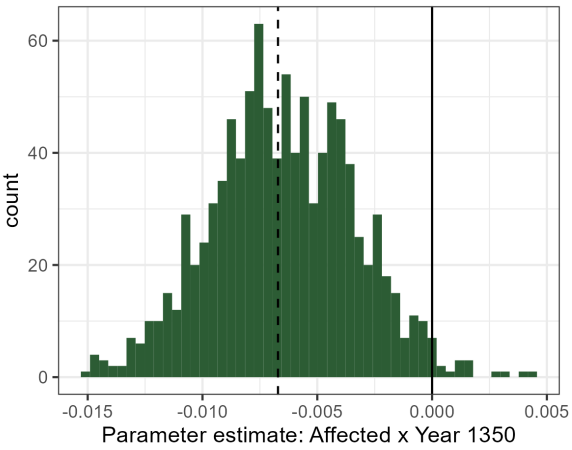
(b) Coins: Dummy approach



(c) Buildings: Market access approach



(d) Buildings: Dummy approach



European Historical Economics Society

EHES Working Paper Series

Recent EHES Working Papers

2024

EHES 261 When London Burned to Sticks: The Economic Impact of the Great Fire of 1666
Philipp Ager, Maja U. Pedersen, Paul Sharp, Xanthi Tsoukli

EHES 260 Lobbying for Industrialization: Theory and Evidence
Dmitry Veselov, Alexander Yarkin

EHES 259 How did the European Marriage Pattern persist? Social versus Familial Inheritance: England and Quebec, 1650-1850
Gregory Clark, Neil Cummins, Matthew Curtis

EHES 258 Quantifying Trade from Renaissance Merchant Letters
Fabio Gatti

EHES 257 Of the bovine ilk: Quantifying the welfare of dairy cattle in history, 1750-1900
Markus Lampe, Paul Sharp

EHES 256 Retaining population with water? Irrigation policies and depopulation in Spain over the long term
Ignacio Cazcarro, Miguel Martín-Retortillo, Guillermo Rodríguez-López, Ana Serrano, Javier Silvestre

EHES 255 Breaking the HISCO Barrier: Automatic Occupational Standardization with OccCANINE
Christian Møller Dahl, Torben Johansen, Christian Vedel

EHES 254 Birth Order and Social Outcomes, England, 1680-2024
Gregory Clark, Neil Cummins

EHES 253 Assimilate for God: The Impact of Religious Divisions on Danish American Communities
Jeanet Sinding Bentzen, Nina Boberg-Fazlić, Paul Sharp, Christian Volmar Skovsgaard, Christian Vedel

EHES 252 Where are the Female Composers? Evidence on the Extent and Causes of Gender Inequality in Music History
Karol Jan Borowiecki, Martin Hørlyk Kristensen, Marc T. Law

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